Developing a wind and solar power data model for Europe with high spatial-temporal resolution

Ingeborg Graabak  
NTNU, Norway  
ingeborg.graabak@ntnu.no

Harald Svendsen  
SINTEF, Norway  
harald.svendsen@sintef.no

Magnus Korpås  
NTNU, Norway  
magnus.korpas@ntnu.no

Abstract—This paper describes a wind and solar power production model for Europe based on the numerical weather prediction model COSMO-EU. The COSMO-EU model has hourly time resolution and a spatial resolution of 7 km x 7 km for Europe. The model is validated against power production information from the system operators in Denmark, Germany and Spain. Mean Average Error (MAE) (hourly error averaged for a year) relative to the wind installed capacity is in the range 4.9%-5.9% for wind power production and 2.4%-5.5% for PV (photovoltaic) power production. Root Mean Square Error is in the range 6.2%-7.6% and 4.5%-9.3% for wind and PV power production respectively. The results are compared with similar modelling based on wind and radiation data from the NCEP reanalysis model. This model has six hourly time resolution for wind resources and daily resolution for radiation data. Modelling of wind power production in Denmark, German and Spain has a MAE in the range 5.6%-8.5% and solar PV production 4.9%-6.4% for the NCEP reanalysis model.

Index Terms—Wind power generation, solar power generation, power system simulation, power system planning.

I. INTRODUCTION

The future power system in Europe will probably include large shares of wind and solar resources. Reduction of conventional capacities based on fossil fuels is likely due to long periods with low power prices and because of the need to reduce emissions of greenhouse gases. Wind and solar resources are variable. If most of the fossil fuel plants are decommissioned, other measures are necessary to balance the variability in the power production.

It is important to understand the variability characteristics of wind and solar resources to be able to develop the future power system in a cost-efficient manner. The variability is dependent on the location of the resources. Thus, it will be very useful to have a data model of wind and solar resources to be able to simulate the variability of a future power system. Based on the simulations, it is possible to obtain increased understanding of the variability and to plan for mitigation of it. In the power system, the net load must be supplied in all times steps. The net load is the load minus the production from wind and solar resources. In the present power system the net load is balanced by dis-patchable power plants. In the future power system it is likely that demand response, different types of storage and more grids will reduce the needs for balancing by conventional power plants.

It is reasonable that a wind and solar resource model with high spatial and temporal resolution. This paper describes the development of a European wind and solar resource model with hourly resolution and spatial resolution of 7 km x 7 km. Furthermore, the present installed capacities of wind turbines and PVs (photovoltaic) are modelled in detail in order to be able to study the variable characteristics of the present system. By upscaling the power production capacities, the future variability can be studied for different configurations of production plants. The developed model are verified by comparison with real production data from the TSOs (Transmission System Operator) in Denmark (DK), Germany (GE) and Spain (ES).

To our knowledge, there are few papers about wind and solar power production models for Europe. One paper describes a bottom-up approach for modelling hourly electricity output based on meteorological data and technical specifications for different reference plants [1]. Since hourly time series for wind and solar power production were not available when the paper was written, a simplified approach was used for validation.

Wind and solar data models are to some degree described in largescale studies of the future European power system. However, according to [2] there are few largescale studies of the future European power system available. The paper describes a study of a future European system with 50% wind and solar power production. The study uses NASA reanalysis data consisting of hourly values of wind speed and solar irradiance at a spatial resolution of 0.5° East/West and 0.66° North/South. NASA reanalysis data has time resolution of 6 hours, so the hourly data has to be constructed. According to [3] previous wind power production models are mainly based on two approaches:

- Up scaling production of single wind farms to cover the real installed capacity of a designated area
- Utilization of wind speed reanalysis data with a relatively wide grid of wind speed data nodes and long term time intervals.

Both approaches have drawbacks in production and/or time resolution accuracy. Due to area clustering, the influence of wind and solar power production on the system simulations might be misrepresented.

The references [3] and [4] describes a wind model for Europe with hourly resolution for the years 2006-2013. Aigner also partly developed a solar PV model for Europe [5]. The work described in [3]-[5] provided a starting point for further extension and improvement by us.
II. MODELLING OF WIND AND SOLAR RESOURCES

The wind and solar resources are from the numerical weather prediction model COSMO EU (Consortium for Small Scale Modelling) [6]. The model is based on thermal-hydrodynamic equations describing the compressible flow in a moist atmosphere. The technical reports on the COSMO website provides detailed description of the model including the basic model design, dynamics, physical parametrization and the data assimilation.

