Forecasting Volatility of Wind Power Production

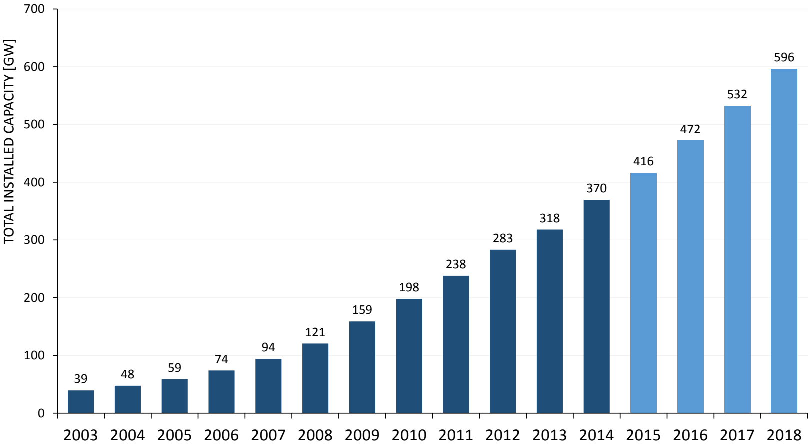
Zhiwei Shen and Matthias Ritter

Department of Agricultural Economics

Humboldt-Universität zu Berlin

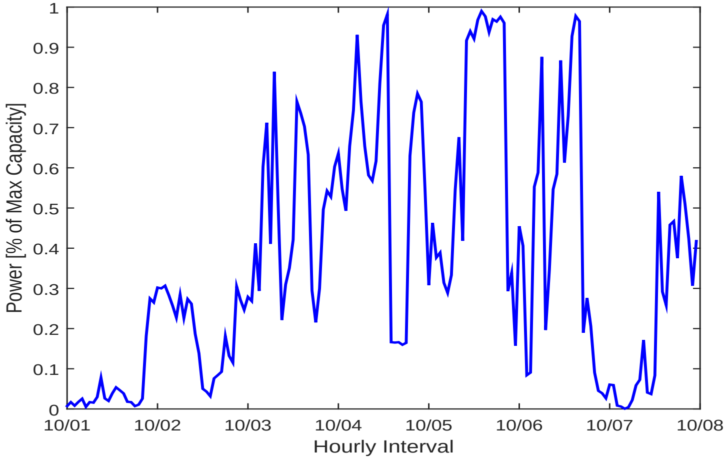
July 18, 2015

Wind Power Development



Data: GWEC (2014,2015)

Volatile Wind Power



Data: 4initia GmbH

Wind power forecasts

Motivation

- It fluctuates significantly due to changing weather condition.
- Volume risk: non-storable commodity.
- Importance in prediction of wind power production:
  - optimize power plant scheduling to balance supply and demand for a regional or national grid.
  - help energy traders make informed decision on how much they can offer or bid in the next trading cycle.

Wind power forecasts

Approaches

- Models
  - Meteorological models: Weather prediction + Power curve(Monterio et al., 2009);
  - Statistical models: time series, data mining such as neural networks or support vector machines (Giebel et al., 2011)

Wind power forecasts

Approaches

- Models
  - Meteorological models: Weather prediction + Power curve(Monterio et al., 2009);
  - Statistical models: time series, data mining such as neural networks or support vector machines (Giebel et al., 2011)
- Forecasts of wind power production
  - Point forecasts /Mean forecasts: a single value of conditional expectation of wind power production.
  - Probabilistic forecasts: such as quantile or interval forecast (Bremnes, 2004; Anastasiades and McSharry, 2013) ,or full predictive density forecasts (Lau and McSharry, 2010)

Wind power forecasts

Volatility forecast of wind power production

- Features of wind power production
  - not only wind speed but also wind volatility is time-varying.
  - time-varying heteroscedasticity similar to financial market (Lau and McSharry, 2010)
  - affected by ramp events[1]
- Volatility modelling
  - Seasonal: Fourier series (Alexandridis and Zapranis, 2013)
  - ARCH or GARCH models (Tastu et al., 2014; Liu et al., 2011; Lau and McSharry, 2010)

[1] ramp event: energy output changes by substantial fraction of the capacity within a short time

Objective

Research questions

- Which model performs best in terms of volatility forecasting?
- Traditional GARCH models may be too restrictive to capture random breaks and non-linear behaviour of wind power data.

