Inference for MRS models of electricity spot prices

Rafał Weron

Joint work with Joanna Janczura, Jakub Nowotarski, Jakub Tomczyk (Wrocław), Stefan Trück (Sydney) and Rodney Wolff (Brisbane)

Institute of Organization and Management

Wrocław University of Technology

Introduction

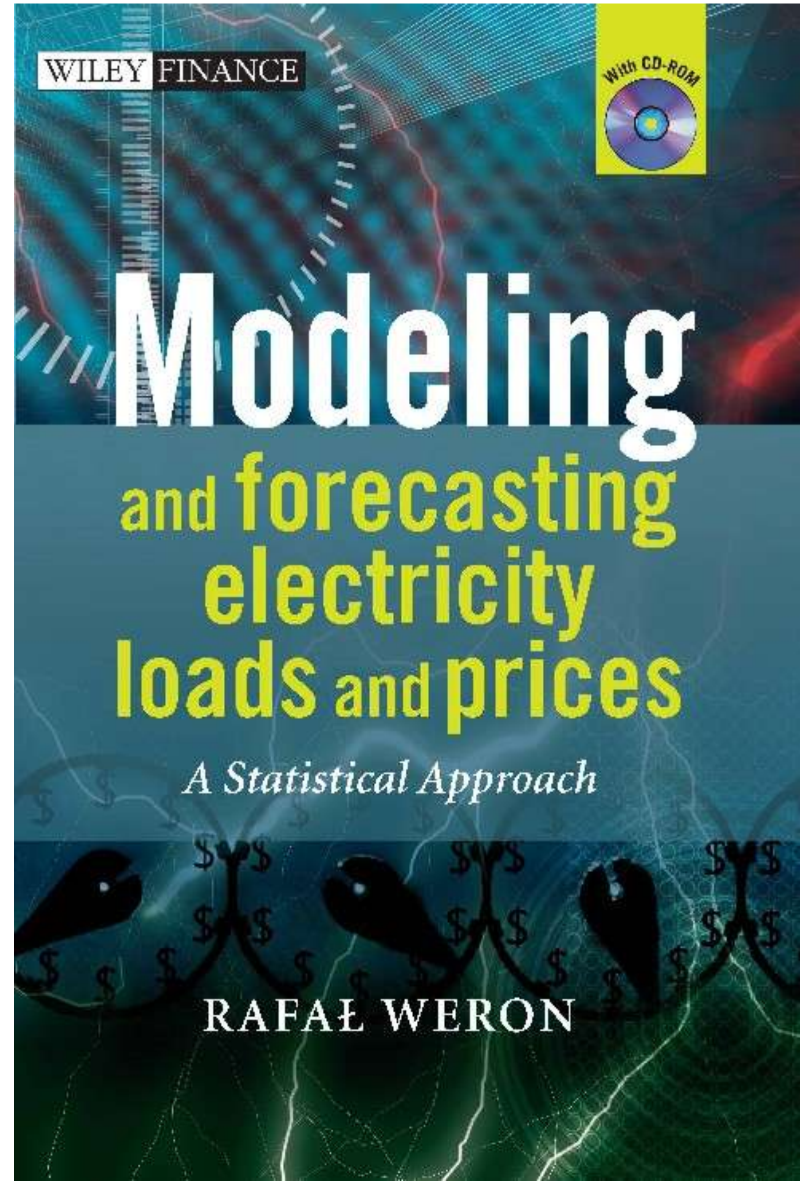
- What is the question ?

  How can we model wholesale electricity spot prices ?

- Why do we care ?

Why do we care?

- Because we can write books about it !



(More serious) motivation

- Risk management and derivatives pricing often require a model for spot prices that is:
- Realistic
  - Why would we want an unrealistic model?!
- Parsimonious
  - Faster simulation, smaller calibration errors
- Statistically sound
  - We can calibrate any model to any dataset …
  - … but does it **really** fit the data? Does it **make sense**?

Agenda

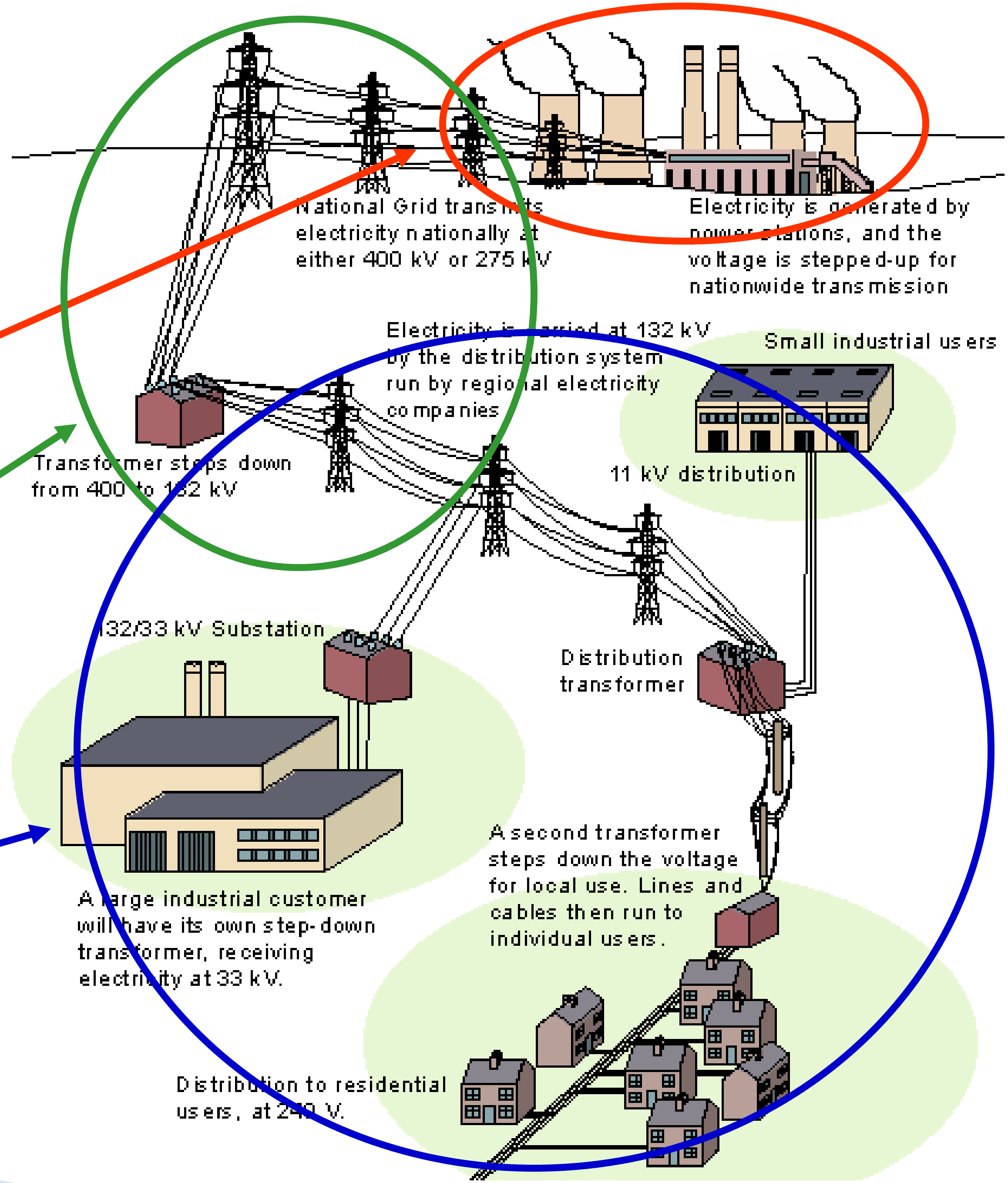
- **Power markets in a nutshell**
  - **Market structure**
  - **The spot**
  - **Electricity is a (special) commodity**
- Overview of modeling approaches
- Second generation MRS models

