Representation of wind power generation in economic models for long-term energy planning

Eimantas Neniškis,

Arvydas Galinis

Lithuanian Energy Institute, Breslaujos St. 3, 44403 Kaunas, Lithuania Email: Eimantas.Neniskis@lei.lt The representation of wind power plants electricity generation in economic models for energy planning is problematic, since generation in wind power plants is variable and it is not possible to predict accurately enough wind fluctuations for more than a few days. Often, wind power plants generation patterns from a single historical year are repeated throughout the modelled time period. Typically, this method is used when analysing power system operation in hourly time intervals for all days in a year for each year. However, hourly time resolution is not feasible in large-scale models, which tend to require considerable amounts of computing power. Thus, some kind of time aggregation is needed. On the other hand, currently used methods for models with less than hourly temporal resolution are becoming inadequate because of increasing share of fluctuating electricity production from wind in total power generation.

In this article, a methodology for evaluation of wind power plants stochastic electricity generation in power system development models is described. The methodology is based on evaluation of how much time single or multiple wind power plants generate a certain output range during a season or some time period within a year, modelling of output distribution for a typical day of the selected time period, and preparation of electricity output curves. These electricity output curves when modelling wind power plants in models with aggregated time allow to assess fluctuations in generation, observable regularities and enhance the objectivity of balancing power demand assessment, also ensure that electricity generated during a typical day corresponds to electricity generated during the selected time period. This methodology will help to determine the rational perspective power generation mix more accurately and make a better assessment of cost-effectiveness of wind power plants in economic models for energy planning, without significantly increasing the size of already large-scale models.

Keywords: wind, stochastic, power system, development, economic model

INTRODUCTION

Oil crisis in 1970s sparked the demand for energy supply and demand forecasts, energy planning and modelling, since every oil-importing country wanted to ensure the security of energy supply. After the end of oil crisis, when oil market reached an equilibrium, a sustainable energy development and reduction of global warming were a new focus for energy models [1]. On 4 November 2016 a Paris Agreement entered into force, in which countries agreed to keep global temperature rise well below 2 degrees Celsius above pre-industrial levels and submitted nationally determined contributions [2]. Economic models for long-term energy planning like TIMES [3], MESSAGE [4], BAL-MOREL [5], PRIMES [6] are necessary for each country to project optimal technology pathways that meet national targets, with the lowest technical and financial risk [7].

In these models the chosen temporal resolution significantly impacts expected electricity

generation from new Renewable Energy Sources (RES) and capacity which can be installed. When using lower temporal resolution wind fluctuations are smoothed and the model tends to overestimate how much new wind capacity should be installed [8, 9]. In order to adequately describe the stochastic nature of wind, high temporal resolution is needed. However, when using high temporal resolution a model becomes much more complex and requires more computing power/time to solve. Stefan Pfenninger [10] compared CPU time required to solve a Calliope energy modelling framework model between different temporal resolutions (8760, 2920, 1460, 730 and 365 time steps in a year resolution) and 3 scenarios: in the first scenario 90% of demand is covered by generation from stochastic RES; the second is like the first scenario, but includes storage; in the third scenario 50% of demand is covered by generation from RES. Results are shown in Fig. 1.

There is a trade-off between the level of detail and ease of solving. For large models it is



Fig. 1. Normalized CPU time and installed offshore wind capacity from 2014 runs with uniformly downsampled time resolution. On the machine used, the 1-hourly run (maximum run time) took about 7.2 h. All CPU times are normalized by dividing by that value [10]

desirable to reduce time steps without significantly reducing accuracy. This can be done by representing each year by typical demand days for each month or season based on average parameter values. One season/month can be represented by several typical days, which corresponds to some characteristic periods, like peak-demand and off-peak-demand days (e.g. weekdays and weekends) [11]. However, it is problematic to represent stochasticity of electricity generation in wind power plants when averaging data to produce a typical day, because it smooths wind fluctuations. It is possible to calculate how many hours wind power plants should produce certain power during a typical day, but it is unknown what power output should be at a specific hour in a typical day. The aim of this article is to describe a methodology on how to create wind curves for each of these typical days.

METHODOLOGY

This methodology is based on the creation of randomized electricity generation from wind curves for each season and day type. In order to do this, data on how many hours wind power plants generate a certain power range is needed for each season. This data can be calculated in several ways:

It can be calcu lated from hourly or higher temporal resolution data of actual electricity generation from wind. This data usually is publicly provided by Transmission System Operators (TSOs). In this method, actual electricity generation from wind data is split according to seasons or other time periods. Time of a certain generated power range p for each season is calculated using this formula:

$$h_{p,t1-t2} = T \sum_{j=1}^{J} \begin{cases} 1, if PR1_p \le x(j) < PR2_p \\ 0, otherwise. \end{cases}$$
(1)

Here $h_{p, t1-t2}$ is the number of x(j), j = 1, J wind power data points (measurements) in time period t_1-t_2 that satisfy $PR1_p \le x(j) < PR2_p$ inequality; x(j) is wind power in particular measurement jwithin time period t_1-t_2 in MW; T is time interval between wind power measurements (it has to be no longer than length of the smallest time slice in the economic model; measurements in this case are also assumed to be made in equal time intervals); p-1, P is the index of wind power range; $PR1_p$ is the lowest power range p value in MW; $PR2_p$ is the highest power range p value in MW; t_1 is hour in a year of season beginning, t_2 is hour in a year of season ending.