Objectives

- Identify the best predictive model for the volatility of wind power production
- Develop Markov regime switching GARCH model to capture the dynamics of wind power production

Contents

Contents

GARCH Models

Considering a time series of wind power y_t, we use the AR (k) model:

$$y_t = c + \sum_{i=1}^{k} \phi_i y_{t-i} + \epsilon_t \tag{1}$$

$$\epsilon_t = \eta_t h_t, \qquad \eta_t \sim (0, 1) \tag{2}$$

- GARCH

$$h_t^2 = \omega + \sum_{i=1}^{m} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{n} \beta_j h_{t-j}^2 \tag{3}$$

- GARCH(1,1)

$$h_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}^2 \tag{4}$$

Asymmetric GARCH

- Exponential GARCH (EGARCH)

$$\ln(h_t) = \omega + \alpha(|\eta_{t-1}| - E[|\eta_{t-1}|]) + \gamma\eta_{t-1} + \beta \ln(h_{t-1}) \quad (5)$$

where α and γ capture the asymmetric effect of magnitude and sign of η on volatility. $\gamma < 0$, the negative value of ϵ results in higher volatility than the positive value. No restriction on the parameters.

Asymmetric GARCH

- Threshold GARCH (Zakoian,1994)[2]

$$h_t = \omega + \alpha(|\epsilon_{t-1}| - \gamma\epsilon_{t-1}) + \beta h_{t-1} \tag{6}$$

- GJR GARCH (Glosten et al., 1993)

$$h_t^2 = \omega + \alpha(|\epsilon_{t-1}| - \gamma\epsilon_{t-1})^2 + \beta h_{t-1}^2. \tag{7}$$

- Nonlinear GARCH (Engle and Ng, 1993)

$$h_t^2 = \omega + \alpha(\epsilon_{t-1} - \gamma h_{t-1})^2 + \beta h_{t-1}^2 \tag{8}$$

where γ reflects the asymmetric effect. $\gamma > 0$, all these three models indicated that negative error terms increase future volatility by larger amount than positive ones.

[2] The following formulas are rearranged and adopted from Ding et al., (1993).

Markov regime switching GARCH Model

Markov regime switching model (Hamilton, 1989)

- Multiple discrete regimes or states governed by a state variable s_t
- Different dynamics and sets of parameters on each regime
- Regime switching by a first-order Markov chain

The state variable s_t is assumed to evolve with transition probability:

$$\mathrm{P}(s_t = j | s_{t-1} = i) = p_{ij} \tag{9}$$

In the case of two regimes $s_t \in \{1, 2\}$, the transition matrix:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} p & 1-q \\ 1-p & q \end{bmatrix} \tag{10}$$

Markov regime switching GARCH model

Markov regime switching GARCH model

- Conditional mean: Markov regime switching AR(p)

$$y_t = c^{(j)} + \sum_{i=1}^{k} \phi_i^{(j)} y_{t-i} + \epsilon_t \qquad \epsilon_t = \eta_t h_t, \quad \eta_t \sim (0,1) \tag{11}$$

where subscript (j) denotes the regime of the process at t.