The power system

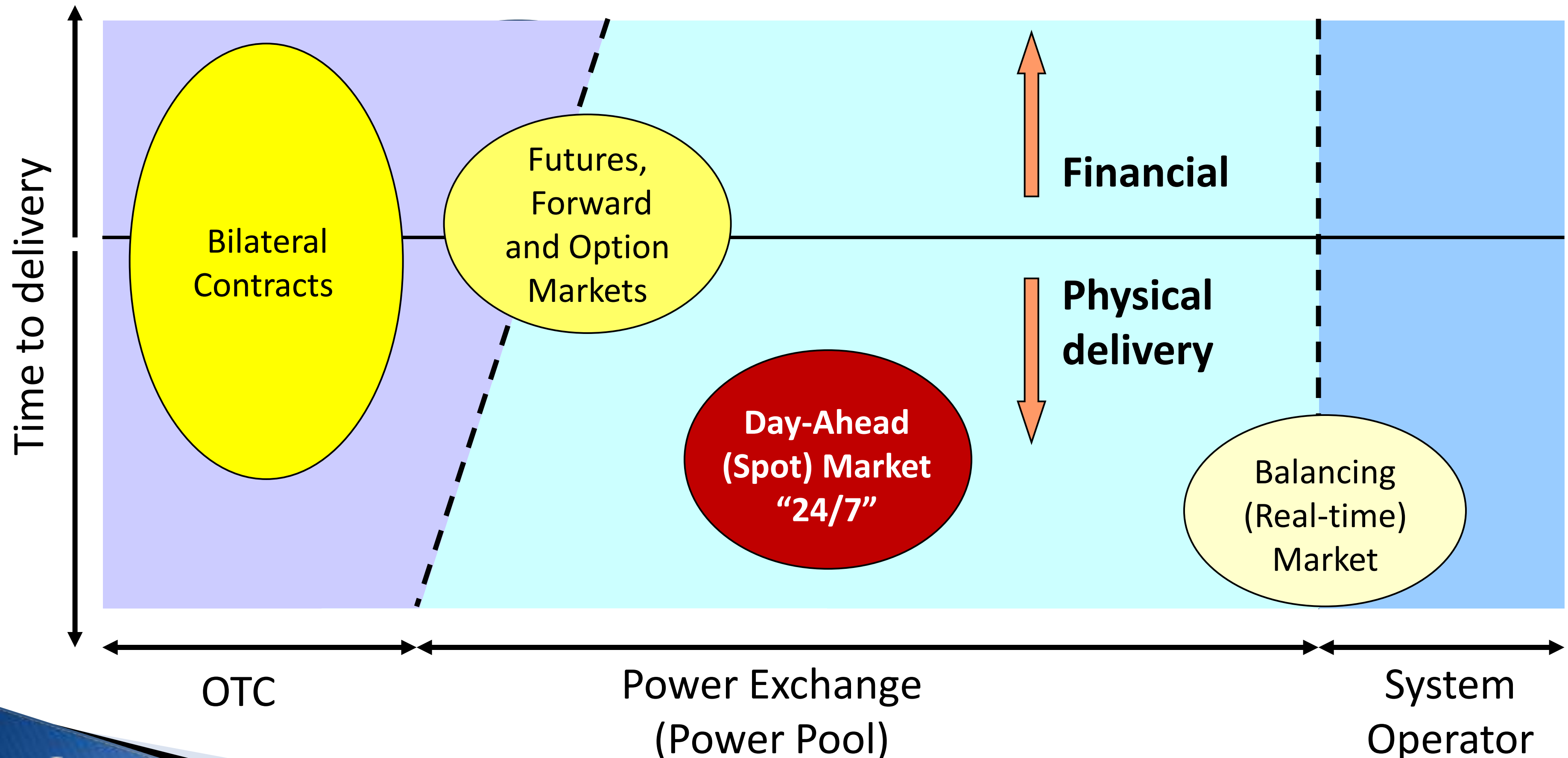
Generation

Transmission

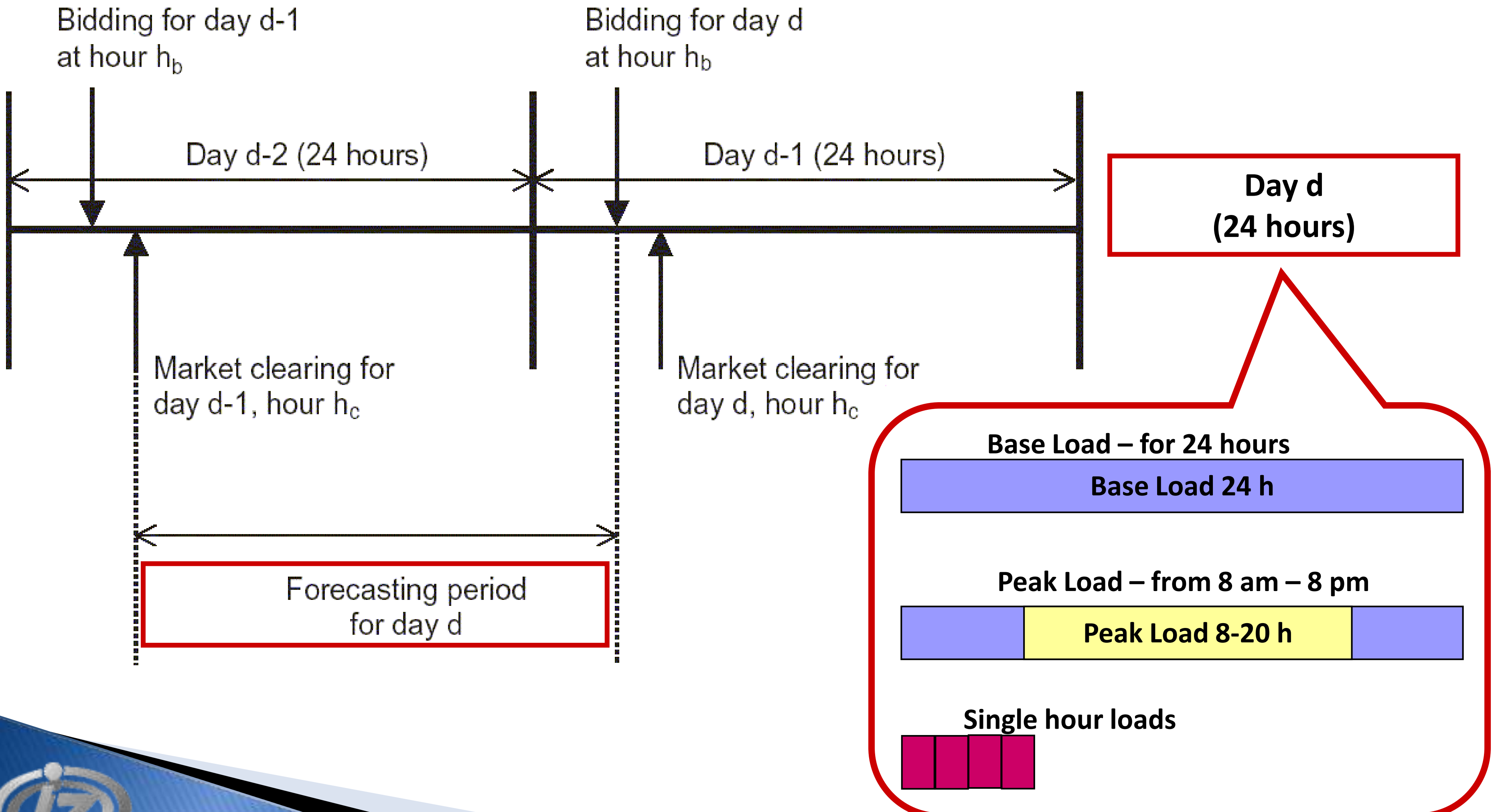
Retail



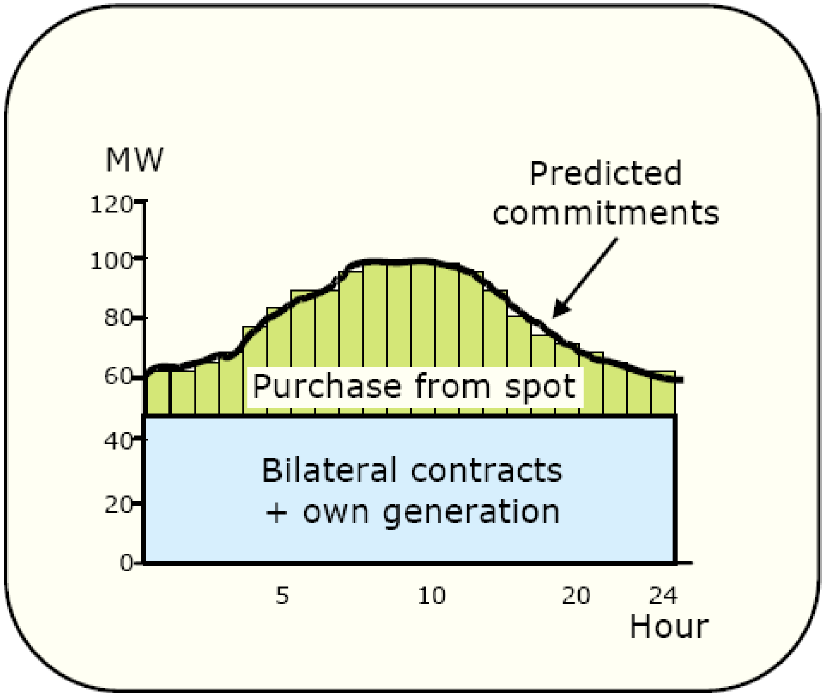
Wholesale electricity market structure

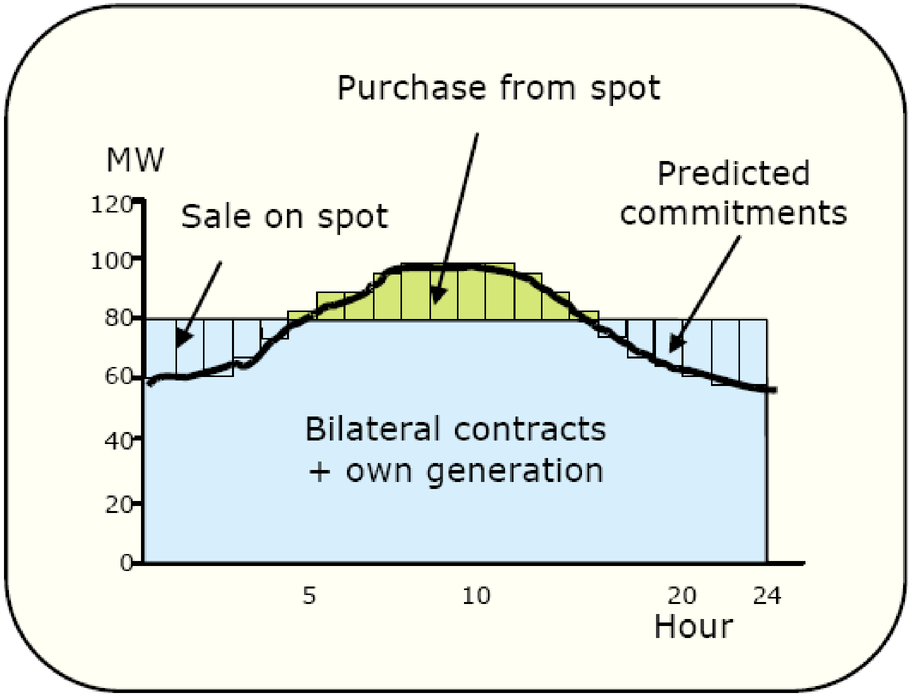


The spot

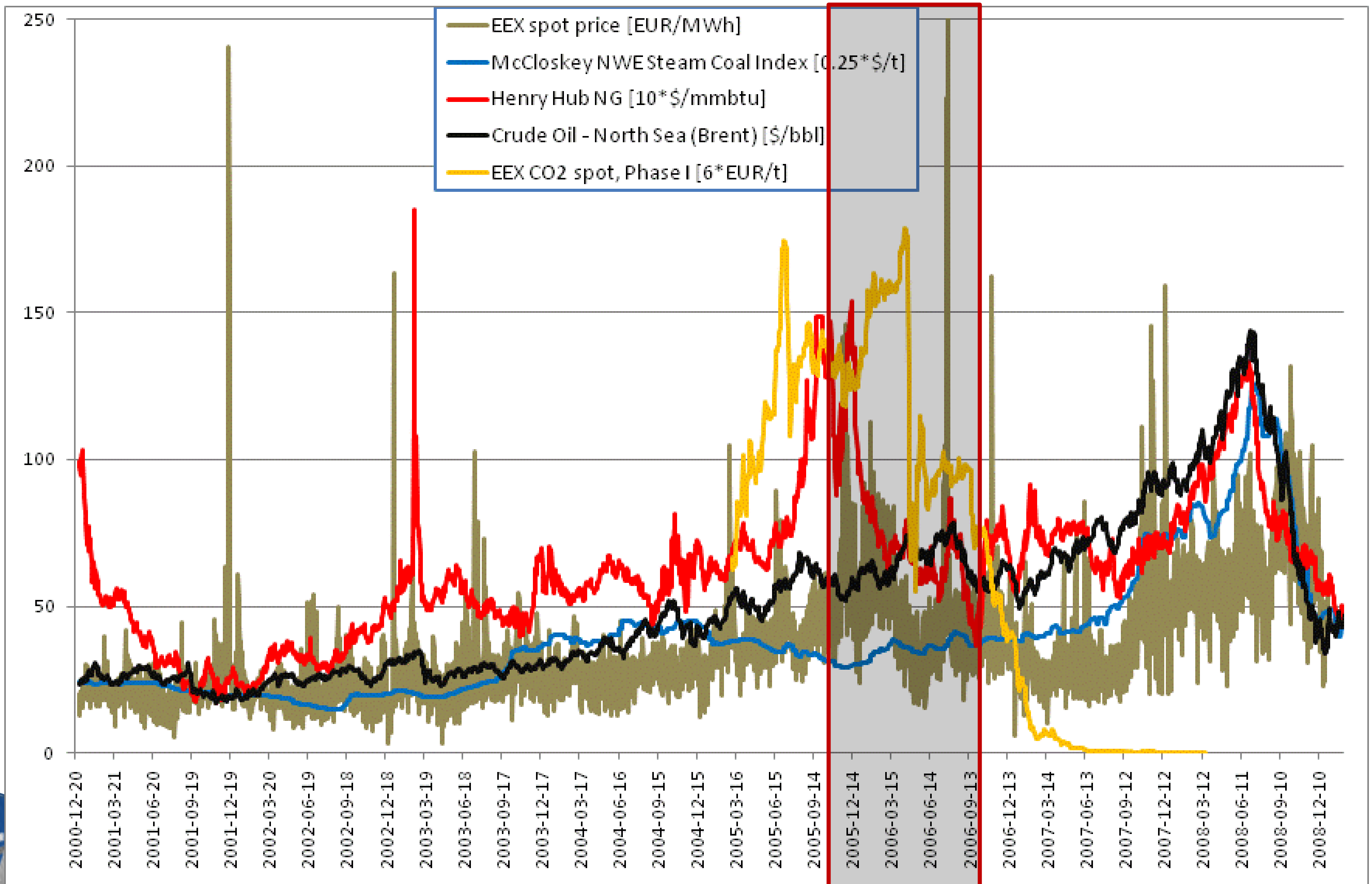

Bidding for day d-1
at hour h_b
Bidding for day d
at hour h_b
Day d-2 (24 hours)
Day d-1 (24 hours)
Day d
(24 hours)
Market clearing for
day d-1, hour h_c
Market clearing for
day d, hour h_c
Forecasting period
for day d
Base Load – for 24 hours
Base Load 24 h
Peak Load – from 8 am – 8 pm
Peak Load 8-20 h
Single hour loads

Use of the spot market

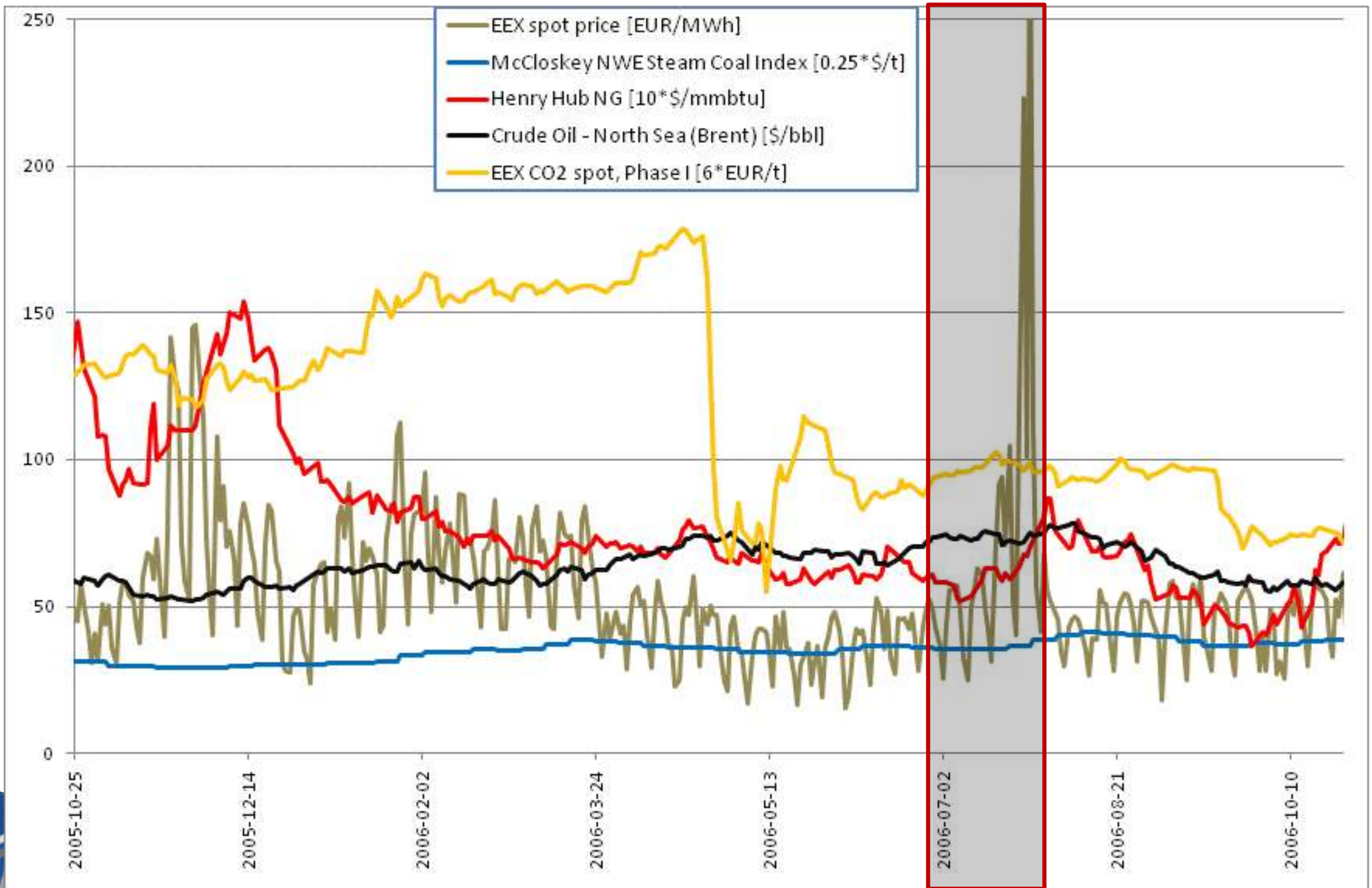




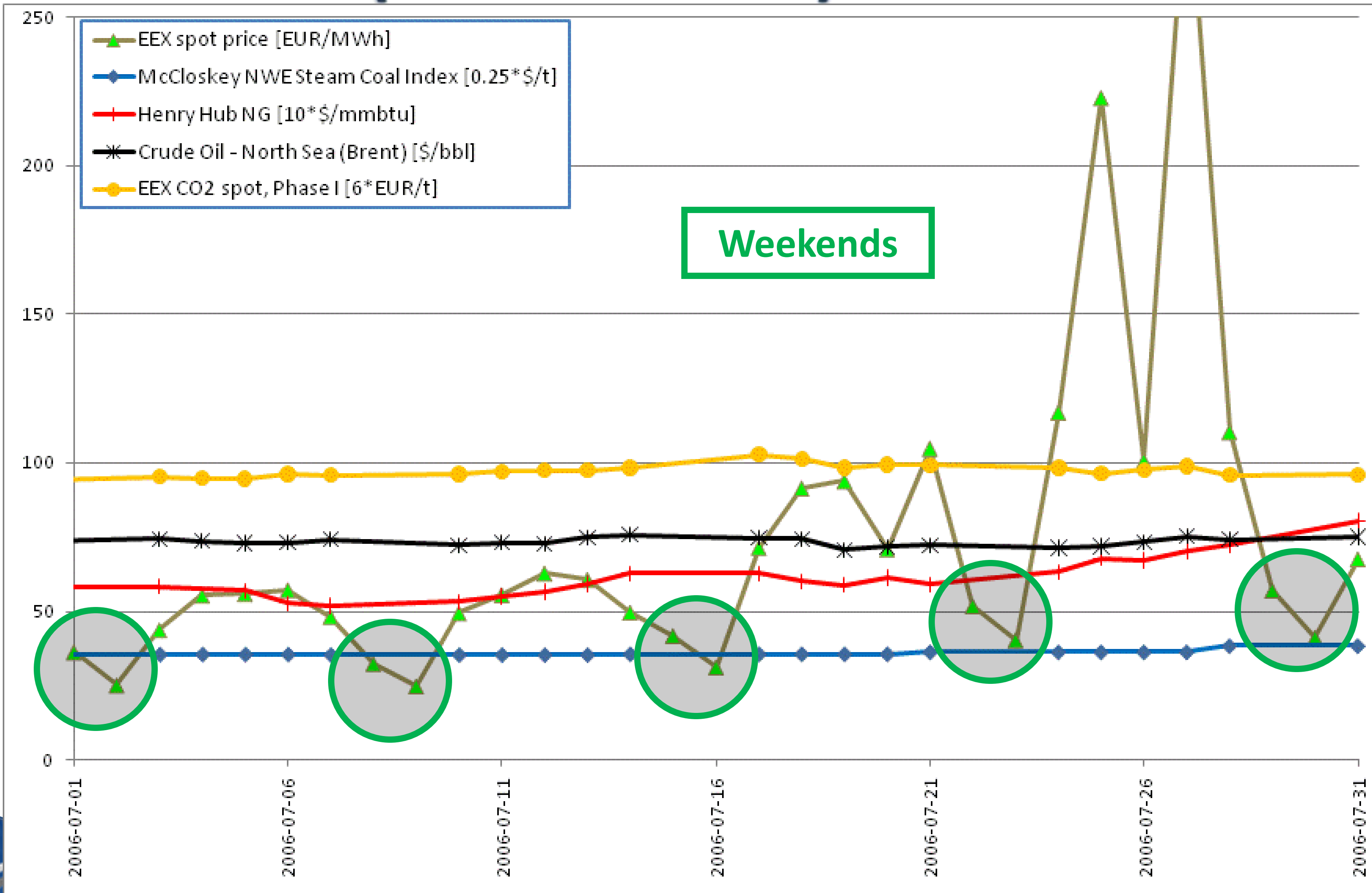
Electricity is a (special) commodity



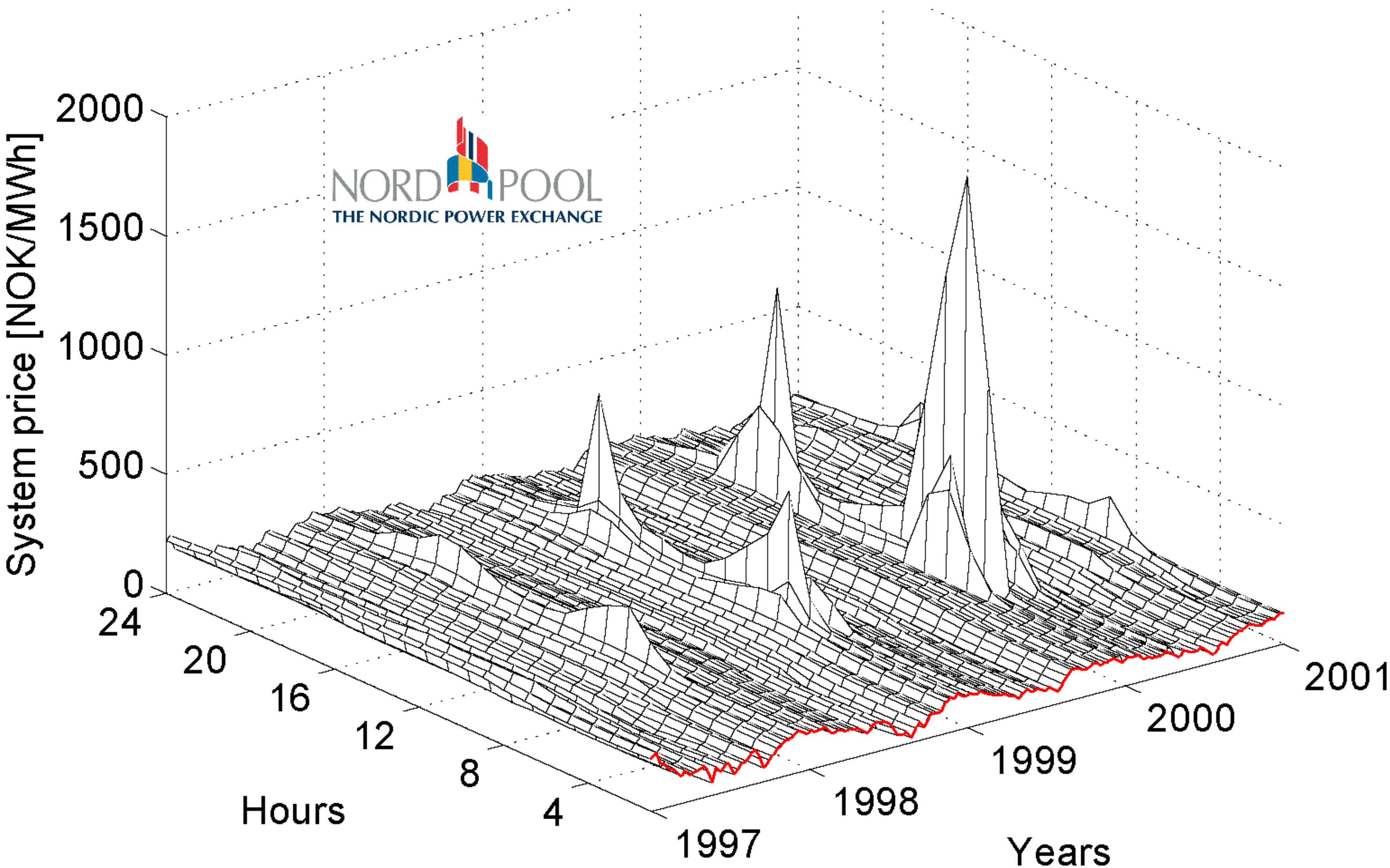
Zoom in (one year)