The version called COSMO-EU is covering the whole Europe, eastern Atlantic and northern Africa. The modelling routine simulates a meshed data grid with a point-to-point resolution of 7 km x 7 km. The resulting 665 x 657 data nodes include among other wind, radiation and temperature data with a time resolution of one hour.

The COSMO data provided by the German meteorological office is stored in a rotated coordinate system. Therefore, a data conversion is necessary to obtain the wind speeds and the radiation data in a spherical system, which can be further processed during the wind and PV power simulations. Pole rotation is a consequence of numerical convergence problems of the meridians resulting in pole singularities in the spherical coordinate system. Therefore, the pole is tilted and transferred so that the equator runs through the centre of the model domain. This allows a minimization of convergence problems for any model domain. Reference [3] describes the coordinate transformation.

A. Conversion to wind power production

Since the COSMO wind speed data is measured in a height of 10 m, a logarithmic scaling is necessary to obtain the wind speed at wind turbine hub height, according to (1). The surface roughness length, describing the roughness characteristic of the terrain, is considered in the scaling process of each wind speed data point:

\[
H_f = \frac{\log_{10}(h_{ref} / z_0)}{\log_{10}(h_{mes} / z_0)}
\]

(1)

where:
- \(H_f\): Scaling factor
- \(h_{ref}\): Reference hub height
- \(Z_0\): Surface roughness length
- \(h_{mes}\): Measurement height

According to (2) the wind speed velocity at hub height \(V_{10}\), called meso wind, is the product of the scaling factor \(H_f\) and the measured wind speed velocity \(V_M\).

\[
V_{10} = H_f V_M
\]

(2)

Each single data grid point is scaled up to hub height according to (1) and (2), taking the respective surface roughness \(Z_0\) (provided by COSMO EU) into account. The geographical wind farm coordinates can now be implemented into the converted COSMO data grid. The wind speeds for each wind facility are interpolated from the surrounding wind speed data points.

The wind speed to power conversion is based on two different turbine power curves, one for turbines with installed capacity less than 10 MW and one for turbines larger than 10 MW (see Fig.1).

Reference [4] provided a starting point for the quantification of the turbine curves. The curves where further improved by simulations and comparison with TSO data and minimization of the Mean Average Error (MAE) and the Root Mean Square Error (RMSE) (see below).

B. Conversion to PV production

Three radiation terms are necessary for the calculation of the overall irradiation hitting the solar panel [5],[7]. These include direct, diffuse and reflected radiation. For being able to calculate the total effective irradiation on a tilted surface like a PV panel, the angular interactions between the sun, the local time and the tilt angle of the surface have to be taken into account. Calculation of the total irradiance is done as in [8]:

\[
I_{tot,p} = I_{direct,p} \cos(\theta_p) + I_{diff,p} \frac{1 + \cos(\alpha)}{2} + I_{diff,p} \frac{1 - \cos(\alpha)}{2} RE
\]

(3)

where:
- \(I_{tot,p}\): Total irradiance on panel \(p\) [W/m²]
- \(I_{direct,p}\): Direct irradiance on panel \(p\) [W/m²]
- \(I_{diff,p}\): Diffuse irradiance on panel \(p\) [W/m²]
- \(RE\): Reflection coefficient
- \(\alpha\): Tilt angle between the normal of the horizontal surface and the normal of the tilted surface.
- \(\theta_p\): Angle between the normal to the tilted panel and the sun’s ray (see Fig.2).

![Fig.2 Angles of tilted surface [7]](image-url)
The tilt angle \( \alpha \) becomes 90 \(^\circ\) for a vertical wall and zero for a horizontal plane. In the simulations, a tilt angle of 40\(^\circ\) has been assumed for the PV installations in Denmark and Germany. This corresponds to the average roof pitch in Germany. The validation process for Spain showed improved results with an angle of 30\(^\circ\) instead of 40\(^\circ\).

