Estimation of wind turbine energy production

Alexandre Brouste

Laboratoire Manceau de Mathématiques
Institut du Risque et de l'Assurance du Mans

Séminaire FIME, IHP

Worldwide installed wind power: 282 GW (2011), 318 GW (2012) mainly China (22.1%), USA (26.4%), Europe (Denmark, Italy, France, UK, Germany, Spain) (28.7%).

France installed wind power: 7.5 GW (2011), 8.2 GW (2012).

France net wind energy production: 14.9 TWh (2012). Since the nominal production is 69.8 TWh, the load factor is about 20%.

This production represent 2.7% of the 541TWh (2012) produced in France.

Manufacturers: Vestas (13.2%), Goldwind (10.3%), Enercon (10.1%), Siemens (8.0%), Suzlon Group - Repower (6.3%), General Electrics (4.9%), Gamesa (4.6%), Guodian United Power (3.9%), Mingyang (3.7%), Nordex (3.4%), others (31.6%).

Energy companies: China Guodian Corp. (4.9%), Iberdrola (4.8%), NextEra Energy (3.7%), China Huaneng Corp. (2.8%), China Datang Corp. (2.6%), EDP (2.3%), Acciona (1.9%), E.on (1.7%), Enel (1.6%), EDF (1.5%), China Huadian (1.5%), Shenua Group (1.5%), Berkshire Hathaway (1.3%), GDF Suez (1.3%), others (66.6%).

In France, the Top 3 is GDF-Suez, Enel and EDF (800 MW).

We are interested in two main applications (at the wind farm level):

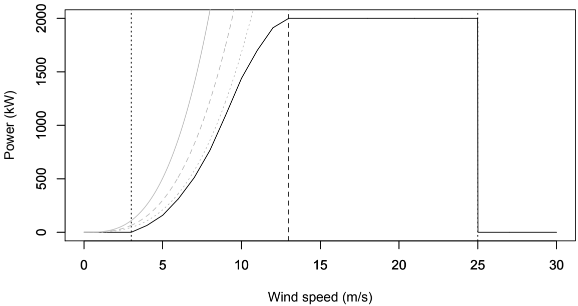
1. Annual wind energy production estimation: pre-installation approach for investment purposes. Investors need the present net value of a wind farm project (production, maintenance costs). Production quantiles can be obtained from the wind speed distribution measured on the site.

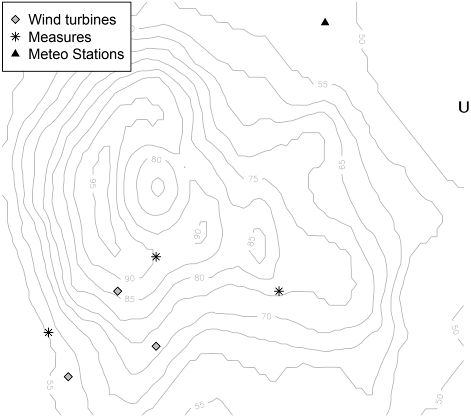
2. Wind farm operational management: the management is highly dependent on the forecasting of wind speed and direction for efficient and safe (regulated) use of the wind farm with storage (at the seconds or hours level), for efficient trading (from 12 hours to 2 days) and for optimal maintenance operations (fews days ahead).

Power of the wind stream $P_w(v) = \frac{1}{2}\rho S v^3$.

Betz' limit of wind turbine power $P_{max}(v) = \frac{16}{27} P_w(v)$.