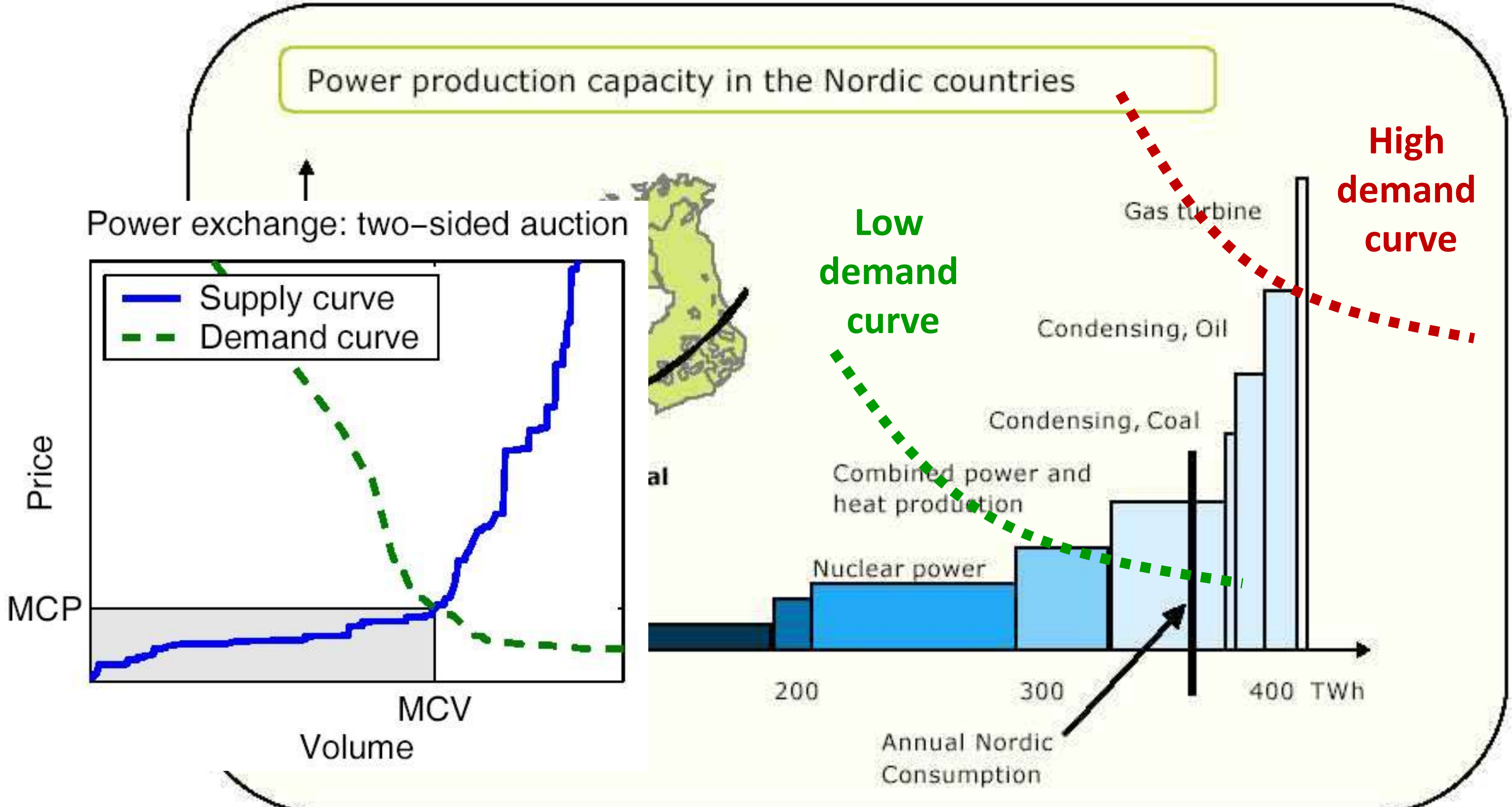
Zoom in (one month)



The seasonalities and spikes



Supply stack and regime-switching



Even mean daily prices can be spiky … and negative!

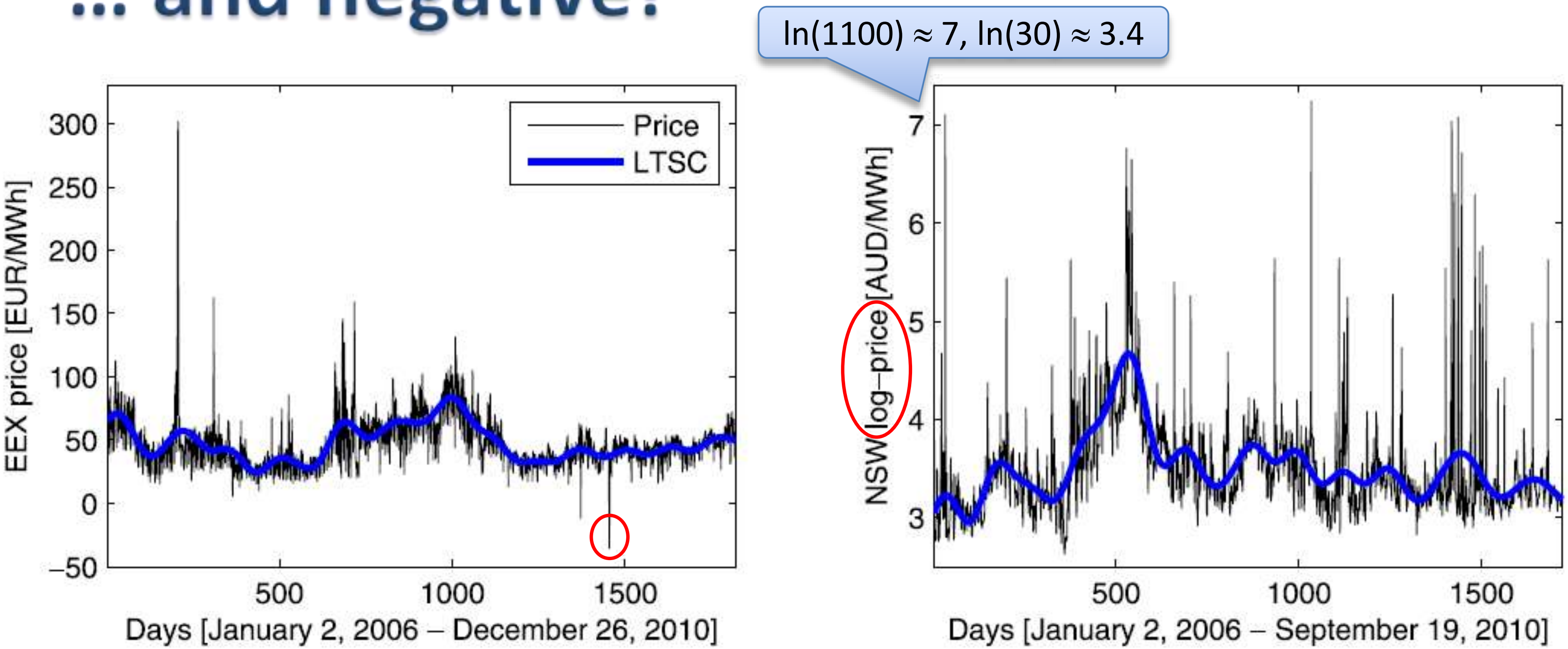
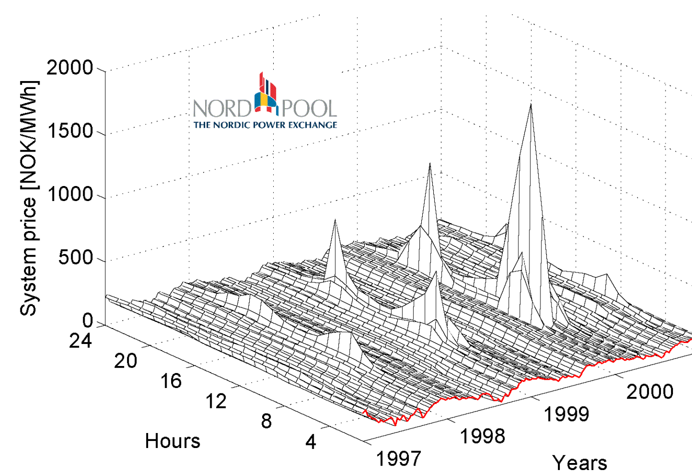


Fig. 6 (Color online) Mean daily spot EEX prices (*left panel*) and NSW log-prices (*right panel*) and the estimated long-term seasonal components (LTSC; *thick blue lines*)

Electricity is a (special) commodity

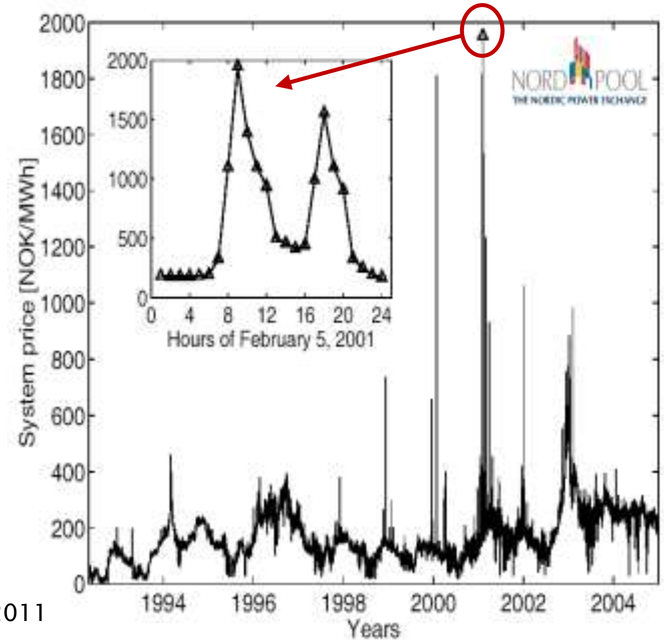
- Limited storability and transmission constraints
- Weather dependency and seasonality (daily, weekly, annual)
- Spikes in prices and loads (consumption)
  - Extreme volatility, up to 50% for daily returns
- *Inverse leverage effect*
  - Prices and volatility are positively correlated; both are negatively related to the inventory level
- *Samuelson effect*
  - Volatility of forward prices decreases with maturity



Agenda

- Power markets in a nutshell
- **Overview of modeling approaches**
  - **The need to deseasonalize**
  - **Jump-diffusions vs. Markov-regime switching**
  - **EM algorithm for MRS models**
  - **Problems with first generation MRS models**
- Second generation MRS models

Reduced form models for the spot



Black-Scholes (1973)
→ Geometric Brownian Motion (GBM)

$$dX_t = \mu X_t dt + \sigma X_t dW_t$$

Vasicek (1977), CIR (1985)
→ Mean-Reverting Diffusion (MRD)

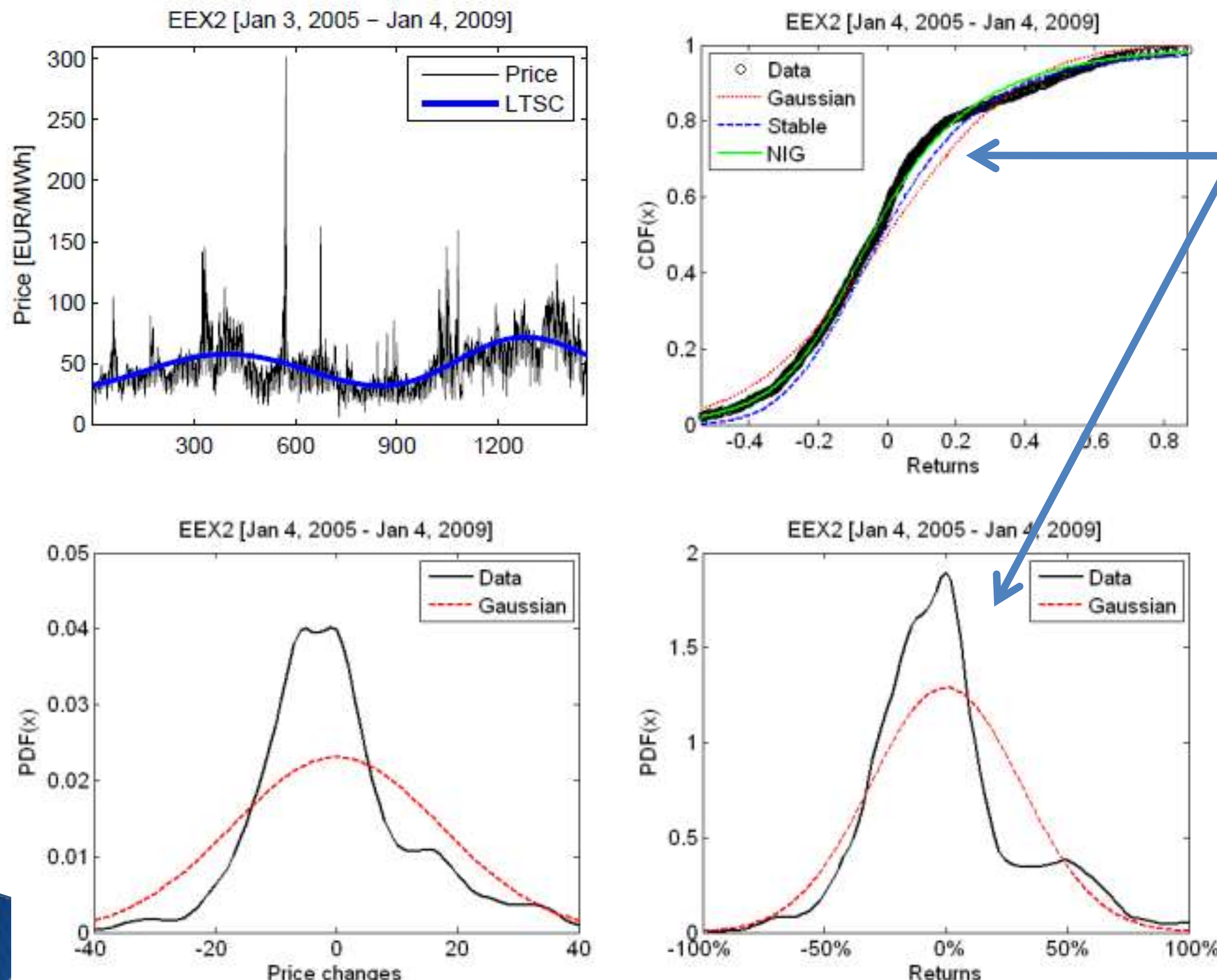
$$dX_t = (\alpha - \beta X_t)dt + \sigma X_t dW_t$$

???

$$dX_t = (\alpha - \beta X_t)dt + \sigma X_t dW_t$$

+ price spikes

The need to deseasonalize …



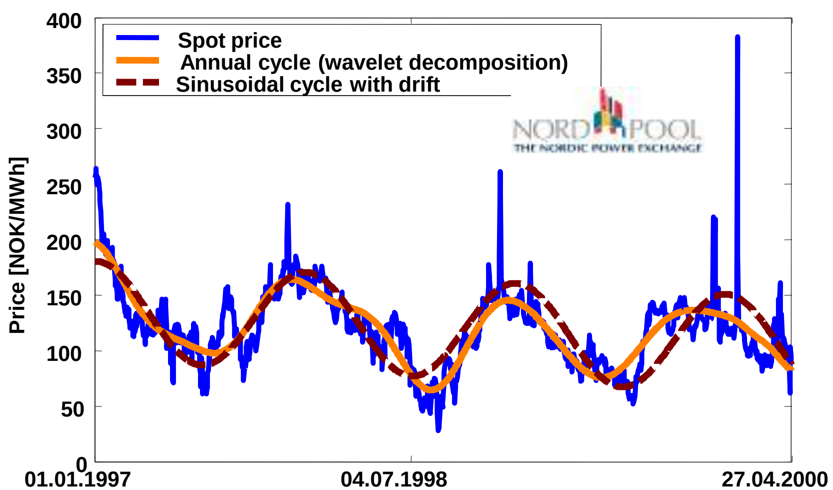
Skewness due to the weekdays-weekend cycle