Due to a lack of information, a reflection coefficient of 0.27 has been assumed which represents the surface of a dark building. The chosen value corresponds to a mean value between a light building surface with a RE of 0.73 and blank earth with an RE of 0.14 [5]. Since the ambient temperature influences the panel efficiency it has to be considered in the simulation of the overall power production. The temperature dependent production for each PV panel therefore becomes [9]:

\[
P_{\text{prod},p} = I_{\text{bat},p} \text{cap}_p \eta_{\text{ref}} [1 + \varepsilon(T_p - T_{\text{ref}})]
\]

where:
- \( P_{\text{prod},p} \): Total panel production
- \( \text{cap}_p \): Installed capacity of panel \( p \)
- \( \eta_{\text{ref}} \): Reference efficiency, how much of the generator power in the PV panel that reaches the grid
- \( \varepsilon \): Solar radiation coefficient
- \( T_p \): Ambient temperature for panel \( p \)
- \( T_{\text{ref}} \): Reference temperature 25 \(^\circ\) C

The Reference efficiency is set to 0.15, which is a typical value for Mo-Si solar panels [9]. Furthermore, the Solar radiation coefficient is set to – 0.0035 [9].

III. MODELLING OF WIND AND SOLAR POWER PRODUCTION IN DENMARK, GERMANY AND SPAIN

A. Modelling of wind power production

Wind power production is modelled based on data from The Windpower database [10]. This database provides information about most of the wind farms in Europe. The information includes among other the following data for each wind farm: longitude and latitude, installed capacity, number and type of turbines and commissioning date. Fig. 3 to the left shows the location of the wind farms used in the COSMO simulations as red dots. The blue stars are location of aggregated wind power production capacities used in the reanalysis simulations. The wind power capacity in the simulations is 52% of the total installed wind power capacity in Europe by the end of 2015 (141.579 MW [14]).

B. Modelling of PV power production

PV production for Denmark was based on information about installed PV capacity for 638 zip codes [11]. The codes where aggregated to 500 geographical locations. The “EEG-Anlagenstammdaten-Gesamtdeutschland zur Jahresabrechnung 2014” gives detailed information about PV installations in Germany in 2014 [12]. The register provides information about 1.535,203 PV plants. The information includes among other installed capacity and PLZ (PostLeitZahl – postal codes) codes for Germany. The information was aggregated to 668 PLZ codes for Germany. The PLZ codes were geocoded to latitudes and longitudes. The Reference [13] provides information about installed PV capacity for 15 regions in the Peninsular Spain for 2014, and this information was used in the simulations. I.e., simulations of the wind power production in Denmark, Germany and Spain is built on very exact information about the wind farms, while information about PV installations are aggregated to regions per country.

IV. VALIDATION OF THE MODEL

The model is validated against TSO data for wind power production in 2015 and PV power production in 2014 [11], [13] and [16]. Fig.4 shows the wind power production for the first 1000 hours of 2015 in Germany. Fig. 5 shows the PV production for Spain for the hours 2200-4400 in 2014. Table 3 shows validation of wind power production. Table 4 shows separate validation for onshore and offshore wind power production. Table 5 shows validation results for onshore and offshore wind power production in Denmark, while Table 5 shows validation results for PV production in Spain.
for PV power production. The following abbreviations are used in the tables: sim – simulations, onsh- onshore, offsh – offshore and prod – production.

Fig.4 Simulated (COSMO) and TSO wind power production GE hour 0-1000 2015

Fig.5 PV power production for ES hour 2200-4400 2014

MAE is the sum of the difference between the simulated value and the TSO value for each hour in the year divided on number of hours in the year. Calculation of the RMSE $[\text{MWh/h}]$ is as follows:

$$ RMSE = \sqrt{\frac{1}{8760} \sum_{j=1}^{8760} (\text{TSO}(j) - \text{COSMO}(j))^2} $$

$(5)$

$\text{TSO}(j)$: real power production in hour $j$

$\text{COSMO}(j)$: simulated power production in hour $j$

Fig.6 shows the delta production for the wind power production in Denmark in 2015. The delta production is the change in production from one hour to the next hour divided on the installed capacity for each hour in the year and sorted from lowest to highest.

Wind power production was also simulated for Denmark, Germany and Spain by aggregating installed capacity to one point for each country and by using the COSMO data for these points. This approach resulted in MEA/installed capacity of 12.1%, 12.4% and 13.9% for Denmark, Germany and Spain respectively. The reason for the low accuracy is that the wind resources are based on one single point in each country, and these points represent only the available wind resources in the given point and not an average value for a larger region.