Time values according formula (1) have to be calculated for all power ranges $p \in P$.

Another way is to calculate it from hourly or higher temporal resolution data of wind speeds. By using the wind turbine power curve it is possible to convert wind speed to power output; however, it is needed to recalculate wind speeds to a hub height first. Wind profile power law describes relationship between wind speeds on two different heights.

$$v_2 = v_1 \left(\frac{z_2}{z_1}\right)^{\alpha}.$$
 (2)



Fig. 2. Electricity generated from wind in Lithuania 2017 [12]

Here v_1 [m/s] and v_2 [m/s] are wind speeds at height z_1 [m] and z_2 [m], correspondingly, α is the wind shear coefficient. The value of α depends on terrain type: for lakes or smooth ground $\alpha = 0.1$, for tall crops and shrubs $\alpha = 0.2$, for small town $\alpha = 0.3$, for city $\alpha = 0.4$ [13].

After wind speed conversion into power output of wind turbine, data should be split according to the time of a certain generated power range for each season calculated using formula (1) as in the first way.

When it is known how much time wind power plants generate at each power range in all seasons, then electricity production for each particular season is calculated as a sum of all power output data points within the season multiplied by duration of time between measurements x(j) in hours *T*. In the case time slices *j* are taken from the set $J_k \in J$, representing time slices of all days belonging to a typical day category $k \in K$ in the season t_1-t_2 , electricity generated during these days $(E_{i-t_2}^k)$ can be calculated as it is shown in formula (3).

$$E_{t_{1-t_{2}}}^{k} = T \cdot \sum_{j \in J_{k}} x(j).$$
(3)

 $E_{typ,t1-t2}^{k}$ – electricity generated from wind during a single day of typical day category k in the season t_1-t_2 is calculated by dividing electricity generated during these typical days in the season by the number of days of category k in that season.

$$E_{typ, t^{1-t^2}}^k = \frac{E_{t^{1-t^2}}^k}{days_{t^{1-t^2}}^k}.$$
(4)

The time when wind power plants generate a certain power range p in a typical season day of category k ($h \downarrow (typ, p, t1-t2)\uparrow k$) is equal to the time when wind power plants generate a certain power range in typical days k in a season divided by the number of category k days in that season. Results of such calculations will be used for creation of curves for electricity generation from wind for typical days of each season.

$$h_{typ, p, t^{1-t^{2}}}^{k} = \frac{h_{p, t^{1-t^{2}}}^{k}}{days_{t^{1-t^{2}}}^{k}}.$$
(5)

Results have to be adjusted in a way that a sum of $h_{typ, p, t1-t2, adi}^k$, which represents the time of

wind power plants operation at all possible power ranges *p* in a typical day *k*, would be equal to L_k – the number of time slices into which typical day *k* is divided in the model. This adjustment can be made by multiplying $h_{typ, p, t1-t2}^k$ by $\frac{T}{\tau}$ where τ is the length of the smallest time slice in an economic model. Then each value is rounded to integer.

$$h_{typ,p,t\,1-t\,2,adj}^{k} = \left[k_{typ,p,t\,1-t\,2}^{k} \cdot \frac{T}{\tau}\right]$$
(6)

$$\sum_{p=1}^{p} h_{typ,p,t1-t2,adj}^{k} \cdot \tau \to E_{typ,t1-t2}^{k}.$$
(7)

The sum of multiplications of each power range average value and corresponding number of hours should be as close as possible to calculated electricity production from wind during a typical day in the season t_1-t_2 . It can be achieved by increasing or decreasing $h_{typ, p, t1-t2, adj}$ values:

$$\sum_{p=1}^{p} \frac{PR1_{p} + PR2_{p}}{2} \cdot h_{typ, p, tl-t2, adj} \cdot \tau \rightarrow E_{typ, tl-t2.}^{k}$$
(8)

When it is known how much time wind power plants generate a certain power range during a typical day of each season, wind power output distribution arrays $X^k(L_k)$ for each typical day k is created. A number of elements (values) in each of these arrays depends on temporal resolution L_k . If, for example, hourly temporal resolution is selected, then each array has 24 elements, one element per hour. Further calculation procedure will be shown only for one typical day, index k will be omitted in notification of elements in array for simplicity. Therefore, wind power output distribution array for one typical day will be noted as X(L).

The number of time steps (hours – for hourly temporal resolution in an economic model) wind power plants generate a certain power range $p \in P$ in X(L) array is represented by the number of elements x(l) that are equal to that power range average value:

$$x(l) = \frac{PR1_p + PR2_p}{2IC}.$$
(9)

Here IC is the installed capacity of wind power plant in MW.

