# A simple scenario generation method

Tue Vissing Jensen

## I. Scenario generation

Denote by $\hat{w}_{ntk}$ the point forecasts of wind production for time $t$ at node $n$ given $k$ hours ahead, and by $w_{nt}$ the observed production. Using this data, and given some new point forecast $\vec{x}$, we wish to construct scenarios for the observed power $\vec{y}_\omega$. As we are not interested in time coupling, we proceed by separately considering the temporal and spatial characteristics of the signal. For simplicity, we assume $y_n \in [0,1]$, and that the generated scenarios will be scaled appropriately after generation.

Define the $k$-hour ahead marginal conditional probability of seeing the production $y_n$ at node $n$ given the point forecast $x_n$ as

$$f_{nk}(y_n|x_n) = \frac{\sum_t K(x_n - \hat{w}_{ntk})K(y_n - w_{nt})}{\sum_t K(x_n - \hat{w}_{ntk})}, \quad (1)$$

where $K$ is a kernel function, here taken to be a Gaussian kernel of bandwidth $\theta = 0.01$:

$$K(x) = \frac{1}{\theta}\exp\left(-x^2/(2\theta^2)\right). \quad (2)$$

To simplify this expression, we compute its conditional mean and standard deviation as

$$\mu_{nk}(x_n) \equiv \int_Y y_n f_{nk}(y_n|x_n) dy_n, \quad (3)$$

$$\sigma_{nk}(x_n) \equiv \sqrt{\int_Y (y_n - \mu_{nk}(x_n))^2 f_{nk}(y_n|x_n) dy_n}, \quad (4)$$

and let our conditional forecast $Y_n(x_n)$ be Beta distributed random variable with the same mean and standard deviation,

$$Y_n(x_n) \sim \text{beta}(\mu_{nk}(x_n), \sigma_{nk}(x_n)) \equiv g_{nk}(x_n). \quad (5)$$

To extract the spatial dependence of the signal, let

$$r_{nt} = \Phi^{-1}\left(M_n(w_{nt})\right), \quad (6)$$

where $M_n$ is the empirical CDF of $w_{nt}$ and $\Phi$ is the standard normal CDF. Define the cross-correlation matrix $\Sigma$ via

$$\Sigma_{nm} = \frac{1}{T}\sum_t (r_{nt} - \bar{r}_n)(r_{mt} - \bar{r}_m), \quad \bar{r}_n = \frac{1}{T}\sum_t r_{nt}, \quad (7)$$

where $n$ and $m$ are nodes.

We now construct scenarios as follows: Given a point forecast $\vec{x}$ for wind production $k$ hours ahead, extract a sample from the correlated standard normal distribution for each scenario $\omega$

$$\vec{r}_\omega \sim\sim \mathbf{N}(\vec{0}, \Sigma). \quad (8)$$

Transform this sample through the conditional marginal distribution to obtain the prediction:

$$w_{n\omega} = G_{nk}^{-1}\left(\Phi\left(r_{\omega n}\right)|x_n\right) \quad (9)$$

where $G_{nk}(\cdot|x_n)$ is the CDF corresponding to $g_{nk}(x_n)$ in eq. (5).

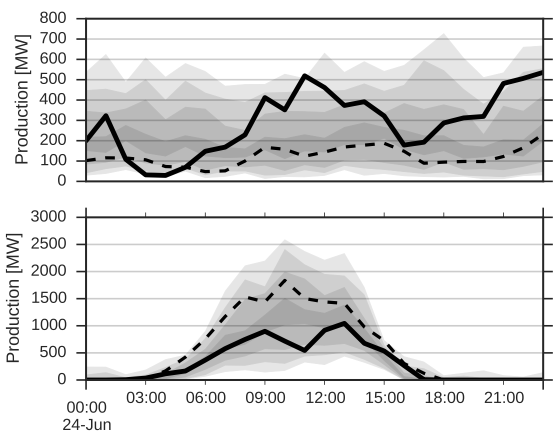


Figure 1. Example wind (top) and solar (bottom) point forecasts (dashed line) and observations (solid line) for node 1124, located in eastern Poland. Shaded bands indicate symmetric 5, 10, 20 and 40% quantile bounds from 100 scenarios. Production is scaled to 100% gross penetration for each resource.

The scenarios constructed in this way will have the correct spatial correlation structure, and their marginal distributions will match those of historical data. They will, however, be uncorrelated in time. This is not a problem for this application, as the market-clearing models do not contain time-coupled terms. Fig. 1 shows examples of the constructed scenarios.