Compared with previous work, a model of wind power production in North Ireland (290 MW installed capacity) using MERRA reanalysis data got a RMSE of 11.9% [17]. A modelling of Swedish wind power production also using MERRA reanalysis data got an RMSE of 3.8% [18]. Two factors explaining the good results for the last study were the

<table>
<thead>
<tr>
<th>TABLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VALIDATION OF COSMO WIND POWER PRODUCTION FOR 2015</strong></td>
</tr>
<tr>
<td><strong>Total production [TWh/year]</strong></td>
</tr>
<tr>
<td>DK TSO</td>
</tr>
<tr>
<td>Sim.</td>
</tr>
<tr>
<td>GE TSO</td>
</tr>
<tr>
<td>Sim.</td>
</tr>
<tr>
<td>ES TSO</td>
</tr>
<tr>
<td>Sim.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VALIDATION OF COSMO ONSHORE AND OFFSHORE WIND POWER PRODUCTION FOR DENMARK</strong></td>
</tr>
<tr>
<td><strong>Total prod [TWh/year]</strong></td>
</tr>
<tr>
<td>TSO onsh</td>
</tr>
<tr>
<td>TSO offsh</td>
</tr>
<tr>
<td>Sim onsh</td>
</tr>
<tr>
<td>Sim offsh</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VALIDATION OF COSMO PV POWER PRODUCTION FOR 2014</strong></td>
</tr>
<tr>
<td><strong>Total production [TWh/year]</strong></td>
</tr>
<tr>
<td>DK TSO</td>
</tr>
<tr>
<td>Sim.</td>
</tr>
<tr>
<td>GE TSO</td>
</tr>
<tr>
<td>Sim.</td>
</tr>
<tr>
<td>ES TSO</td>
</tr>
<tr>
<td>Sim.</td>
</tr>
</tbody>
</table>

Fig.6 Delta production for the wind power production in DK for the COSMO simulated results (blue curve) and the TSO data sorted (red curve).
use of a globally optimised power curve smoothing parameter and correction of seasonal and diurnal bias.

V. COMPARISON WITH REANALYSIS DATA

Wind and PV power production in Denmark, Germany and Spain is calculated based on reanalysis data. Wind speed time series and irradiation time series are from NCEP reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder Colorado USA from their Web site [19]. The Reanalysis wind data set has a temporal resolution of 6 hours and a spatial resolution of 2.5 degrees in both latitude and longitude.

The wind energy is computed from the wind speed using the same method as in the TradeWind project [20]. In order to get hourly time series for wind speed, a linear interpolation of the 6-hourly values has been applied. Since the wind speed is the average and smoothed out wind speed for a wide area, and because the wind energy represents many wind turbines, a regional power curve is used for computation. Fig. 7 shows the normalised power curve.

Often there will initially be significant discrepancy between the computed wind and the actual wind energy with the use of coarse data. To correct for these discrepancies a constant adjustment factor (per point) on the wind speed series has been used such that historical capacity factors for 2011 were matched as close as possible. Table 6 shows the adjustment factors.

The ratios of hourly to daily total radiation \( r_t \) and hourly to daily total diffuse radiation \( r_d \) are then given by the empirical expression:

\[
\begin{align*}
 r_t^{(h)} &= \frac{H_t}{24(a + b \cos \omega_t)} f_{\text{crp}} \\
 r_d^{(h)} &= \frac{H_d}{24(a + b \cos \omega_t)} f_{\text{crp}}
\end{align*}
\]

Then the hourly global radiation \( H \), diffuse irradiation \( H_d \) and beam radiation \( H_b \) at a particular hour of the day is

\[
H_t = r_t(h)H_b, \\
H_d = r_d(h)H_d, \\
H_b = H - H_d.
\]

The total irradiance is calculated as (3) and the total production from the PV panel is calculated as in (4) except that there is no corrections for temperature variations. Table 6 shows aggregation of production capacities and the adjustment factors calculated to match yearly production for a region. Table 7 shows validation of the simulation results. Fig. 8 shows simulations results for both COSMO and reanalysis data compared with TSO data for Denmark hour 0-500 for 2015.

![Wind speed curve](image)

**Fig. 7 Regional power curve used in reanalysis simulations [20], [21].**

The reanalysis radiation data set covers 1948-today [22]. The irradiation time series provides the following daily average values with a spatial resolution of 1.875° in longitude and 1.904° in latitude:

- \( H_{by} \): visible beam downwards solar flux (W/m²)
- \( H_{bn} \): near infrared beam downward solar flux (W/m²)
- \( H_{dy} \): visible diffuse downwards solar flux (W/m²)
- \( H_{dn} \): near infrared diffuse downward solar flux (W/m²)