Solution:
fit short-term seasonal component (STSC)

e.g. by finding the *mean week*

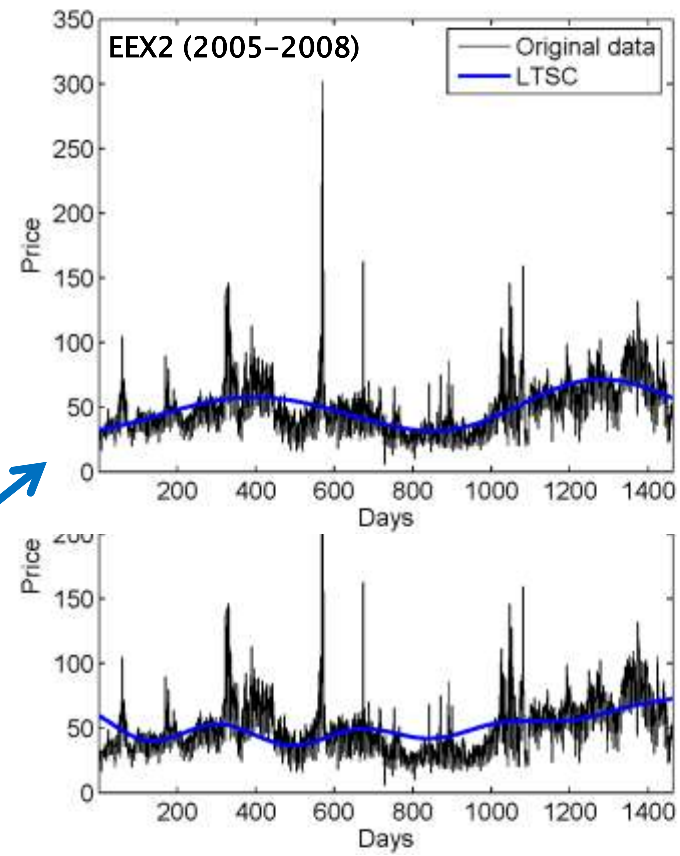
Long-term seasonality (LTSC)

- Fitting piecewise constant functions (dummy variables) for each month
  - Bhanot (2000), Haldrup & Nielsen (2006), Knittel & Roberts (2005), Lucia & Schwartz (2002)
  - For each day of the week → corresponds to mean week or MA method
- Fitting (a sum of) sinusoids with trend
  - Bierbrauer et al. (2007), Borovkova & Permana (2006), Cartea & Figueroa (2005), De Jong (2006), Geman & Roncoroni (2006), Lucia & Schwartz (2002), Pilipovic (1997), Weron (2006)
- Wavelet smoothing
  - Weron et al. (2004), Trück et al. (2007), Weron (2008), Janczura & Weron (2009,2010)

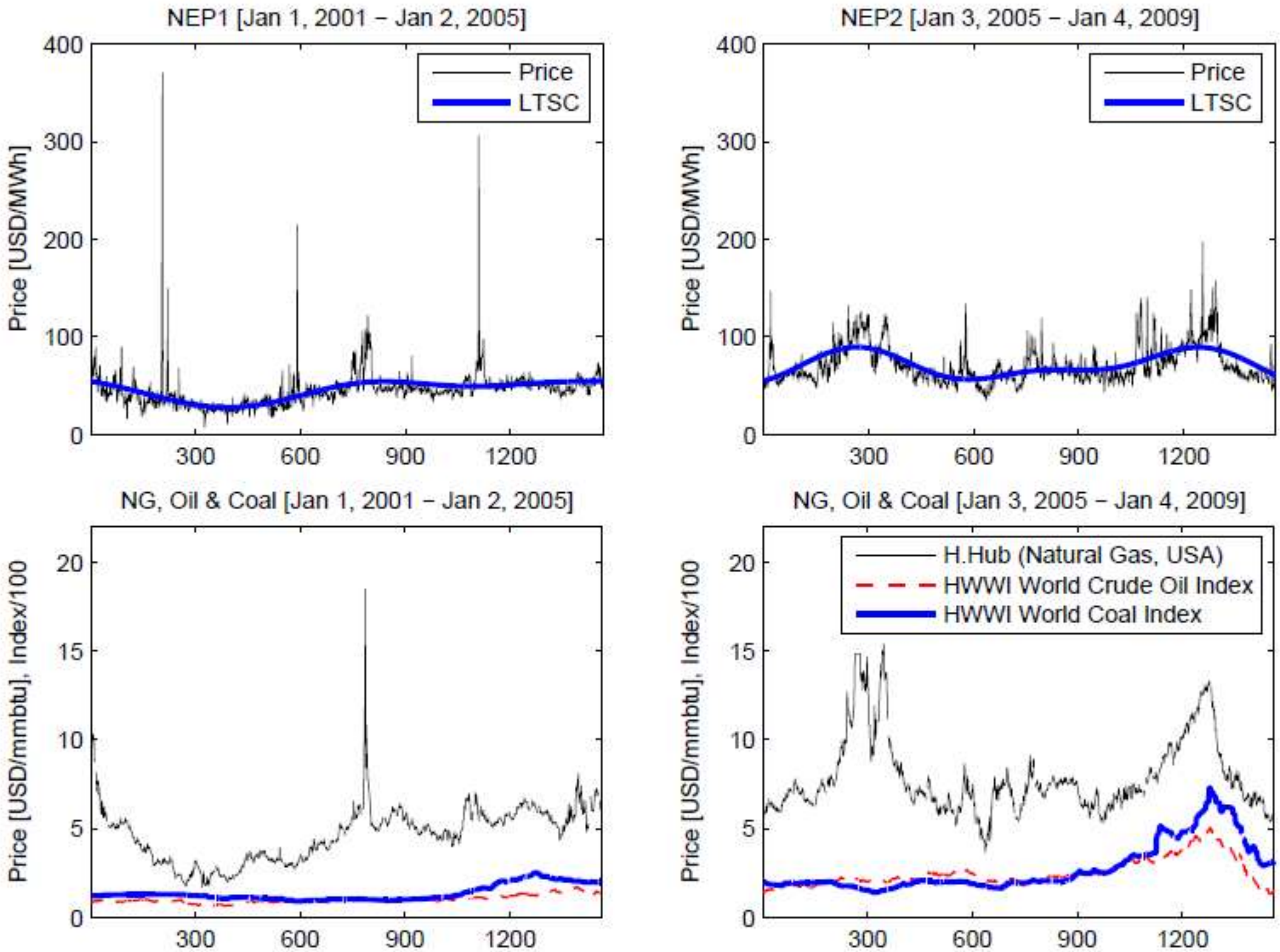


Wavelet smoothing

- Decompose: $x(t) = S_J + D_J + D_{J-1} + \dots + D_1$, with $2^J <$ #obs
- At the coarsest scale the signal can be estimated by S_J
    - By adding a mother wavelet D_j of a lower scale $j = J-1, J-2, \dots$, we obtain a better estimate of the original signal → **lowpass filtering**
- For daily data the S_3, S_5 and S_8 approximations roughly correspond to
    - weekly ($2^3 = 8$ days),
    - monthly ($2^5 = 32$ days) and
    - **annual ($2^8 = 256$ days) smoothing**



Rationale for the wavelet smoother



Jump-diffusion (JD) models

- First modeling attempts
  - Clewlow & Strickland (2000), Eydeland & Geman (2000), Kaminski (1999)
  - The deseasonalized spot electricity price X_t was assumed to follow some kind of a jump-diffusion (JD) process:

$$dX = \mu(X,t)dt + \sigma(X,t)dB + dq(X,t)$$

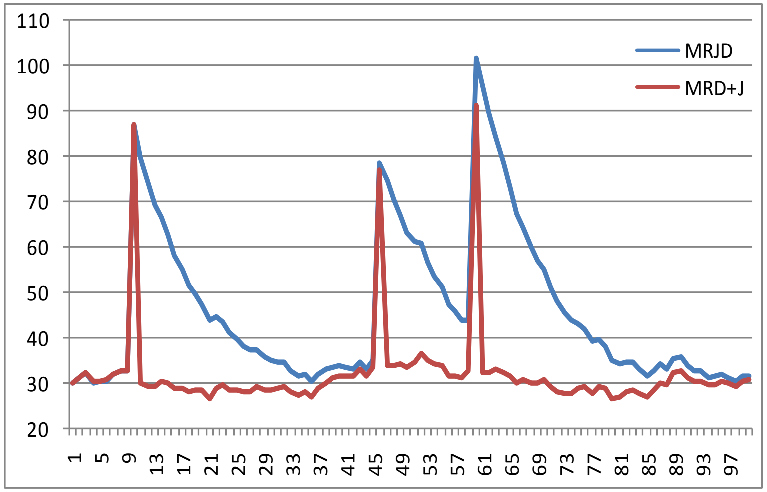
Generalized drift (possibly mean reverting)

Generalized volatility (possibly heteroskedastic)

Pure jump process with given intensity and severity – e.g. a compound Poisson process: $dq(X,t) = XdN_t$

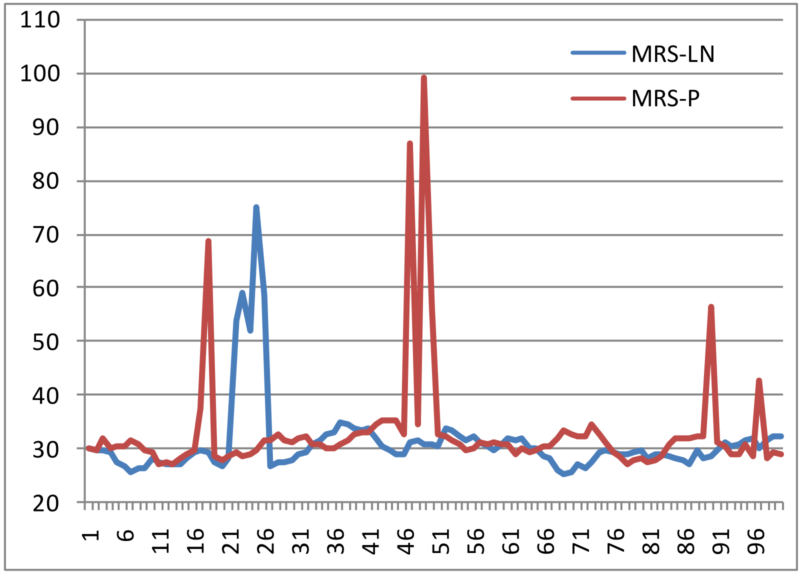
Problems with JD models

- After a jump the price is forced back to its normal level
  - by mean reversion (**MRJD**)
  - by mean reversion coupled with downward jumps
    - Deng (1999), Escribano et al. (2002), Geman & Roncoroni (2006)
  - by a negative jump of approximately the same size – on the daily scale (**MRD+J**)
    - Weron et al. (2004), Weron (2008)
  - by a combination of mean reversions with different rates
    - Benth et al. (2007), Benth et al. (2008)



Problems with JD models cont.

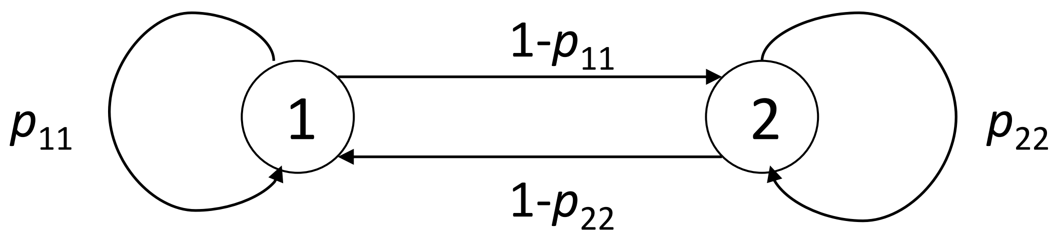
- What about periods of consecutive jumps?
  - Grid congestion, outage
- Solution:
  - Allow the process to *stay* in the *jump regime* with some probability
  - Regime-switching models:
    - **Markov regime-switching (MRS)**
    - Threshold autoregressions (TAR, SETAR, STAR)



Markov regime switching models

- The switching mechanism is driven by a latent random variable that follows a Markov chain with two (or more) possible states
  - The regimes are only latent, not directly observable
  - Estimation via the EM (expectation–maximization) algorithm

A two-state regime model: $X_t = \{1,2\}$



Transition probabilities:

$$\mathbf{P} = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} 1 - p_{12} & p_{12} \\ p_{21} & 1 - p_{21} \end{pmatrix}$$

First generation MRS models

For electricity spot log-prices

- Ethier & Mount (1998) proposed a model with 2 regimes governed by AR(1) processes
  - Conclusion: strong support for the existence of different means and variances in the two regimes
  - **Parameter-switching (MRS) model**

$$X_t = \alpha_{R_t} + (1 - \beta_{R_t})X_{t-1} + \sigma_{R_t}\epsilon_t$$

    - with the same set of random innovations in both regimes
    - Note, that the current value of the process depends **on the last observation only**, no matter from which regime it originated

First generation MRS models cont.

For electricity spot log-prices

- Huisman & de Jong (2003) proposed a 2-regime model
  - With a stable, mean-reverting AR(1) regime

    $$X_{t,i} = \alpha_i + (1-\beta_i)X_{t-1,i} + \sigma_i \epsilon_{t,i}$$

    and an **independent spike (IS)** regime – a normal variable with a higher mean and variance
  - Bierbrauer et al. (2004), Weron et al. (2004) used log-normal and Pareto distributed spike regimes
  - De Jong (2006) introduced autoregressive, Poisson driven spike regime dynamics

EM algorithm for MRS models

(Dempster et al., 1977; Hamilton, 1990; Kim, 1993)

- 'E-step': The **smoothed inferences** $P(R_t = i|\mathbf{x}_T;\theta^{(n)})$, or expectations, for the process being in regime *j* at time *t* are calculated based on some starting values $\theta^{(0)}$
- 'M-step': more exact ML estimates of η are calculated
  - Each component of the log-likelihood has to be weighted with the corresponding smoothed inference

$$\log\left[L(\eta^{(n+1)})\right] = \sum_{i=1}^{L}\sum_{t=1}^{T} P(R_t = i|\mathbf{x_T};\theta^{(n)}) \log\left[f(x_t|R_t = i, \mathbf{x_{t-1}};\eta^{(n+1)})\right]$$

$$\theta^{(n+1)} = (\eta^{(n+1)}, \mathbf{P}^{(n+1)}, \rho_i^{(n+1)})$$

- With each new vector $\theta^{(n+1)}$, the next cycle of the algorithm is started in order to reevaluate the smoothed inferences

MRS model with IS

- The values of the mean-reverting regime become latent when the process is in the other state
  - Distribution of X_t is dependent on the whole history $(x_1, \ldots, x_{t-1})$ → all possible paths of the state process should be used in the E-step
    - The number of possible state process paths is 2^T (2-regime model)

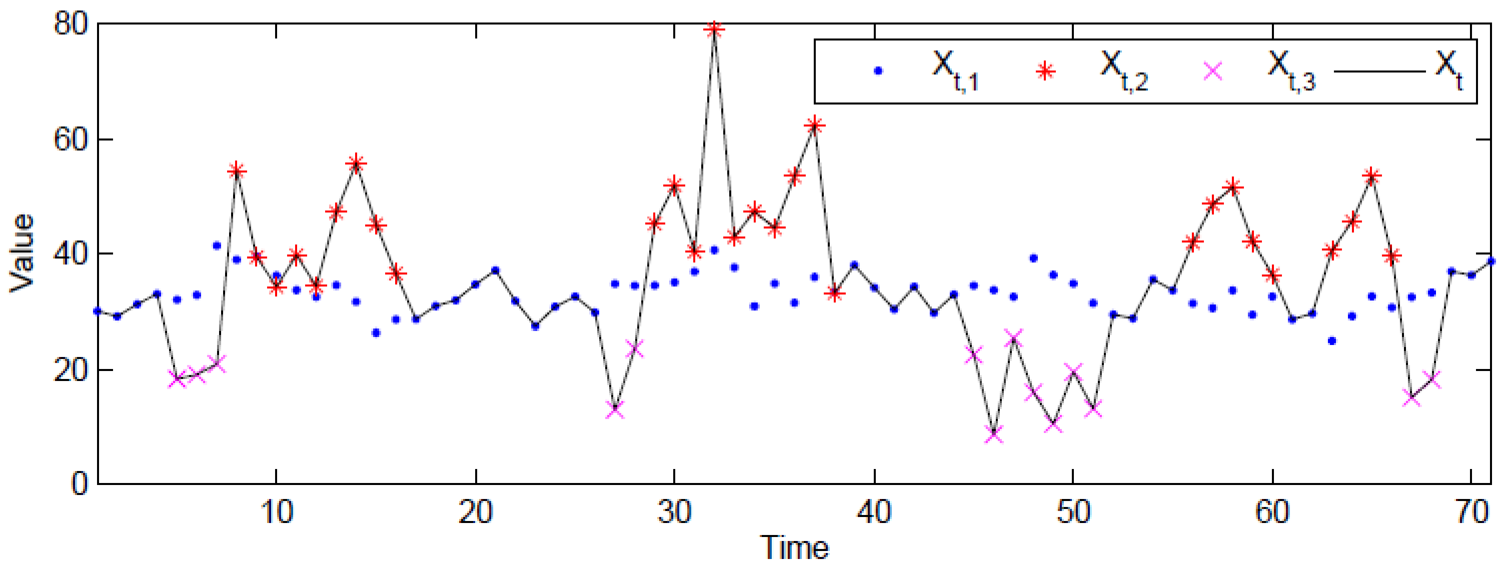


Fig. 2 A sample trajectory of a MRS model with three independent regimes (black solid line)

MRS model with IS cont.