- Conditional variance: Markov regime switching GARCH(1,1)

$$h_t^{2|(j)} = \alpha_0^{(j)} + \alpha_1^{(j)} \epsilon_{t-1}^2 + \beta_1^{(j)} h_{t-1}^2 \tag{12}$$

h_{t-1} is a state-independent average of past conditional variances. Using expression of Klaassen (2002):

$$h_t^{2|(j)} = \alpha_0^{(j)} + \alpha_1^{(j)} \epsilon_{t-1}^2 + \beta_1^{(j)} \mathbb{E}_{t-1}\{h_{t-1}^{2|(j)} | s_t\} \tag{13}$$

Estimation: Markov regime switching GARCH model

Estimation

The conditional probability of s_t being in regime $j = 1, 2$, given the information set $\Omega_t = \{y_t, y_{t-1}...y_1\}$ and parameters Θ:

$$\xi^{(j)}_{s_t|\Omega_t} = \mathrm{P}(s_t = j|\Omega_t, \Theta) \tag{14}$$

To derive $\xi^{(j)}_{s_t|\Omega_t}$, reformulating (14) via conditional probability of y_t at $s_t = j$:

$$\xi^{(j)}_{s_t|\Omega_t} = \frac{f(y_t, s_t = j|\Omega_{t-1}, \Theta)}{f(y_t|\Omega_{t-1}, \Theta)} \tag{15}$$

where $f(y_t, s_t = j|\cdot)$: joint conditional density of y_t and regime j.

$$f(y_t, s_t = j|\Omega_{t-1}, \Theta) = \xi^{(j)}_{s_t|\Omega_{t-1}} f(y_t|s_t = j, \Omega_{t-1}, \Theta) \tag{16}$$

$$\xi^{(j)}_{s_t|\Omega_{t-1}} = \mathrm{P}(s_t = j|\Omega_{t-1}, \Theta) \tag{17}$$

$$\boldsymbol{\xi}_{s_t|\Omega_{t-1}} = \mathbf{P}^T \boldsymbol{\xi}_{s_{t-1}|\Omega_{t-1}} \tag{18}$$

where $\xi^{(j)}_{s_t|\Omega_{t-1}}$ forecast of probability of s_t being in j; $\mathbf{P}^T$ transition matrix .

Estimation: Markov regime switching GARCH model

Log-likelihood function

The conditional density of y_t:

$$f(y_t|\Omega_{t-1},\Theta) = \sum_{j=1}^{2} f(y_t, s_t = j|\Omega_{t-1},\Theta)$$
$$= \sum_{j=1}^{2}\sum_{i=1}^{2} p_{ij}\xi_{t-1}^{j} f(y_t|s_t = j, \Omega_{t-1},\Theta) \quad (19)$$

The log-likelihood function:

$$L(y_1, y_2, ...y_T|\Theta) = \sum_{t=1}^{T} \log f(y_t|\Omega_{t-1},\Theta) \quad (20)$$

Forecasting Evaluation

Measures (Patton, 2011; Byun and Cho, 2013)

$\hat{h}_t^2$: conditional volatility forecast at t and $\hat{\sigma}_t^2$ is the *ex post* proxy of conditional variance representing the actual volatility at t. Statistical loss functions:

- Root Mean Square Error (RMSE) $\text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}(\hat{\sigma}_t^2 - \hat{h}_t^2)^2}$
- RMSE-LOG $\text{RMSE-LOG} = \sqrt{\frac{1}{N}\sum_{t=T+1}^{T+N}(\log\hat{\sigma}_t^2 - \log\hat{h}_t^2)^2}$
- Mean Absolute Error (MAE) $\text{MAE} = \frac{1}{N}\sum_{t=T+1}^{T+N}|\hat{\sigma}_{t+1}^2 - \hat{h}_{t+1|t-k}|$
- MAE-LOG $\text{MAE-LOG} = \frac{1}{N}\sum_{t=T+1}^{T+N}|log\hat{\sigma}_{t+1}^2 - \log\hat{h}_{t+1|t-k}|$
- QLIKE $\text{QLIKE} = \frac{1}{N}\sum_{t=T+1}^{T+N}(\log(\hat{h}_t^2) + \frac{\hat{\sigma}_t^2}{\hat{h}_t^2})$

In this paper, we use realized volatility as $\hat{\sigma}_t^2$.

Forecasting Evaluation

Diebold-Mariano (DM) test

The DM test is applied to determine if the predictive accuracies of competing models are significantly different.