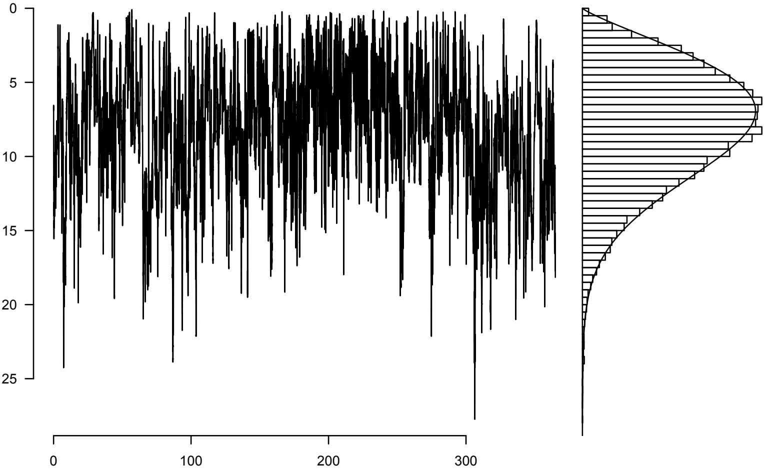
This power is transformed and we consider the power (transfer) function $P(v)$ (cannot reach in practice 70% of the Betz' limit).



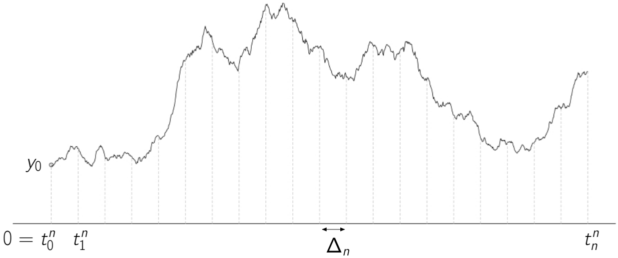


Uncertainties:

- wind uncertainty,
- measures,
- power function,
- vertical extrapolation,
- horizontal extrapolation,
- long-term extrapolation,
- cut-out, unavailability.


0
5
10
15
20
25
0
100
200
300

Let $(Y_t, t \geq 0)$ be the solution of a (fractional) SDE whose law depends on the unknown parameter ϑ.



Our aim is to give asymptotical properties of estimators of ϑ given the observation of the path on a discrete grid $0 < t_1^n < \ldots < t_n^n$, as $n \to \infty$. Asymptotic properties depends on the convergence scheme.

For homogeneous diffusion processes, the loglikelihood is given by

$$\ell(\vartheta, Y^{(n)}) = \sum_{i=1}^{n} \log p^{\vartheta}(\Delta_n, Y_{t_{i-1}^n}, Y_{t_i^n})$$

where $Y^{(n)} = (Y_{t_1^n}, \dots, Y_{t_n^n})$ and the transition densities $p^{\vartheta}(t, x, y)$ satisfies the Fokker-Planck-Kolmogorov equation

$$\frac{\partial}{\partial t} p^{\vartheta} = -\frac{\partial}{\partial y}(v_0(y, \vartheta) p^{\vartheta}) + \frac{1}{2} \frac{\partial^2}{\partial y^2}(v_1(y, \vartheta)^2 p^{\vartheta}), \quad y \in \mathbb{R}, \quad t \in (0, \Delta_n],$$

with initial condition $p(0, x, y) = \delta_x(y)$.

Large sample. Here $\Delta_n = \Delta > 0$ is fixed and, under proper assumptions (smoothness, ergodicity, uniform ellipticity), the LAN property of the likelihoods is satisfied (Roussas 72) with rate $\varphi(n) = \frac{1}{\sqrt{n}}$ and Fisher information matrix equals to

$$\mathcal{I}(\Delta, \vartheta)_{i,j} = \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{\partial}{\partial \vartheta_i} \log p^{\vartheta} \frac{\partial}{\partial \vartheta_j} \log p^{\vartheta} \cdot p^{\vartheta} dy \mu_{\vartheta}(dx) \qquad (1)$$

where μ_{ϑ} is the invariant measure of the diffusion process. Consequently lower bound for the variance of the estimators can be derived, precisely

$$\liminf_{\epsilon \to 0} \liminf_{n \to \infty} \inf_{\widetilde{\vartheta}_n} \sup_{|\vartheta - \vartheta_0| < \epsilon} \mathbf{E}_{\vartheta} \ell \left(\varphi_n^{-1}(\vartheta_0) \left(\widetilde{\vartheta}_n - \vartheta \right) \right) \geq \mathbf{E}_{\vartheta_0} \ell \left(\mathcal{I}(\vartheta_0)^{-1} \xi \right)$$

with $\xi \sim \mathcal{N}(0, I)$ and ℓ is a polynomial cost function.