- Huisman & de Jong (2002): use probabilities of the last 10 observations
  - Can be used only if the probability of more than 10 consecutive observations from the other regimes is negligible
- Janczura & Weron (2012): in the M-step replace the latent variables x_t from the mean-reverting regime(s) with their expectations based on the whole information available at time *t* using the following recursive formula

$$E\left(X_{t,i}|\mathbf{x}_t;\theta^{(n)}\right) = P\left(R_t = i|\mathbf{x}_t;\theta^{(n)}\right)x_t + P\left(R_t \neq i|\mathbf{x}_t;\theta^{(n)}\right)\cdot \\ \cdot\left\{\alpha_i^{(n)} + \left(1-\beta_i^{(n)}\right)E\left(X_{t-1,i}|\mathbf{x}_{t-1};\theta^{(n)}\right)\right\}$$

  - A similar approach was used by Gray (1996) to avoid the problem of conditional volatility path dependence in RS-GARCH models

MRS model with IS cont.

- Total number of probabilities (real numbers) to be stored in computer memory:
  - The EM algorithm: $2(2^{T+1}-1)$
  - Huisman & de Jong (2002): $2\{2^{10}(T-9)-1\}$
  - Janczura & Weron (2012): $4T$
    - This allows for a 100 to over 1000 times faster calibration for samples of typical size (a few thousand observations) than in case of the Huisman & de Jong algorithm

Problems

With first generation MRS models for log-prices

- Some authors reported that the '*expected spike sizes*' ($\equiv E(Y_{t,\text{spike}}) - E(Y_{t,\text{base}})$) were **negative**
  - See e.g. De Jong (2006), Bierbrauer et al. (2007)
  - … but were not considered as evidence for model misspecification
- Regime classification was not checked but …
  - … the calibration scheme generally assigns all extreme prices to the spike regime
    - The '*sudden drops*' in the log-price are not that interesting for price modeling and derivatives valuation
    - They appear extreme only because of the log transform

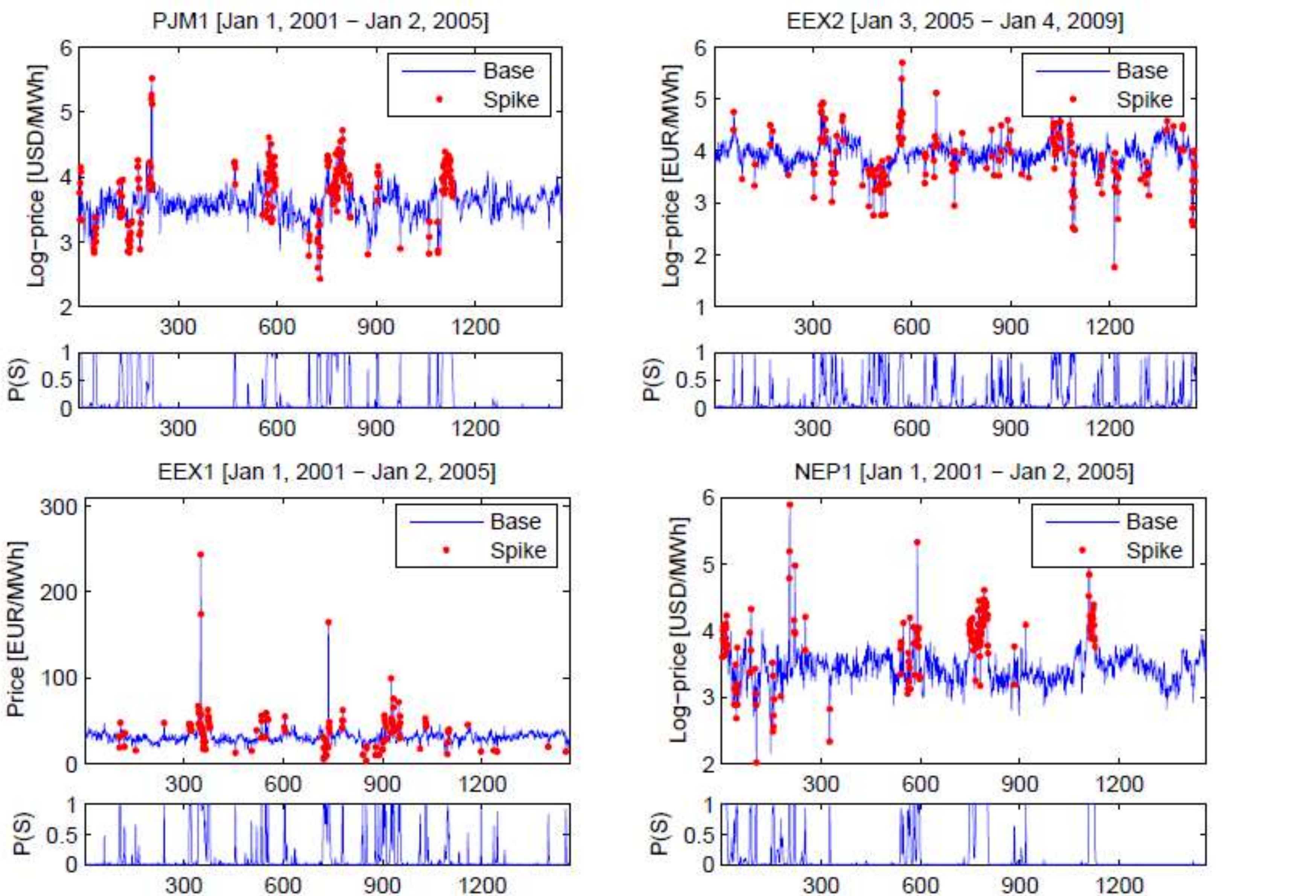


Figure 3: Sample calibration results for 2-regime models with Vasicek, i.e. AR(1), base regime dynamics and alternative spike regimes fitted to deseasonalized prices or log-prices from three major power markets. *Top left*: An independent spike (IS) model with normal spikes fitted to PJM log-prices. *Top right*: The Ethier and Mount (1998) model with AR(1) spike regime fitted to EEX log-prices. *Bottom left and right*: An IS model with lognormal spikes fitted to EEX prices and NEPOOL log-prices, respectively. The corresponding lower panels display the conditional probability $P(S) \equiv P(R_t = s|x_1, x_2, ..., x_T)$ of being in the spike regime. The prices or log-prices classified as spikes, i.e. with $P(S) > 0.5$, are additionally denoted by dots in the upper panels. For descriptions of the datasets see Section 2 and Figures 1-2.

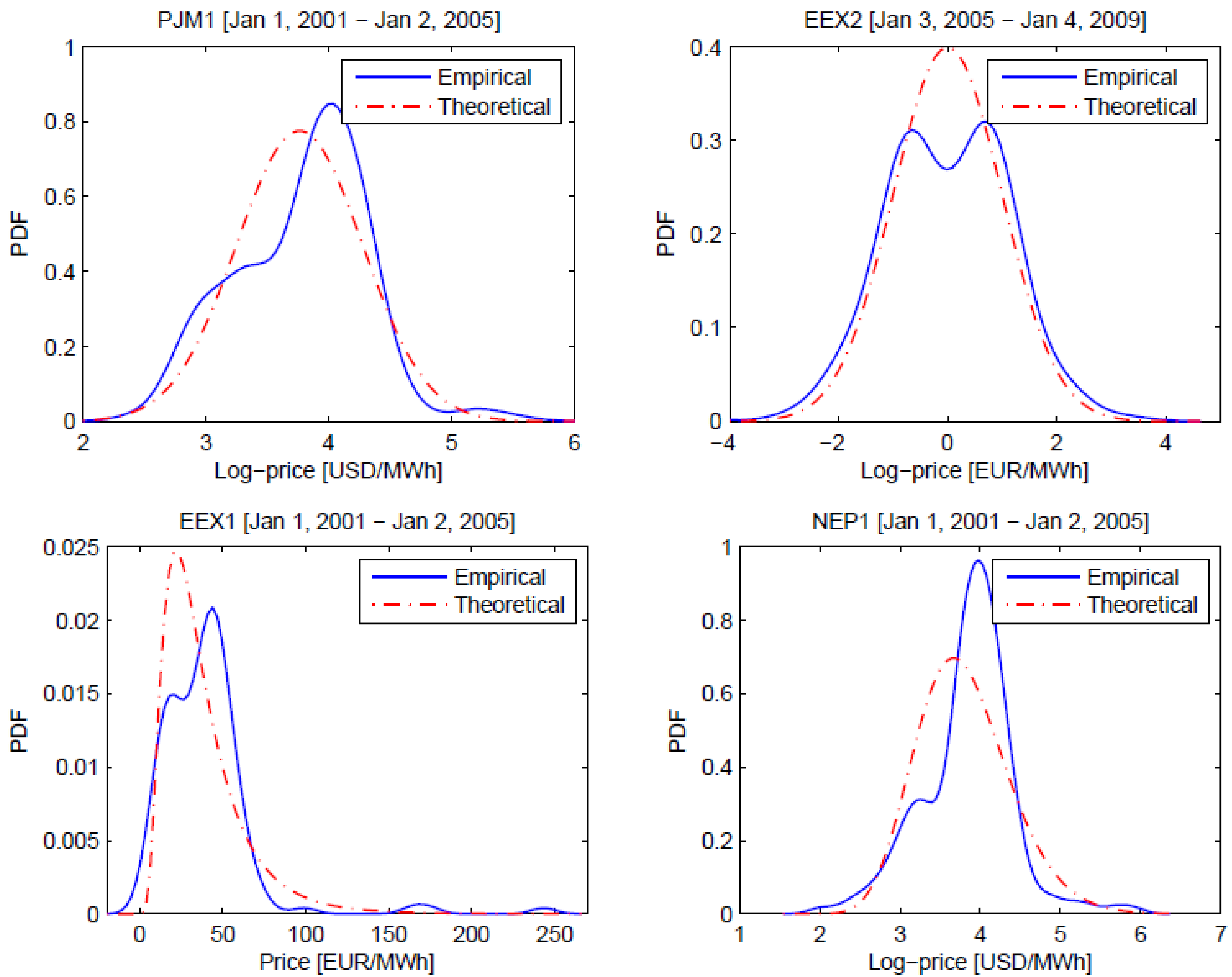


Figure 4: Comparison of empirical (sample) and theoretical (model implied) spike regime probability distribution functions in the first generation 2-regime models. The models and datasets are the same as in Figure 3. Note, that for the Ethier and Mount (1998) model the distributions of the noise in the AR(1) process driving the spike regime are plotted (*top right*).

Agenda

- Power markets in a nutshell
- Overview of modeling approaches
- **Second generation MRS models**
  - **Shifted spike distributions**
  - **Heteroskedastic base regime processes**
  - **Time-varying transition probabilities**
  - **Goodness-of-fit testing**

Second generation MRS models

For electricity spot and log-spot prices

- Fundamental extensions to improve spike occurrence:
  - Mount et al. (2006): Two AR(1)-regime model for log-prices
    - With transition probabilities dependent on the reserve margin
  - Huisman (2008) extended the IS 2-regime model for log-prices
    - Considered temperature dependent transition probabilities
- Statistical refinements to improve goodness-of-fit:
  - Weron (2009) suggested to fit prices, not log-prices
  - Janczura & Weron (2009, 2010)
    - Introduced median-shifted spike regime distributions and
    - Heteroskedastic dynamics for the base regime
    - Advocated the IS 3-regime model

Shifted spike distributions

- Perhaps spike distributions should assign zero probability to prices below a certain quantile
- Let m = median(X_t)
  - Shifted log-normal (SLN), for x > m

$$\log\left(X_{t,2} - m\right) \sim N(\mu, \sigma^2)$$

  - Shifted Pareto (SP), for x > λ ≥ m

$$X_{t,2} \sim F_P(x; \alpha, \lambda) = 1 - \left(\frac{\lambda}{x}\right)^{\alpha}$$

- Is the median cutoff optimal?
  - In general, no …

Table 1: Goodness-of-fit statistics for 2-regime models with Vasicek, see eqns. (4)-(5), base regime dynamics and median-shifted lognormal or Pareto spike distributions. Models for prices are summarized in columns 2-7, for log-prices in columns 8-13. p-values of 0.05 or more are emphasized in bold.

	Prices						Log-prices					
	Simulation			K-S test p-value			Simulation			K-S test p-value		
Data	IQR	IDR	LogL	Base	Spike	Model	IQR	IDR	LogL	Base	Spike	Model
Shifted lognormal spikes												
EEX1	9%	11%	-4193.7	0.0012	**0.4061**	0.0032	30%	30%	403.0	0.0000	**0.7371**	0.0000
EEX2	13%	3%	-5066.9	0.0090	**0.4732**	0.0149	27%	16%	399.6	0.0000	**0.6313**	0.0000
PJM1	9%	-1%	-4385.9	0.0341	**0.4346**	**0.0530**	19%	8%	777.4	0.0007	**0.3219**	0.0012
PJM2	-3%	3%	-5012.1	**0.0887**	**0.9196**	**0.0893**	3%	6%	780.1	**0.0747**	**0.4147**	**0.0696**
NEP1	2%	2%	-4327.1	0.0247	**0.5093**	**0.0561**	9%	8%	610.4	0.0002	**0.8316**	0.0003
NEP2	0%	-2%	-4665.9	**0.0823**	**0.8416**	**0.1251**	8%	0%	1417.9	0.0088	**0.7430**	0.0170
Shifted Pareto spikes												
EEX1	7%	9%	-4218.6	0.0000	0.0000	0.0000	27%	27%	436.9	0.0000	0.0123	0.0000
EEX2	10%	1%	-5101.8	0.0188	0.0008	0.0412	26%	16%	374.6	0.0000	0.0166	0.0000
PJM1	8%	-5%	-4447.2	**0.0500**	0.0000	**0.0500**	20%	6%	755.2	0.0007	0.0000	0.0012
PJM2	0%	1%	-5161.6	0.0041	0.0000	0.0007	6%	7%	744.9	0.0262	**0.3508**	0.0147
NEP1	-2	-6%	-4366.1	0.0300	0.0000	0.0300	8%	3%	546.7	0.0001	0.0000	0.0003
NEP2	13%	0%	-4703.1	0.0230	0.0000	0.0097	9%	0%	1409.7	0.0024	0.0000	0.0083

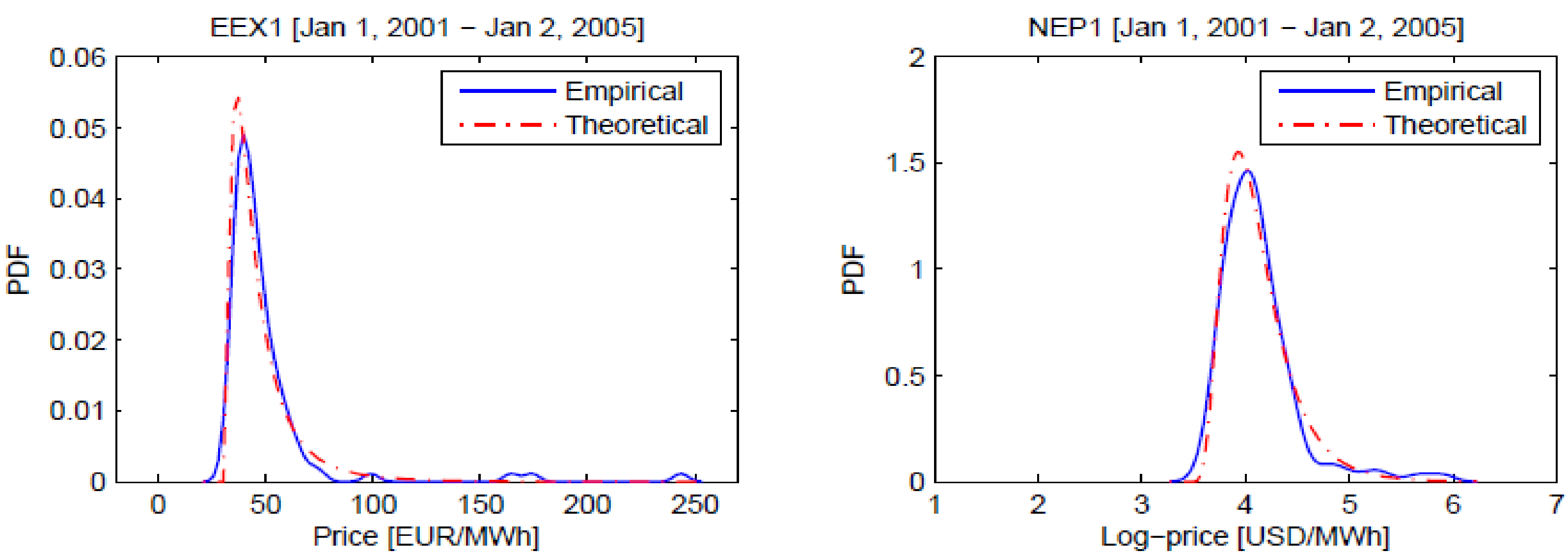


Figure 5: Comparison of empirical (sample) and theoretical (model implied) spike regime probability distribution functions in the 2-regime model with median-shifted lognormal spikes and Vasicek base regime dynamics. The fits are much better than for the models with non-shifted spike regime distributions, see Figure 4.