Visible and near infrared irradiation is added and multiplied by 24 hours to give daily beam, diffuse and global irradiation:

\[
\begin{align*}
\bar{H}_b &= 24 \times (H_{by} + H_{bn}) \\
\bar{H}_d &= 24 \times (H_{dy} + H_{dn}) \\
\bar{H} &= \bar{H}_b + \bar{H}_d
\end{align*}
\]

A daily profile is computed from the obtained daily average values, taking into account time of year, time of day, and latitude and longitude. The Collares-Pereira-Rabl (CPR) factor \( f_{\text{crp}} \) and the parameters \( a \) and \( b \) are calculated as follows:

\[
\begin{align*}
\begin{cases}
0, & \omega > \omega_s \\
\frac{c}{\sin \omega - \omega_s c}, & 0 \leq \omega < \omega_s \\
0, & \text{otherwise}
\end{cases}

f_{\text{crp}}^{(h)} = \frac{\omega - \omega_s}{\sin \omega - \omega_s c} \\
an = 0.409 + 0.5016 \sin (\omega_s - R/t) \\
b = 0.66609 + 0.4767 \sin (\omega_s - R/t)
\end{align*}
\]

The ratios of hourly to daily total radiation \( r_t \) and hourly to daily total diffuse radiation \( r_d \) are then given by the empirical expression:

\[
\begin{align*}
 r_t^{(h)} &= \frac{H_t}{24(a + b \cos \omega_t)} f_{\text{crp}} \\
 r_d^{(h)} &= \frac{H_d}{24(a + b \cos \omega_t)} f_{\text{crp}}
\end{align*}
\]

The total irradiance is calculated as (3) and the total production from the PV panel is calculated as in (4) except that there is no corrections for temperature variations. Table 6 shows aggregation of production capacities and the adjustment factors calculated to match yearly production for a region. Table 7 shows validation of the simulation results. Fig. 8 shows simulations results for both COSMO and reanalysis data compared with TSO data for Denmark hour 0-500 for 2015.
VI. DISCUSSIONS

Wind and solar PV power production is modelled for Denmark, Germany and Spain based on two different models: the COSMO EU model with hourly resolution and a spatial resolution of 7 km x 7 km and a NCEP reanalysis model with six hourly resolution for wind resources and daily resolution for solar resources.

For wind power production simulated by the COSMO data the MAE/RMS are 5.5%/7.1% for Denmark, 4.9%/6.2% for Germany and 6.0%/7.5% for Spain. All numbers are in percent of installed wind power capacity in the respective country. The corresponding percentages are 3.9%/7.6% for Denmark, 2.4%/4.5% for Germany and 5.5%/9.3% for Spain for PV power production. One obvious contribution to the better results for PV power production is that for many hours through a year the MAE/RMS will be zero due to nights and dark winter months. For both the COSMO and the reanalysis simulations, the results for Spain has lower accuracy than for Denmark and Germany. For wind, one reason is that it is more difficult to model the varying terrain in Spain than the flatter landscape in Denmark and the northern part of Germany (Z0 in (1)).

As expected, the simulation results from the COSMO EU model is closer to the TSO data than the NCEP reanalysis data for all simulated cases except for PV power production in Spain where simulations based on reanalysis data is slightly more accurate than COSMO. The reason for this is probably aggregation of installed PV capacities to represent a large area, while the COSMO resource data are not averaged to represent the same region.

Simulations of onshore wind power production in Denmark seems to be closer to the real production than offshore production. One possible reason for this is that there are larger turbines in the offshore power production. I.e. deviations in the turbine curve from the real curve will have larger impact per point than for onshore production. It is also probably possible to reduce the deviations both for the onshore and offshore simulations by better tuning of the power curves.

All in all, both COSMO and reanalysis simulations follow the real production curves quite well. Even though there are deviations between the simulated and real production, the COSMO model represents the changes in power production from one hour to another very close to the real production (Fig.6), and this is important in studies of how to balance future variable wind and PV power production (ramps).

Further improvement of the wind power production simulations are probably possible e.g. by more advanced modelling of the turbine curves. Another possibility is to investigate the wind and PV power simulations for diurnal or seasonal bias compared to the real production. Further work should focus on these topics.

ACKNOWLEDGEMENTS

We acknowledge Tobias Aigner's PhD work. His work is a basis for the COSMO results in this paper. Furthermore, we would like to thank Deutsche Wetterdienst for providing us wind, radiation and temperature data.

REFERENCES

[12] www.netztransparenz.de
[14] www.eewa.org
[16] www.eex.com