Mixed scheme: Here $n\Delta_n \to \infty$ and $\Delta_n \to 0$ and the LAN property of the likelihoods has been established (Gobet, 2002) under proper conditions (smoothness, ergodicity, uniform ellipticity) with different rates for ϑ_1 (drift parameter) and ϑ_2 (diffusion coefficient parameter). Namely $\varphi(n)_{1,1} = \frac{1}{\sqrt{n\Delta_n}}$ and $\varphi(n)_{2,2} = \frac{1}{\sqrt{n}}$ respectively and Fisher information matrix is given by

$$\mathcal{I}(\vartheta)_{i,j} = \int_{\mathbb{R}} \frac{\partial}{\partial \vartheta_{1,i}} v_0(y,\vartheta_1) \frac{\partial}{\partial \vartheta_{1,j}} v_0(y,\vartheta_1) \cdot v_1(y,\vartheta_2)^{-2} \mu(dy)$$

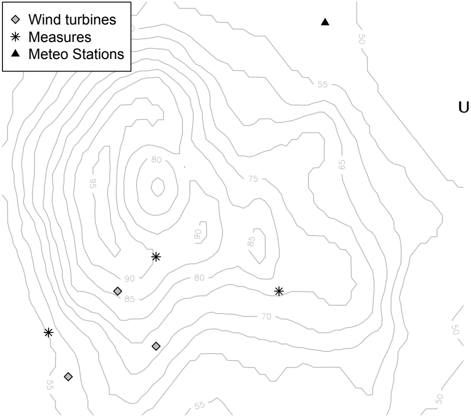
and

$$\mathcal{I}(\vartheta)_{q+i,q+j} = 2\int_{\mathbb{R}} \frac{\partial}{\partial \vartheta_{2,i}} v_1(y,\vartheta_2) \frac{\partial}{\partial \vartheta_{2,j}} v_1(y,\vartheta_2) \cdot v_1(y,\vartheta_2)^{-2} \mu(dy).$$

The YUIMA Project is mainly developed by statisticians who actively publish in the field of inference for stochastic differential equations.

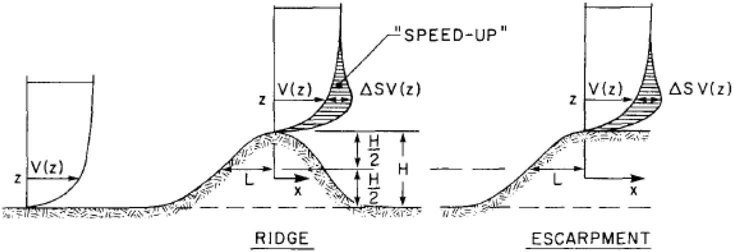
The YUIMA Project Core Team with write access to the source code, currently consists of:

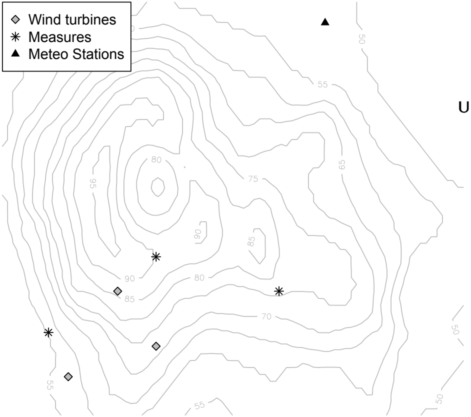
- Alexandre Brouste (Le Mans)
- Masaaki Fukasawa (Osaka)
- Hideitsu Hino (Waseda U.)
- Stefano M. Iacus (Milan)
- Kengo Kamatani (Tokyo)
- Hiroki Masuda (Kyushu U.)
- Yasutaka Shimizu (Osaka)
- Masayuki Uchida (Osaka)
- Nakahiro Yoshida (Tokyo)