- Define loss function differential between model a and model b:

$$d_t = g(e_{a,t}) - g(e_{b,t})$$

where $g(\cdot)$ means a loss function and $e_{\cdot,t}$ corresponds to forecast errors from the competing models.

- Two typical loss functions: $g(e_t) = e_t^2$ or $g(e_t) = |e_t|$.
- H_0: $E(d_t) = 0$, DM test statistic:

$$DM = \bar{d}/\sqrt{\hat{V}(\bar{d})} \sim N(0,1)$$

where $\bar{d} = \frac{1}{n}\sum_{t-1}^{n} d_t$ and $\hat{V}(\bar{d}) = \frac{2\pi \hat{f}_d(0)}{T}$

Contents

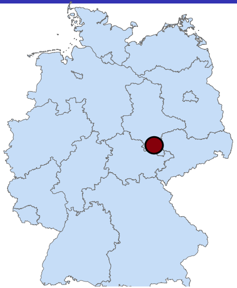
Data

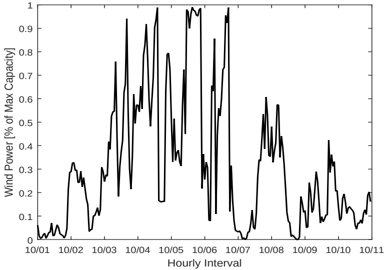
Wind power production

- Hourly interval data for 1.10.2012–31.12.2013 (Estimation) 1.1.2014–07.01.2014 (Evaluation)
- 6 turbines with capacity 2.3MW
- Normalization by max. capacity

Realized volatility

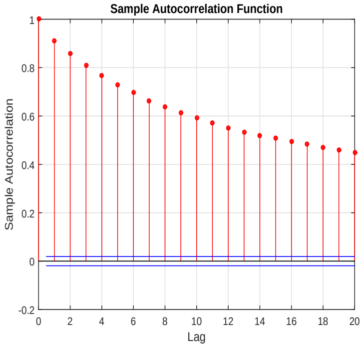
- Cumulative squared deviation over different time intervals.
- 10 mins interval data: $y_{t,i}$ $(i = 0, ..., 6)$
- RV: $\hat{\sigma}^2_{RV,t} = \sum_{i=1}^{6}(y_{t,i} - y_{t,i-1})^2$

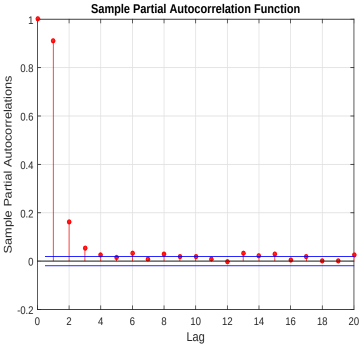




Data analysis

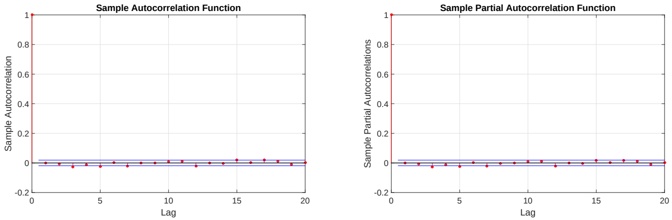
Wind power data:





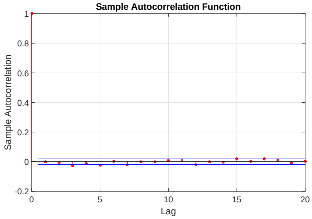
Data analysis

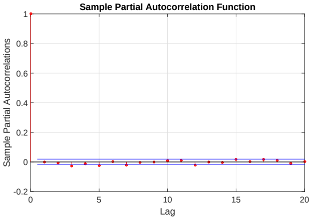
AR(3): ϵ



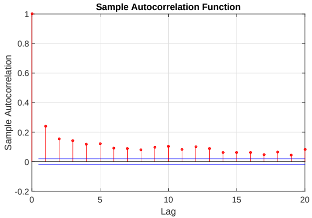
Data analysis

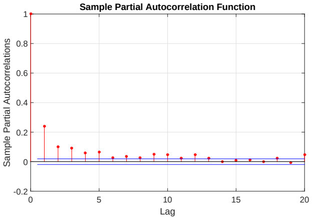
AR(3): ϵ



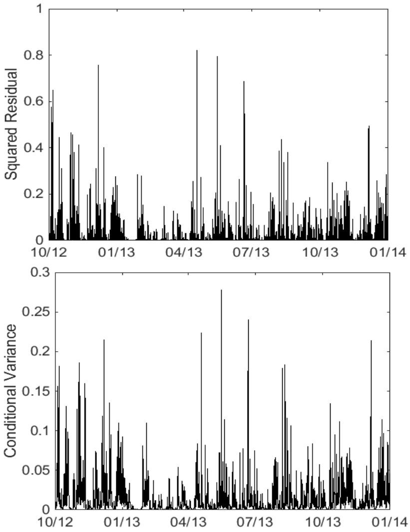


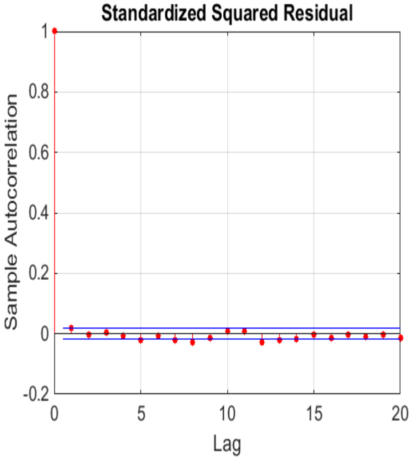
AR(3): ϵ^2





SGARCH models: AR(3)-GARCH(1,1)





Model comparison: In-sample

Table : In-sample goodness-of-fit statistics

Model	# Par.	Persistence	AIC	BIC	Log-likelihood	Rank
SGARCH	7	0.999	-2.167	-2.163	11690.55	6
EGARCH	8	0.959	-2.290	-2.284	12351.78	3
TGARCH	8	0.993	-2.297	-2.292	12394.37	2
GJR-GARCH	8	0.999	-2.233	-2.227	12043.92	5
NGARCH	8	0.945	-2.269	-2.264	12240.36	4
MRS-GARCH	16	0.999	-2.783	-2.773	15021.21	1

Note: for MRS-GARCH only highest persistence is reported. AIC is calculated as $(-2\log(L) + 2k)/T$, where k is the number of parameters and T is the number of observations. BIC is calculated as $(-2\log(L) + k\log(T))/T$.