Case study: Optimizing the cutoffs

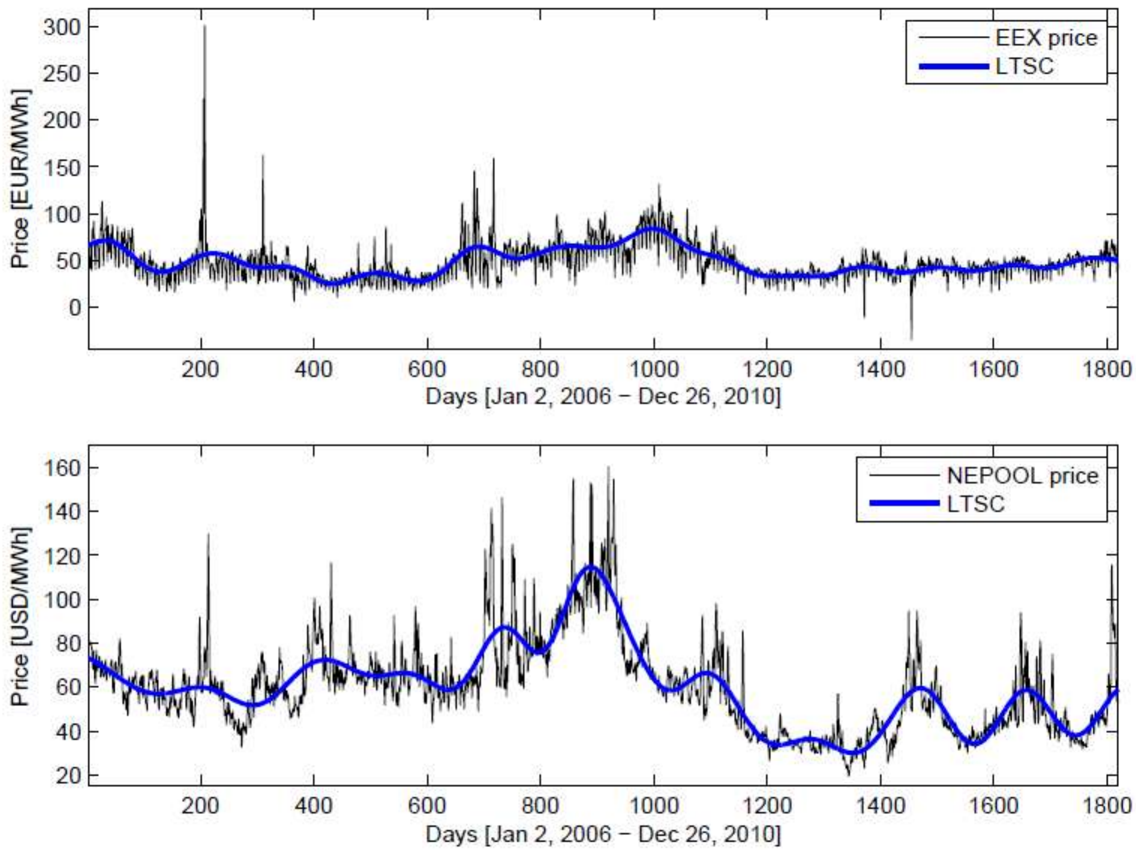
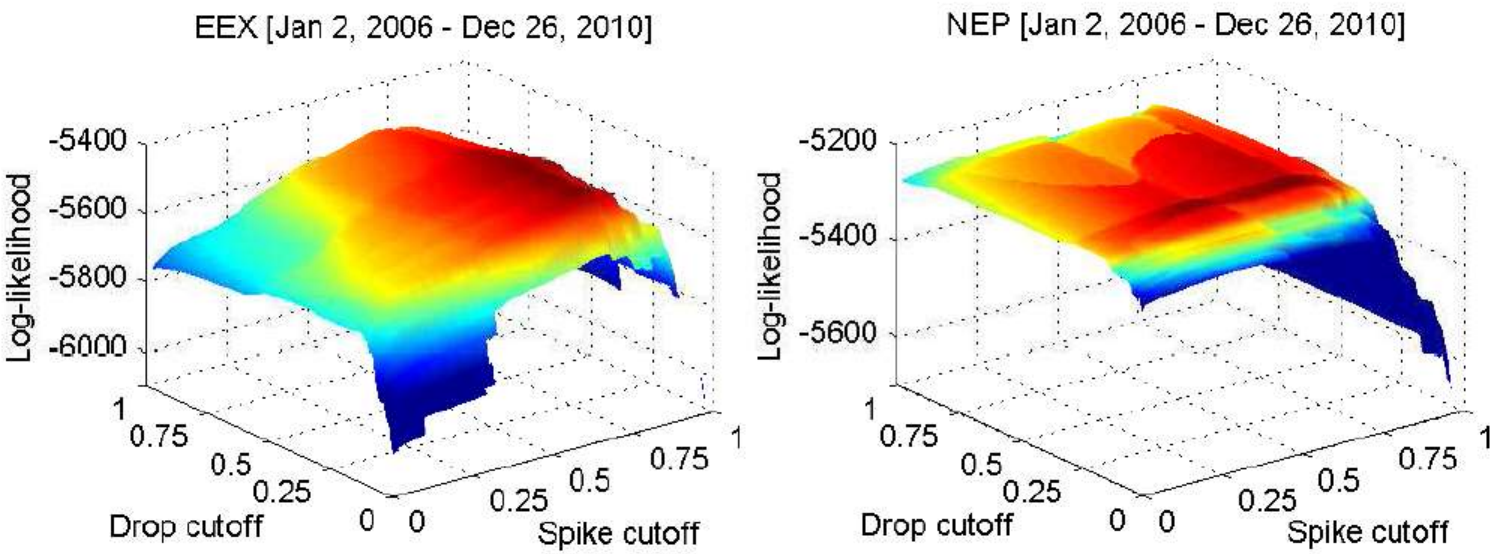


Fig. 3 Mean daily spot EEX (*top*) and NEPOOL (*bottom*) prices and the estimated long-term seasonal components (LTSC; thick blue lines).

Optimizing the cutoffs cont.

- The computational cost is not overwhelming
  - Typically <100 calibrations have to be performed before a (local) maximum is reached
    - Using the default parameters of the simplex routine in Matlab



Optimizing the cutoffs cont.

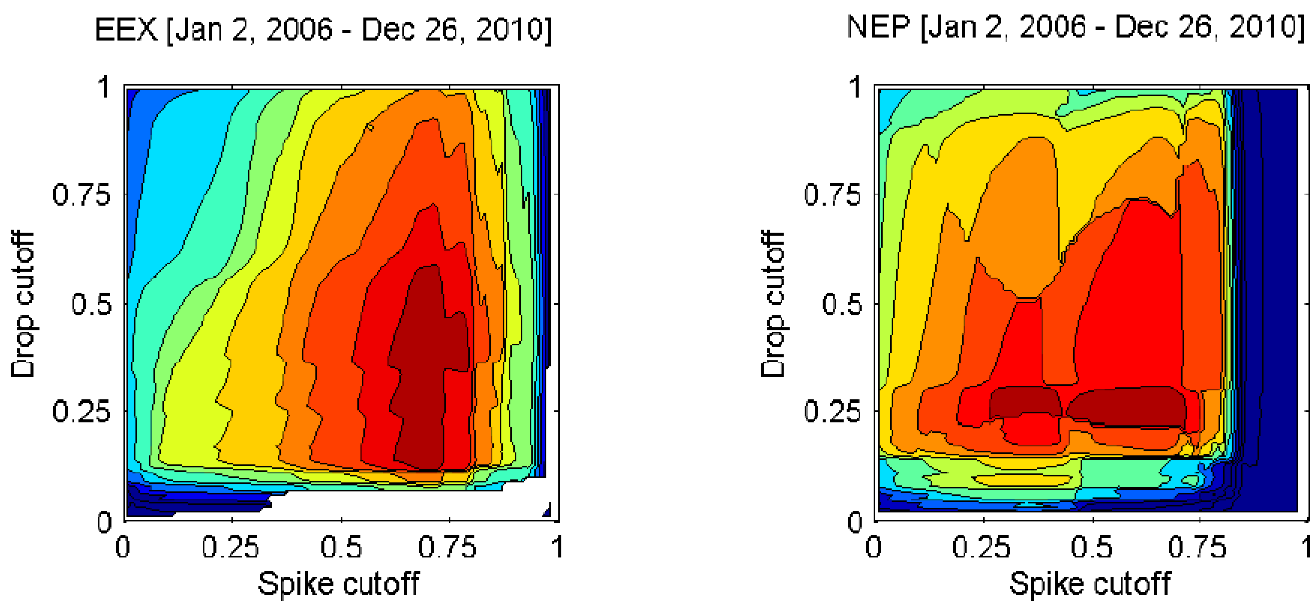


Fig. 6 Log-likelihoods of the best fitted models for each of the possible spike and drop regime cutoffs (0, 0.01, ..., 1) for EEX (*left panels*) and NEPOOL (*right panels*) deseasonalized daily

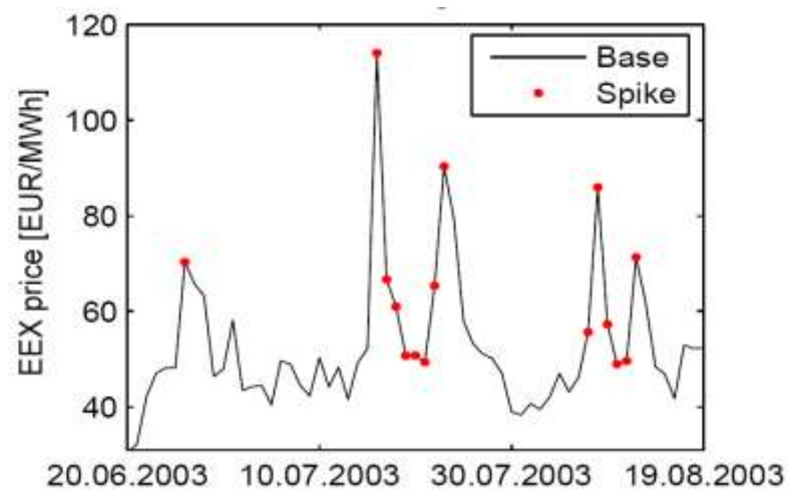
Optimizing the cutoffs cont.

Table 2 Goodness-of-fit statistics for the MRS model with three independent regimes (25)-(27) fitted to the deseasonalized EEX and NEPOOL prices. For parameter estimates see Table 1.

Calibration scenario	K-S test p-values				LogL
	Base	Spike	Drop	Model	
EEX					
optimal (0.25%, 0.69%)	0.8433	0.3106	0.9323	0.5887	-5432.17
quartiles (0.25%, 0.75%)	0.8912	0.1940	0.7898	0.4370	-5459.58
median (0.5%, 0.5%)	0.8857	0.0156	0.2767	0.1355	-5492.59
NEPOOL					
optimal (0.24%, 0.65%)	0.1480	0.4556	0.9529	0.3690	-5222.08
quartiles (0.25%, 0.75%)	0.1159	0.3233	0.9655	0.2339	-5233.92
median (0.5%, 0.5%)	0.1925	0.7821	0.9195	0.2988	-5232.62

Heteroskedastic base regime

- Shifted distributions are more suitable for modeling the spikes
  - But … there are clusters of 'normal' prices classified as spikes



  - Perhaps different dynamics should be used for the base regime
- Introduce heteroskedasticity ($\gamma \neq 0$)

$$X_{t,i} = \alpha_i + (1 - \beta_i) X_{t-1,i} + \sigma_i |X_{t-1,i}|^{\gamma_i} \epsilon_{t,i}$$

Vasicek vs. heteroscedastic base regime dynamics

Table 2: Goodness-of-fit statistics for 2-regime models with heteroscedastic base regime dynamics and median-shifted lognormal spike distributions. Models for prices are summarized in columns 2-8, for log-prices in columns 9-15. p-values of 0.05 or more are emphasized in bold.

	Prices							Log-prices						
		Simulation			K-S test p-value				Simulation			K-S test p-value		
Data	γ	IQR	IDR	LogL	Base	Spike	Model	γ	IQR	IDR	LogL	Base	Spike	Model
						Shifted lognormal spikes								
EEX1	-0.43	0%	0%	-4169.3	0.0022	**0.2365**	0.0050	-4.08	22%	26%	625.5	0.0000	**0.9865**	0.0000
EEX2	-0.32	10%	2%	-5041.7	0.0125	**0.2306**	0.0276	-3.69	22%	12%	551.8	0.0000	**0.5875**	0.0000
PJM1	0.10	5%	1%	-4356.4	**0.0853**	**0.5408**	**0.1607**	-1.02	17%	6%	793.1	0.0006	**0.1924**	0.0011
PJM2	0.16	1%	-1%	-4989.3	**0.5882**	**0.1802**	**0.5435**	-0.01	1%	2%	804.2	**0.0582**	**0.1843**	**0.0995**
NEP1	0.22	2%	0%	-4326.3	0.0317	**0.4754**	**0.0742**	-1.35	9%	12%	643.1	0.0003	**0.8524**	0.0003
NEP2	0.62	0%	0%	-4654.0	**0.0828**	**0.3566**	**0.0983**	-2.37	1%	-1%	1445.2	0.0368	**0.1724**	**0.0980**

Leverage effect ?!