Uncertainties:

- wind uncertainty,
- measures,
- power function,
- vertical extrapolation,
- horizontal extrapolation,
- long-term extrapolation,
- cut-out, unavailability.


"SPEED-UP"
V(z)
z
ΔSV(z)
V(z)
z
H/2
H/2
H
L
x
V(z)
z
ΔS V(z)
L
x
RIDGE
ESCARPMENT



Uncertainties:

- wind uncertainty,
- measures,
- power function,
- vertical extrapolation,
- horizontal extrapolation,
- long-term extrapolation,
- cut-out, unavailability.

Incertitude de mesure	1.5 %
Incertitude sur l'extrapolation verticale	3 %
Incertitude reconstitution des données long terme	3.1 %
Incertitude d'extrapolation horizontale	3.1 %
Incertitude de variabilité interannuelle	6%
$\sigma_{p,p}$	8.15 %
Incertitude totale vitesse (méthode EEN)	17.7 %
Incertitude sur la courbe de puissance	5%
Incertitude fonctionnement réel (givre, cut-out,..)	0.5%
Incertitude totale	18.4 %

The general problem of interest relates to estimating the annual production of a windmill, denoted $\mathcal{P}^a$. In particular, it is important to provide the 90% quantile, denoted $P90$, defined by

$$P(\mathcal{P}^a > P90) = 0.90.$$

Practitioners assume $\mathcal{P}^a$ is a Gaussian variable. In that case,

$$P90 = \mu - 1.2815\,\sigma$$

depends only on 2 parameters. In the presintallation approach, we have no production sample on a site.

Let us consider the annual wind power production

$$\mathcal{P}_{annual} = \sum_{s=1}^{S} g(V_s)$$

where $S = 52416$ is the number of subperiods (of 10 min) within a year of 364 days, $(V_1, \ldots, V_S)$ is the wind speed measurements, $v \longmapsto g(v)$ is the wind speed/turbine power relation.

We considered parametric models for the wind speed datasets (France, Hong-Kong, US) : dependent Weibull, Seasonal Weibull, Stationary SDE, ...

Using dependent variable central limit theorem, we obtain that

$$P90 = T \times \mathcal{P}(\vartheta) - 1.2815\sqrt{T \times \mathcal{V}(\vartheta) \times \Gamma_T^2}$$

where

$$\mathcal{P}(\vartheta) = \int_0^{\infty} P(v) f_\vartheta(v) dv, \quad \mathcal{V}(\vartheta) = \int_0^{+\infty} P^2(v) f_\vartheta(v) dv - \mathcal{P}(\vartheta)^2$$

and

$$\Gamma_T = \left(1 + 2\sum_{k=1}^{T} \rho_P(k) \left(1 - \frac{k}{T}\right)\right)^{\frac{1}{2}}.$$

Here T is the number of data, f_ϑ is the Weibull density function, P is the transfer function and $\rho_P(k)$ is the correlation function of the production time series.