Model comparison: Out-of-sample

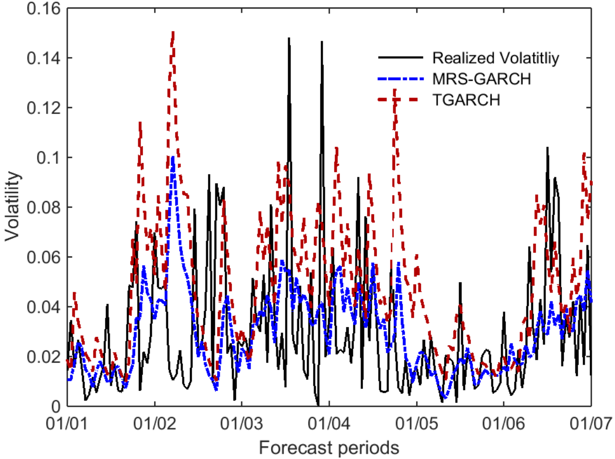


Fig. Realized volatility and 1-step ahead volatility from MRS-GARCH and TGARCH

Model comparison: Out-of-sample

Table : Out-of-sample evaluation for 1-step ahead volatility forecast

Model	RMSE (Rank)	RMSE-LOG (Rank)	MAE (Rank)	MAE-LOG (Rank)	QLIKE (Rank)
SGARCH	0.0335 (3)	1.9332 (2)	0.0244 (3)	0.9195 (4)	-2.443 (3)
EGARCH	0.0328 (2)	1.9845 (3)	0.0243 (2)	0.9117 (2)	-2.4592 (2)
TGARCH	0.0416 (6)	2.3421 (6)	0.0308 (6)	1.0139 (6)	-2.4227 (6)
GJR-GARCH	0.039 (5)	2.1738 (5)	0.0283 (5)	0.9687 (5)	-2.4321 (4)
NGARCH	0.0344 (4)	2.0626 (4)	0.0254 (4)	0.9134 (3)	-2.4813 (1)
MRS-GARCH	0.0295 (1)	1.7597 (1)	0.0207 (1)	0.8406 (1)	-2.4244 (5)

Model comparison: Out-of-sample

Table : Out-of-sample evaluation for 5-step ahead volatility forecast

Model	RMSE (Rank)	RMSE-LOG (Rank)	MAE (Rank)	MAE-LOG (Rank)	QLIKE (Rank)
SGARCH	0.1260 (3)	0.5487 (4)	0.0882 (4)	0.5694 (4)	-2.3679 (4)
EGARCH	0.1001 (2)	0.4279 (2)	0.0701 (2)	0.4770 (2)	-2.3737 (3)
TGARCH	0.1732 (6)	0.6709 (6)	0.1219 (6)	0.6300 (6)	-2.3677 (5)
GJR-GARCH	0.16 (5)	0.6468 (5)	0.1062 (5)	0.5951 (5)	-2.3507 (6)
NGARCH	0.1281 (4)	0.4805 (3)	0.0881 (3)	0.5133 (3)	-2.4104 (2)
MRS-GARCH	0.0982 (1)	0.4066 (1)	0.0676 (1)	0.4696 (1)	-2.4287 (1)

Model comparison: Out-of-sample

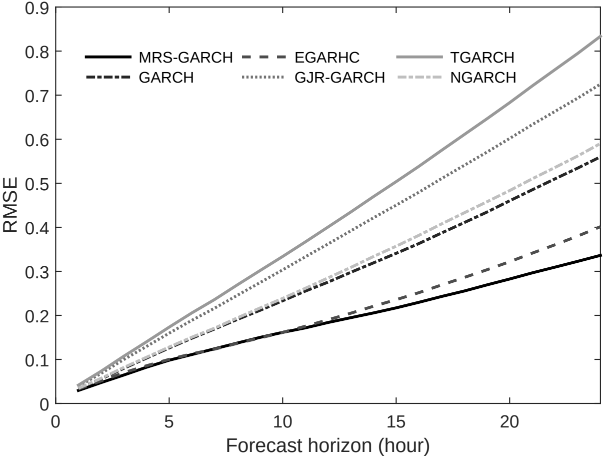


Fig. Forecast horizon and root mean square error

Model comparison: Out-of-sample

Table : Diebold-Mariano Test (1-step-ahead)

Model	Square error loss	Absolute error loss
MRS-GARCH	Benchmark	
SGARCH	-2.49^{***}	-3.64^{***}
EGARCH	-2.02^{***}	-3.44^{***}
TGARCH	-4.24^{***}	-5.54^{***}
GJR-GARCH	-3.86^{***}	-4.82^{***}
NGARCH	-2.86^{***}	-3.79^{***}

Note: the negative sign implies that benchmark's loss is lower than loss implied by other models. Asterisks *** denote significance at the 1% level .