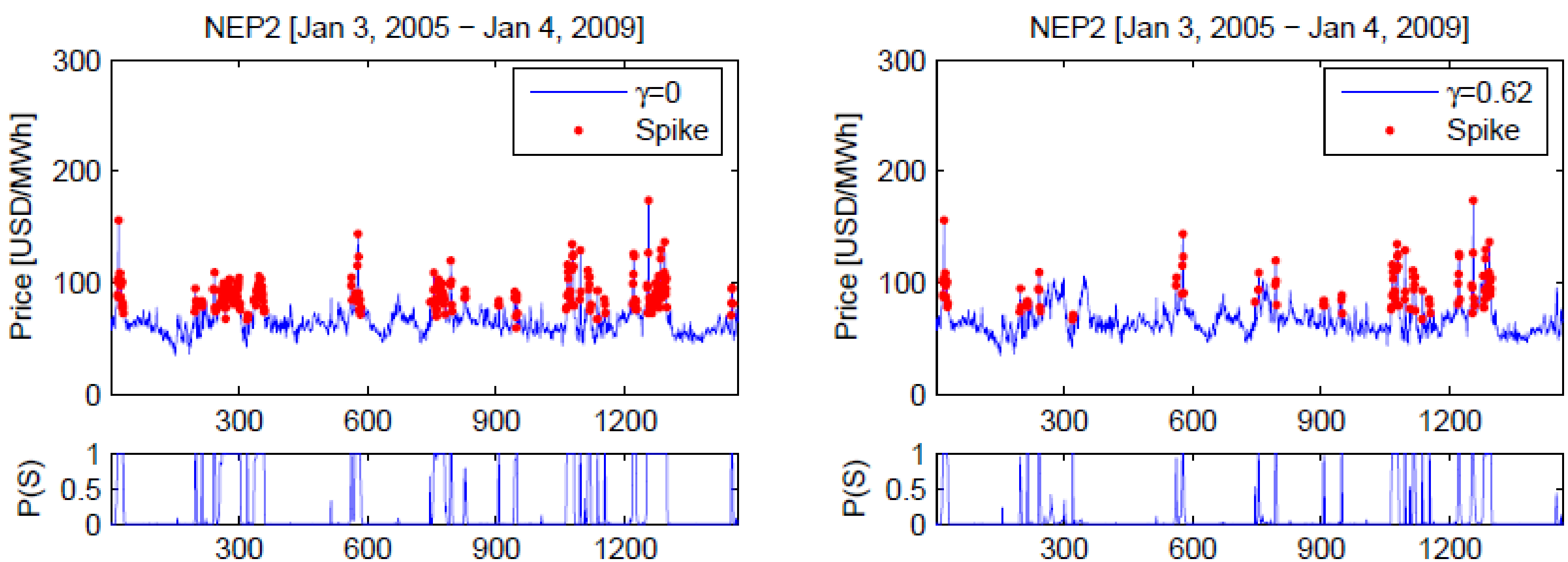
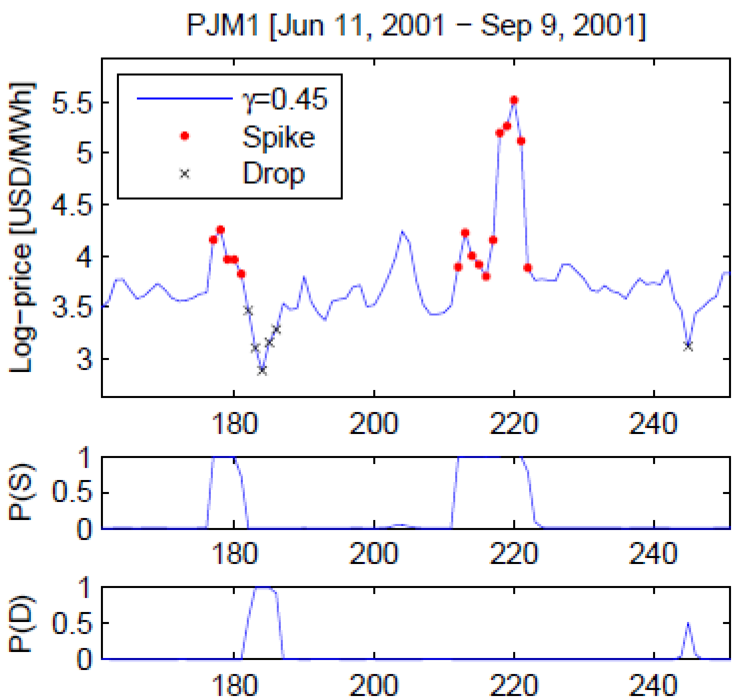


Figure 6: Sample calibration results for the 2-regime model with median-shifted lognormal spikes fitted to NEP2 prices. The difference between Vasicek (*left*) and heteroscedastic (*right*) base regime dynamics is clearly visible. Note, that due to the cutoff at the median, none of the price 'drops' are classified as spikes anymore.

3-regime models revisited

- Perhaps we need a 3rd regime to model 'price drops'
- Introduce a 3rd 'drop' regime
  - Contrary to the Huisman & Mahieu (2003) model, the price **can stay** in the 'excited' regimes ('spike' and 'drop')
  - Use a 'mirror image' or 'reflected' shifted log-normal distribution



IS 3-regime models

Table 4: Goodness of fit statistics for the IS 3-regime models with heteroscedastic base regime dynamics and median-shifted lognormal spikes and drops. p-values of 0.05 or more are emphasized in bold.

Data	Simulation			K-S test p-values			
	IQR	IDR	LogL	Base	Spike	Drop	Model
Prices							
EEX1	-1%	3%	-3798.2	**0.8371**	**0.0726**	**0.8576**	**0.5719**
EEX2	5%	-1%	-4848.5	**0.5510**	**0.2920**	**0.9196**	**0.3168**
PJM1	2%	0%	-4153.9	**0.3876**	**0.7052**	**0.7715**	**0.4072**
PJM2	2%	0%	-4723.2	**0.4824**	**0.3273**	0.0244	**0.4828**
NEP1	0%	0%	-4266.6	0.0404	**0.6121**	**0.9092**	**0.0609**
NEP2	5%	0%	-4610.3	**0.1359**	**0.7911**	**0.8771**	**0.1059**

Data	Simulation			K-S test p-values			
	IQR	IDR	LogL	Base	Spike	Drop	Model
Log-prices							
EEX1	-2%	5%	1181.2	**0.7297**	**0.3341**	**0.1971**	**0.5001**
EEX2	7%	1%	944.6	**0.2933**	**0.5090**	**0.5204**	**0.6517**
PJM1	8%	1%	1002.9	**0.3413**	**0.2726**	**0.9080**	**0.4640**
PJM2	1%	2%	1030.2	**0.2165**	**0.4604**	**0.2887**	**0.5258**
NEP1	1%	2%	853.0	**0.1084**	**0.8940**	**0.1948**	**0.1362**
NEP2	6%	0%	1573.7	**0.6858**	**0.7338**	**0.2334**	**0.8796**

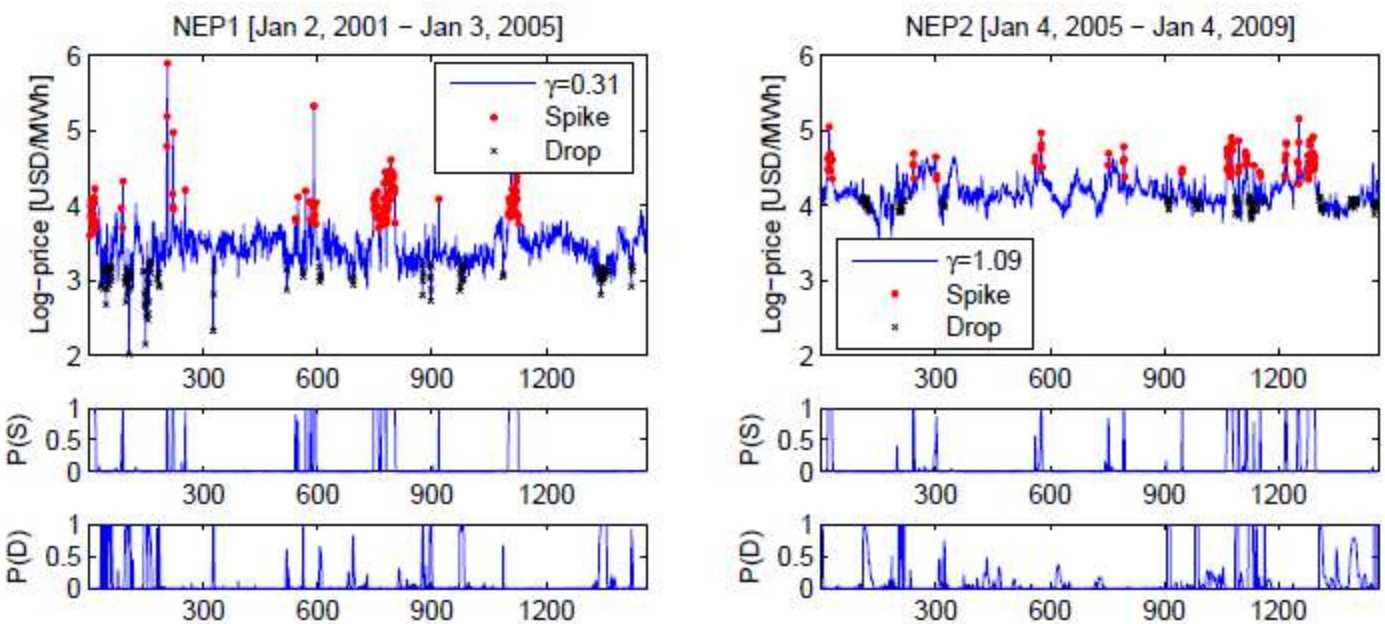


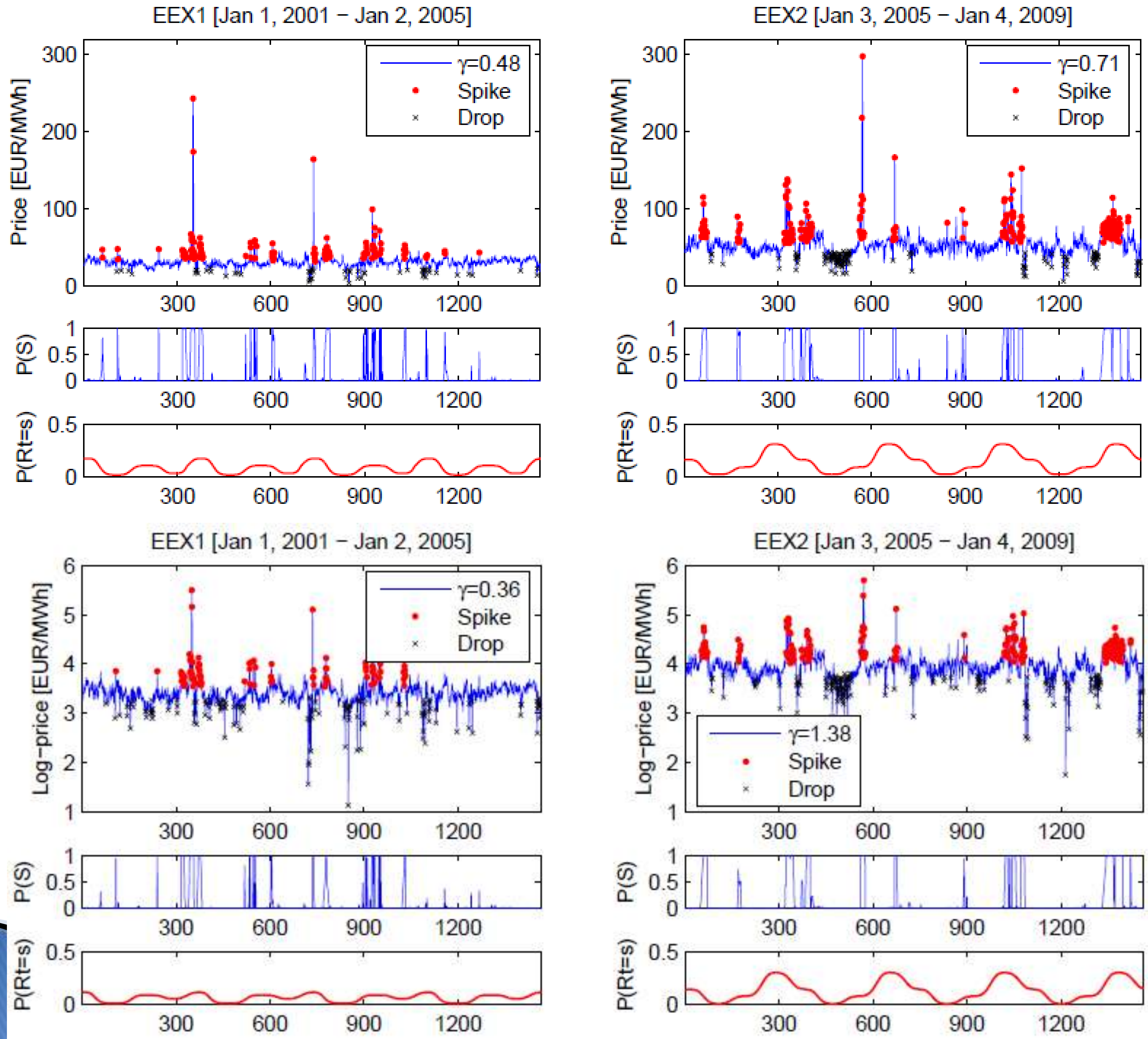
Figure 9: Calibration results for the IS 3-regime models with heteroscedastic base regime dynamics and median-shifted lognormal spikes and drops fitted to log-prices.

Inverse leverage effect !

Data	γ
Prices	
EEX1	0.6309
EEX2	0.3070
PJM1	0.6595
PJM2	0.1724
NEP1	0.5262
NEP2	0.0742
Log-prices	
EEX1	0.4102
EEX2	0.9115
PJM1	0.4481
PJM2	0.5057
NEP1	0.3068
NEP2	1.0923

IS 3-regime models with time-varying transition probabilities

- Admit a transition matrix with **time-varying (periodic) probabilities** $p_{ij}(t)$
- Calibrated in a two-step procedure in the last part of the E-step of the EM algorithm:
  - The probabilities are estimated independently for each season: Winter (XII-II), Spring (III-V), Summer (VI-VIII) and Autumn (IX-XI)
  - Then they are smoothed using a kernel density estimator with a Gaussian kernel
- This modification complicates gof testing
  - Only p-values for individual regimes are reported


EEX1 [Jan 1, 2001 − Jan 2, 2005]
γ=0.48
Spike
Drop
Price [EUR/MWh]
P(S)
P(Rt=s)
EEX2 [Jan 3, 2005 − Jan 4, 2009]
γ=0.71
Spike
Drop
Price [EUR/MWh]
P(S)
P(Rt=s)
EEX1 [Jan 1, 2001 − Jan 2, 2005]
γ=0.36
Spike
Drop
Log−price [EUR/MWh]
P(S)
P(Rt=s)
EEX2 [Jan 3, 2005 − Jan 4, 2009]
γ=1.38
Spike
Drop
Log−price [EUR/MWh]
P(S)
P(Rt=s)

Transition probabilities:

constant vs. time-varying

Data	Simulation IQR	Simulation IDR	LogL	K-S test p-values Base	K-S test p-values Spike	K-S test p-values Drop
			Prices			
EEX1	-0%	-2%	-3896.2	**0.9283**	**0.2444**	**0.5846**
EEX2	-4%	-1%	-4839.5	**0.1551**	**0.2536**	**0.8980**
PJM1	-5%	-2%	-4215.9	**0.1957**	**0.5965**	**0.3705**
PJM2	-4%	-4%	-4879.7	0.0424	**0.4312**	**0.1007**
NEP1	0%	0%	-4211.1	0.0134	**0.5258**	**0.7442**
NEP2	-3%	-2%	-4566.5	**0.2981**	**0.1140**	**0.4922**
			Log-prices			
EEX1	-1%	-3%	1132.7	**0.6898**	**0.8781**	0.0417
EEX2	-3%	0%	959.8	**0.3952**	**0.3190**	**0.2340**
PJM1	-2%	-3%	1018.7	**0.3139**	**0.1204**	**0.8435**
PJM2	-2%	-2%	904.2	**0.1272**	**0.0953**	**0.3402**
NEP1	0%	0%	863.6	**0.1881**	**0.8957**	**0.7062**
NEP2	-1%	-3%	1593.9	**0.7883**	**0.1881**	0.0409

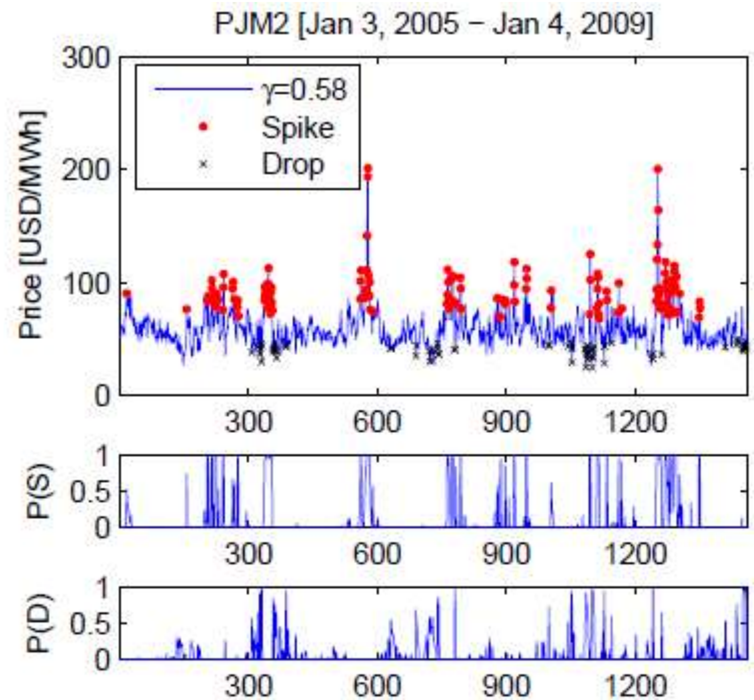


Figure 8: Comparison of calibration results for the IS 3-regime models with constant (*left*) and time-varying (*right*) transition probabilities. Note, the time-varying (periodic) intensity of spikes and price drops and the overall visually better fit of the latter model. The corresponding lower panels display the conditional probabilities $P(S) \equiv P(R_t = s|x_1, x_2, ..., x_T)$ and $P(D) \equiv P(R_t = d|x_1, x_2, ..., x_T)$ of being in the spike or drop regime, respectively.

Goodness-of-fit

- Testing the goodness-of-fit for processes is not straightforward
- We can
  - Test the marginal distributions using an EDF-type test, like the Kolmogorov-Smirnov (K-S) test
- … but
  - The K-S test cannot be applied directly …
  - In the considered models neither the prices themselves nor their differences or returns are i.i.d.