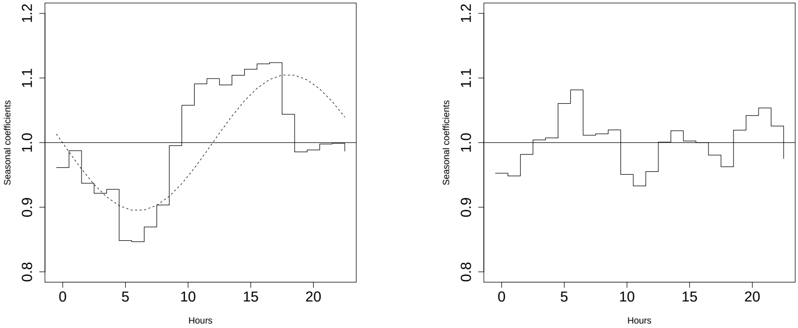
Let us recall that $n = (k-1)L + s \stackrel{\Delta}{=} (s,k)$ where $k = 1,\ldots,K$ is the index of the period (days) and $s = 1,\ldots,L$ the index of the subperiod (hours for instance). We assume, for the wind speed time series, the following model

$$V_{(s,k)} \sim \mathcal{W}(\lambda_{(s,k)}, p), \quad \lambda_{(s,k)} = a\,g(k) + s_s\,b,$$

where $a, b \in \mathbb{R}$ are unknown constants, $s = (s_s, s = 1,\ldots,L)$ are the unknown seasonal factors, the shape parameter p is known and the trend function $g(.)$ is also known. We have moreover

$$\sum_{s=1}^{L} s_s = L$$

and $\vartheta = (a, b, s) \in \Theta \subset \mathbb{P} \subset \mathbb{R}^{L+2}$.


Seasonal coefficients
0.8
0.9
1.0
1.1
1.2
0
5
10
15
20
Hours
Seasonal coefficients
0.8
0.9
1.0
1.1
1.2
0
5
10
15
20
Hours

The type of problem to be faced by the provider is :

1. to regulated power provided to the grid and optimize the (second, minutes) storage ;
2. to guarantee a given amount of energy, for a given duration, at a given price for optimal trading ;
3. to manage the maintenance operations.

Wind model for short-term forecasting is a key element in the operational management of a wind farm.

The type of problem to be faced by the provider is :

1. to regulated power provided to the grid and optimize the (second, minutes) storage ;
2. to guarantee a given amount of energy, for a given duration, at a given price for optimal trading ;
3. to manage the maintenance operations.

Wind model for short-term forecasting is a key element in the operational management of a wind farm.

Lots of studies of dynamical models for forecasting:

- statistical models (times series, neural networks, sde ...) for seconds, minutes, hours ;
- meteorological models for days and week.

For the moment, we are studying sde models and consider their evaluation on a topography (with fluid mechanics (approximate linear) formulas). It allows in the future to connect correlated it with the meteorological models and stochastic controls problems for storage and trading.

1– A. Bensoussan, P. Bertrand and A. Brouste (2012) *Forecasting the energy produced by a windmill on a yearly basis*, Stochastic Environmental Research and Risk Assessment, 26(8), 1109–1122.

2– A. Bensoussan, P. Bertrand, A. Brouste, N. Haouas, M. Fhima and D. Koulibaly (2014) *Confidence intervals for annual wind power production*, ESAIM Proceedings, 44, 150–158.

3*– A. Bensoussan, P. Bertrand, A. Brouste and N. Haouas (2014) *Impact of Seasonality on interquartile range for annual wind power production*, submitted.

4– Bensoussan A., Bertrand P. and Brouste A. "A generalized linear model approach to seasonal aspects of wind speed modeling", Journal of Applied Statistics, 41(8), 1694-1707, 2014.

5– Bensoussan A., Bertrand P. and Brouste A., "Estimation theory for GLM in Future Perspectives" in Risk Models and Finance edited by A. Bensoussan, D. Guegan and C. Tapiero, Springer-Verlag.

6– A. Brouste, M. Fukasawa, H. Hino, S. Iacus, K. Kamatani Y. Koike, H. Masuda, R. Nomura, Y. Shimuzu, M. Uchida and N. Yoshida (2014) *The YUIMA Project : a Computational Framework for Simulation and Inference of Stochastic Differential Equations*, Journal of Statistical Software, 57(4), 1–51