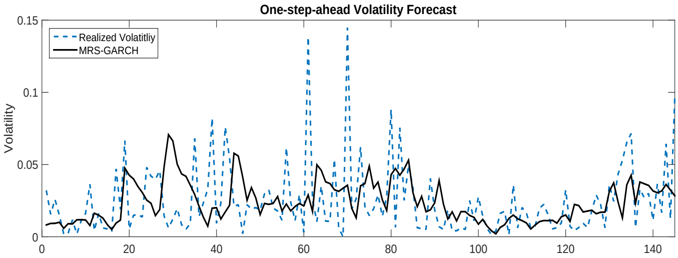
Model comparison: Out-of-sample

Table : Diebold-Mariano Test (1-step-ahead)

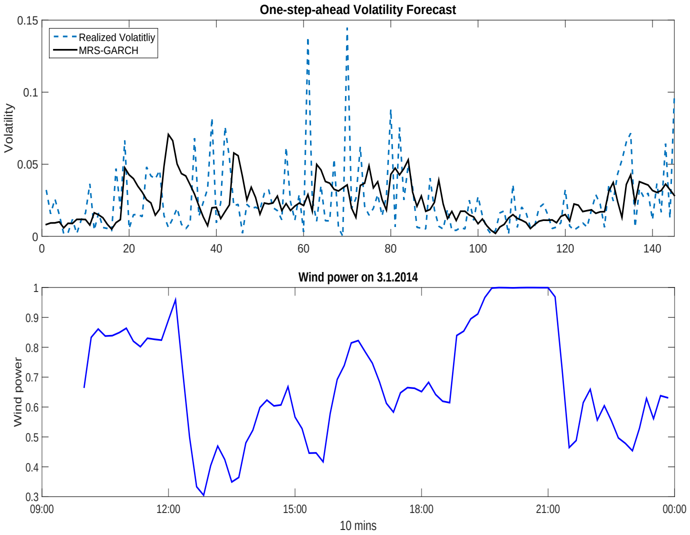
Model	Square error loss	Absolute error loss
EGARCH	Benchmark	
MRS-GARCH	2.01^{***}	3.44^{***}
SGARCH	-0.48^{*}	-0.06^{*}
TGARCH	-4.60^{***}	-5.81^{***}
GJR-GARCH	-4.13^{***}	-4.67^{***}
NGARCH	-1.12^{***}	-1.58^{***}

Note: the negative sign implies that benchmark's loss is lower than loss implied by other models. Asterisks *** and * denote significance at the 1% level and 10% level, respectively.

Model comparison: Jumps in realized volatility



Model comparison: Jumps in realized volatility



Conclusion

Findings

- MRS-GARCH model outperforms traditional GARCH models in in-sample and out-of-sample comparisons.
- EGARCH model might also be a fair choice to forecast the volatility of wind power production due to computational effort and information gain.
- The abrupt changes in realized volatility cannot be predicted due to the big instantaneous jumps in higher frequency wind power production.

Extensions

- Non-normal heavy tailed distributions
- Incorporating Markov regime switching with asymmetric GARCH models
- Validation using other wind farm data

Forecasting Volatility of Wind Power Production

Zhiwei Shen and Matthias Ritter

Department of Agricultural Economics

Humboldt-Universität zu Berlin

July 18, 2015

Reference I

Alexandriidis, A. and Zapranis, A. (2013).
Wind derivatives: Modeling and pricing.
Computational Economics, 41(3):299–326.

Anastasiades, G. and McSharry, P. (2013).
Quantile forecasting of wind power using variability indices.
Energies, 6(2):662–695.

Benth, J. Š. and Benth, F. E. (2010).
Analysis and modelling of wind speed in new york.
Journal of Applied Statistics, 37(6):893–909.

Bremnes, J. B. (2004).
Probabilistic wind power forecasts using local quantile regression.
Wind Energy, 7(1):47–54.

Byun, S. J. and Cho, H. (2013).
Forecasting carbon futures volatility using garch models with energy volatilities.
Energy economics, 40:207–221.

Ding, Z., Granger, C. W., and Engle, R. F. (1993).
A long memory property of stock market returns and a new model.
Journal of empirical finance, 1(1):83–106.

Reference II

Engle, R. F. and Ng, V. K. (1993).
Measuring and testing the impact of news on volatility.
The journal of finance, 48(5):1749–1778.

Giebel, G., Brownsword, R., Kariniotakis, G., Denhard, M., and Draxl, C. (2011).
The state-of-the-art in short-term prediction of wind power: A literature overview.
Technical report, ANEMOS. plus.