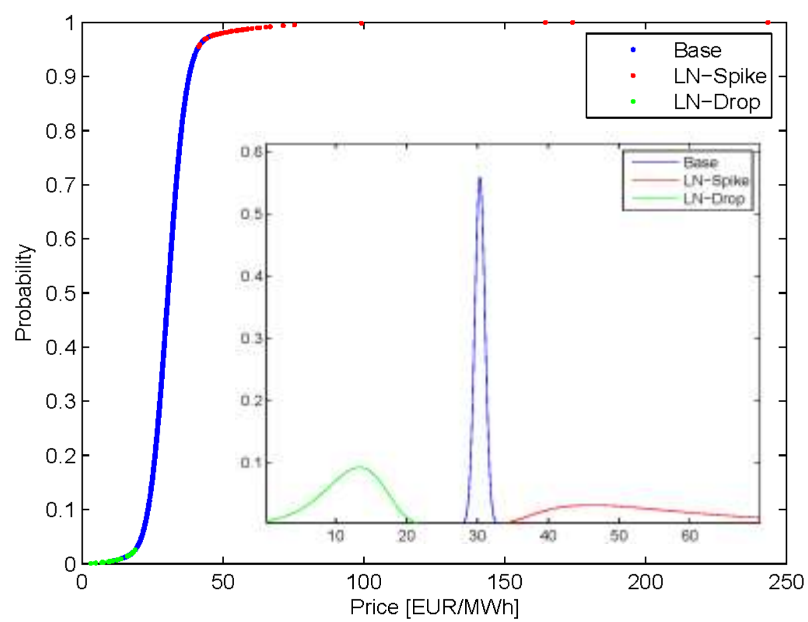
Testing procedure #1:

‘Equally weighted edf (ewedf)’

- Data is split into 2 subsets (3 for the 3-regime model)
  - **Spikes**, i.e. prices with probability $P(R_t = 2) > 0.5$
    - f_S-distributed, e.g. lognormal or Pareto
    - Price **drops**, i.e. prices with probability $P(R_t = 3) > 0.5$, f_D-distributed
  - Residuals of the **base** regime, i.e. remaining prices → N(0,1)

$$\varepsilon_{t,1} = \frac{X_t - (1-\beta_1 \Delta t) X_{t-\Delta t} - \alpha_1 \Delta t}{\sqrt{\Delta t}\sigma_1 |X_{t-\Delta t}|^{\gamma_1}}$$

- Run the K-S test for the subsets and the sample



Testing procedure #2:

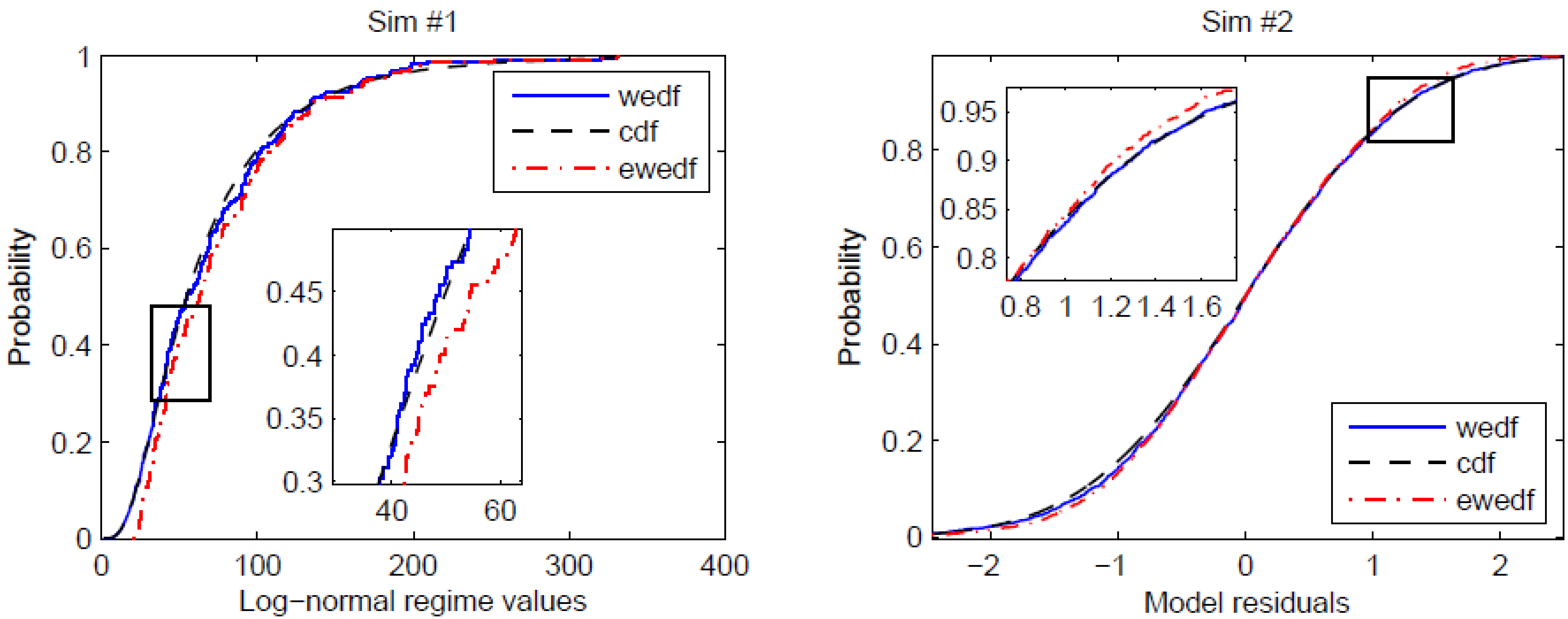
'Weighted edf (wedf)'

- Weighted empirical distribution function (edf)

$$F_n^w(x) = \sum_{t=1}^{n} \frac{w_t \mathbb{I}_{\{y_t < x\}}}{\sum_{t=1}^{n} w_t}$$

  - with $w_t = P(R_t = i | x_1, x_2, \ldots, x_T) = E(\mathbb{I}_{\{R_t = i\}} | x_1, x_2, \ldots, x_T)$
  - An unbiased and consistent estimator of F(t)
  - The statistics $D_n^w = \sqrt{n} \sup_{x \in \mathbb{R}} |F_n^w(x) - F(x)|$ converges (weakly) to the Kolmogorov-Smirnov distribution
    - For γ_1=0 → Janczura & Weron (2012)

Comparison: ewedf vs. wedf



	Parameters						Probabilities	
	α_1	β_1	σ_1^2	α_2	β_2	σ_2^2	p_{11}	p_{22}
Sim #1	10.0	0.8	10.0	4.0	-	0.5	0.9	0.2
Sim #2	1.0	0.8	1.0	3.0	0.4	0.5	0.6	0.5

Monte Carlo testing scheme

- If the parameters are estimated from the data, the *p*-values might be overestimated
- Stephens (1978) proposed the `half-sample' approach
  - Use half the data to estimate the parameters, but then the entire data set to conduct the test
    - Dependent on the choice of the `half sample'
    - There is no way of increasing the accuracy
- Ross (2002) advocates the use of Monte Carlo simulations

Monte Carlo testing scheme cont.

- Testing scheme:
  - Parameters are estimated for a sample of size *n* → $\hat{\theta}$
  - Assuming that the sample is $F(x;\hat{\theta})$-distributed, the EDF test statistics is calculated → *d*
  - Generate an `$F(x;\hat{\theta})$-distributed' trajectory of size *n*
    - The parameter vector is estimated → $\hat{\theta}_1$
    - d_1 is calculated assuming that the sample is $F(x;\hat{\theta}_1)$-distributed
  - This is repeated as many times as required (500, 1000, …)
  - The *p*-value is obtained as the proportion of times that $d_i \geq d$

Example: 2 or 3 regimes?

- NEPOOL mean daily day-ahead spot prices
- January 2, 2006 – January 2, 2011 (1827 observations)
- MRS models for deseasonalized log-prices:
  - 2-regime

$$\begin{cases} dX_{t,b} = (\alpha - \beta X_{t,b})dt + \sigma_b dW_t \\ \log(X_{t,s} - m) \sim N(\mu_s, \sigma_s{}^2) \end{cases}$$

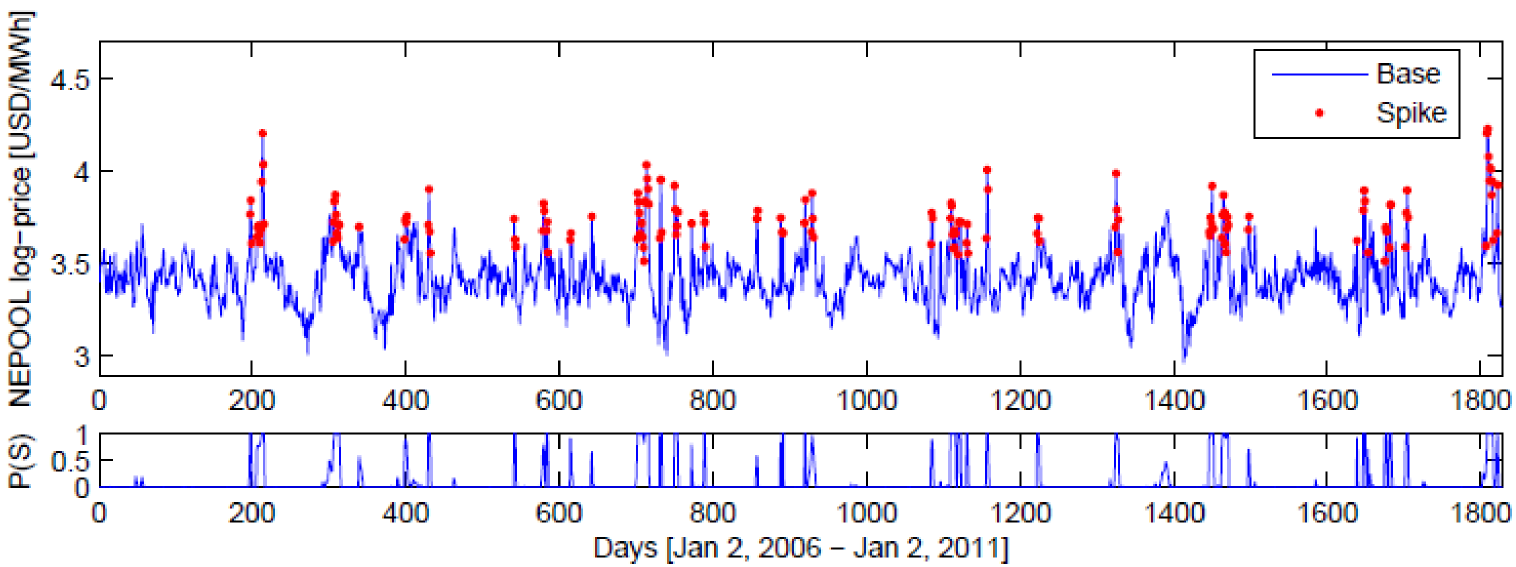
  - 3-regime

$$\begin{cases} dX_{t,b} = (\alpha - \beta X_{t,b})dt + \sigma_b dW_t \\ \log(X_{t,s} - m) \sim N(\mu_s, \sigma_s{}^2), \;\; X_{t,s} > m \\ \log(m - X_{t,d}) \sim N(\mu_d, \sigma_d{}^2), \;\; X_{t,d} < m \end{cases}$$

Example cont.

- 2-regime model for NEPOOL log-prices

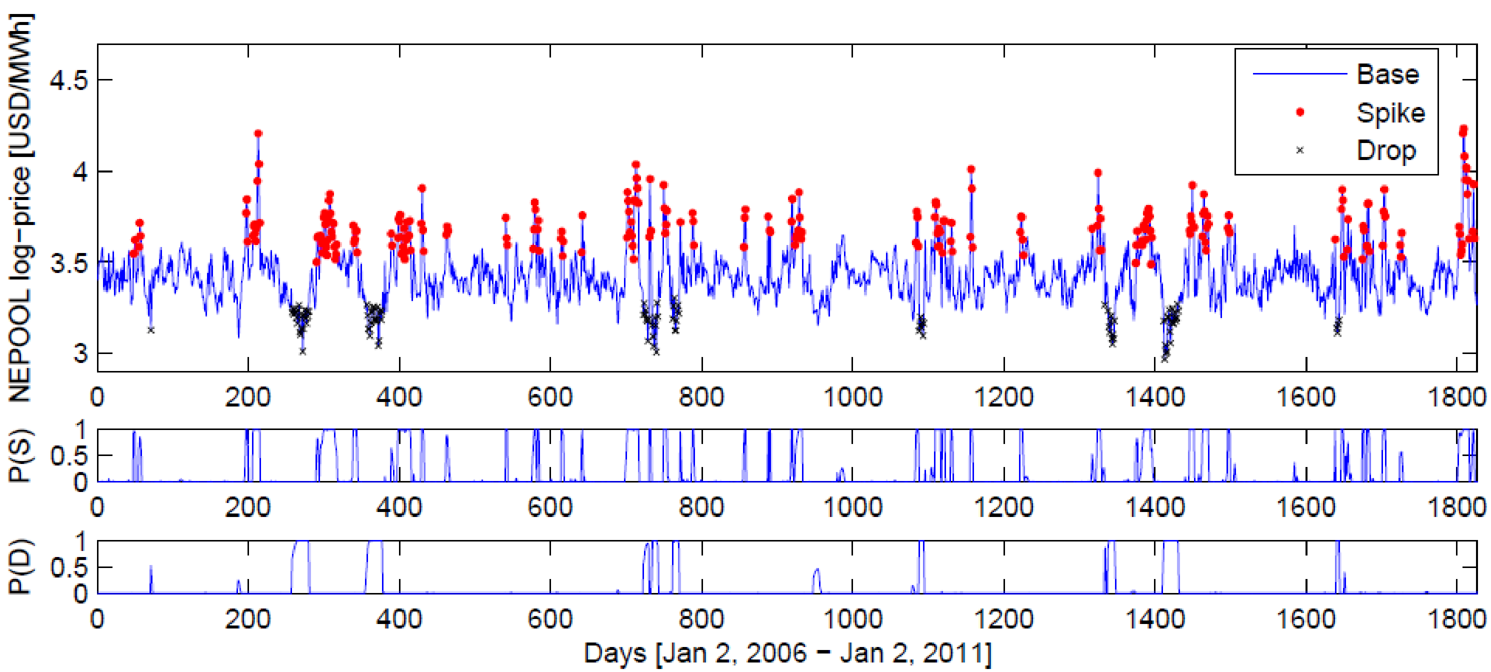
	ewedf				wedf			
Regime	Base	Spike	Drop	Model	Base	Spike	Drop	Model
	2-regime model							
K-S test tables	**0.21**	**0.27**	-	**0.12**	**0.08**	**0.93**	-	**0.10**
MC simulations	0.01	**0.07**	-	0.01	0.00	**0.57**	-	0.00



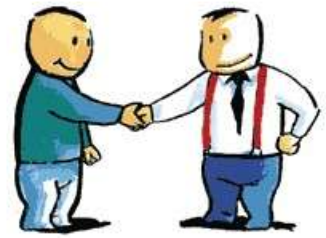
Example cont.

- 3-regime model for NEPOOL log-prices

	ewedf				wedf			
Regime	Base	Spike	Drop	Model	Base	Spike	Drop	Model
				3-regime model				
K-S test tables	**0.56**	**0.25**	**0.98**	**0.68**	**0.38**	**0.71**	**0.92**	**0.51**
MC simulations	**0.19**	**0.06**	**0.81**	**0.20**	**0.31**	**0.27**	**0.49**	**0.25**



Conclusions



- The quest for the model is not over
  - Improve the timing of spikes
- The **devil** is in **deseasonalization**
  - Preprocess data before fitting the seasonal components?
  - Use fundamental data to better fit long term seasonality?

The end

References

- J. Janczura, R. Weron (2010) An empirical comparison of alternate regime-switching models for electricity spot prices, Energy Economics 32, 1059-1073
- J. Janczura, R. Weron (2012) Efficient estimation of Markov regime-switching models: An application to electricity spot prices, AStA, DOI 10.1007/s10182-011-0181-2
- J. Janczura, R. Weron (2012) Goodness-of-fit testing for regime-switching models, Working Paper
- J. Nowotarski, J. Tomczyk, R. Weron (2011) Wavelet-based modeling and forecasting of the seasonal component of spot electricity prices, Paper presented at EFC11
- S. Trück, R. Weron, R. Wolff (2007) Outlier treatment and robust approaches for modeling electricity spot prices, Bulletin of the ISI 62, 1524; revised version in 2008

- Reviews

  - F.E. Benth, J.S. Benth, S. Koekebakker (2008) *Stochastic Modeling of Electricity and Related Markets*, World Scientific
  - M. Burger, B. Graeber, G. Schindlmayr (2007) *Managing Energy Risk*, Wiley
  - A. Eydeland, K. Wolyniec (2003) *Energy and Power Risk Management*, Wiley
  - R. Weron (2006) *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, Wiley