Glosten, L. R., Jagannathan, R., and Runkle, D. E. (1993).
On the relation between the expected value and the volatility of the nominal excess return on stocks.
The journal of finance, 48(5):1779–1801.

Hamilton, J. D. (1989).
A new approach to the economic analysis of nonstationary time series and the business cycle.
Econometrica: Journal of the Econometric Society, pages 357–384.

Klaassen, F. (2002).
Improving garch volatility forecasts with regime-switching garch.
In *Advances in Markov-Switching Models*, pages 223–254. Springer.

Lau, A. and McSharry, P. (2010).
Approaches for multi-step density forecasts with application to aggregated wind power.
The Annals of Applied Statistics, pages 1311–1341.

Reference III

Liu, H., Erdem, E., and Shi, J. (2011).
Comprehensive evaluation of arma-garch(-m) approaches for modeling the mean and volatility of wind speed.
Applied Energy, 88(3):724–732.

Monteiro, C., Bessa, R., Miranda, V., Botterud, A., Wang, J., Conzelmann, G., et al. (2009).
Wind power forecasting: state-of-the-art 2009.
Technical report, Argonne National Laboratory (ANL).

Patton, A. J. (2011).
Volatility forecast comparison using imperfect volatility proxies.
Journal of Econometrics, 160(1):246–256.

Tastu, J., Pinson, P., Trombe, P.-J., and Madsen, H. (2014).
Probabilistic forecasts of wind power generation accounting for geographically dispersed information.
Smart Grid, IEEE Transactions on, 5(1):480–489.

Zakoian, J.-M. (1994).
Threshold heteroskedastic models.
Journal of Economic Dynamics and control, 18(5):931–955.

Appendx: Equations

$$\mathbb{E}_{t-1}\{h_{t-1}^{2|(j)} \mid s_t\} = \sum_{i=1}^{2} \bar{p}_{ij,t-1}[(\mu_{t-1}^{(i)})^2 + h_{t-1}^{2|(i)}] - [\sum_{i=1}^{2} \bar{p}_{ij,t-1}\mu_{t-1}^{(i)}]^2$$

$$\bar{p}_{ij,t-1} = \Pr(s_t = i \mid s_{t+1} = j, \Omega_{t-1}) = \frac{p_{ij}\Pr(s_t = i \mid \Omega_{t-1})}{\Pr(s_{t+1} = j \mid \Omega_{t-1})} = \frac{p_{ij}p_{i,t}}{p_{j,t+1}}$$

Appendix: Equations

In Diebold Mariano Test,

$$f_d(0) = \frac{1}{2\pi}(\sum_{k=-\infty}^{\infty} \gamma_d(k)) \tag{21}$$

is the spectral density of loss differential at frequency 0. $\gamma_d(k)$ is autocovariance of loss differential at lag k.

MRS-GARCH: estimated sequence of regimes probabilities

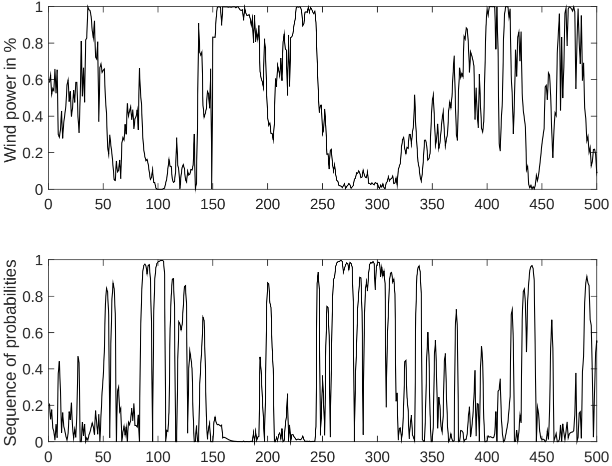


Fig. Wind power production and the estimated sequence of probabilities of being in regime 1 (i.e., low regime)