A computer code is used to generate wind curves for these typical days by randomly rearranging X(L) elements. A block diagram for wind curves generation code is shown in Fig. 3 also description of variables used in the code are given in Table 1. This rearrangement is done by looping a number of times (l = 1, L) equal to the number of elements x(l) in X(L) array. In each loop (l = 1, L) one of X(L) values x(y) is entered into array D(Q), (where d(q) = x(y) and Q = L) representing the wind power plant output curve. Which value x(y) is entered into D(Q) as d(q) is determined by variable *y* = *rand_int* – randomly generated integer between 1 and the number of elements in X(L). In order to ensure that all generated random integers are unique, an array

C(O), O = Q is used to track these integers. If randomly generated integer value y already is in C(O) array, then a new random integer is generated and checked if it has a matching value in C(O). If not, random integer value y is entered into C(O) array element c(o), where o = q, then if it is a first entry into D(Q) array x(y) value is entered into d(1) and then the program goes to the beginning of the loop; if it is not a first entry, then the program checks if difference between x(y) and d(q-1) is less or equal to the set tolerance. If so, the computer code enters x(y) into array D(Q) (d(q) = x(y)). The set tolerance ensures that the created wind output curve, represented by array D(Q), would not exceed the desired variability, i.e. output would not change

Variable	Туре	Description			
X(L)	Array, integer or real	Wind power plant's output distribution points, which will be used to create the output curve			
A(M)	Array, integer or real	Different output points			
B(N)	Array, integer	Repetition of each different output in $X(L)$ array. Example: If $a(m) = 3$ and there are four "3" values in $X(L)$ array then $b(n) = 4$, where $n = m$			
C(O)	Array, integer	Array for generated different random numbers			
D(Q)	Array, integer or real	Output curve array			
q	Integer	= 0 Variable used to track loops in FOR EACH cycle			
gen_nr	Integer	= 0 Variable used to track how many times random integer was generated			
gen_max	Integer	This value sets how many times a random number can be generated be- fore exiting FOR EACH loop			
iter	Integer	= 1 Variable used to track iterations of this program			
iter_max	Integer	A maximum number of iterations before the program is terminated			
у	Integer	Random integer between 1 and number of elements in $X(L)$			
0	Integer	Variable used to track first empty position in array C(O)			
set_tol	Integer or real	This value sets what is the biggest difference of nearby values allowed in $D(Q)$			
tol	Integer or real	Biggest difference of nearby values allowed in $D(Q)$ that is in force			
min_B	Integer	This value sets minimum $b(b)$ value that is required for $tol = set_tol$			
add_tol	Integer or real	$tol = set_tol + add_tol$ if $b(n) < min_B$			
т	Integer	Variable used to track loops in FOR EACH cycle			
n	Integer	This variable tracks at which $n a(m = n) = x(y)$,			
n_minus1	Integer	This variable tracks at which $n_{minus1} a(m = n_{minus1}) = D(nr-1)$,			
result	Boolean	If result = 0, it means the program has failed to generate the output curve; if result = 1, the program succeeded			

Table 1. Variables used in code



Fig. 3. Block diagram for wind curves generation code

unrealistically fast. However, there are instances when there is just a single element in array X(L) with a certain value and because of it this element fits only in certain D(Q) positions (see Tolerance example). In order to solve this issue program checks how many elements in X(L)array have x(y) value and how many elements in the same X(L) have the value of the last element entered into array D(Q). If any of these two numbers is lower than set minimum value min_B, then tolerance is increased by additional value *add_tol*. If difference between x(y) and last entry in D(Q) array is greater than tolerance, then the process goes back to generation of new random integer variable y, otherwise x(y) is entered into the first empty D(Q) array's position and the program returns to the beginning of the loop. The program ends when a loop has been made for each element in X(L). However, since positions in D(Q) for each X(L) element are assigned randomly, often after several x(y)values have been entered into D(Q) array it becomes impossible to find next unused value from X(L) array, which meets tolerance constrains, because of bad positioning. To avoid unending searching a variable *gen_nr* is used to track how many times a random integer has been generated and if this number exceeds a set maximum limit (gen_max) then the program goes back to the beginning and starts over. To avoid program running indefinitely, when it is impossible to reposition all X(L) elements in a way which meets the restrictions, a variable *iter* tracks iterations of this program and after a set maximum number of iterations (*iter_max*) the program is terminated.

Tolerance example

This unevenness in the averaged curve can be decreased by increasing tolerance by *add_tol* value when the number of X(L) elements, which are equal to X(y) or last D(Q) entry, is less than *min_B*. See Fig 5. In this example *add_tol* = 0.125 and *min_B* = 2.

The code for this methodology was developed by using Visual Basic for Applications (VBA) programming language within Microsoft Office Excel [14].



Fig. 4. Averaged output curve without additional tolerance



Fig. 5. Averaged output curve with additional tolerance

RESULTS

The methodology was tested by running a simple MESSAGE model of isolated electricity and district heat sector of Lithuania. In this model electricity generation from solar power plants was not evaluated, in order to ensure that time aggregation impact on solar power plants generation would not affect electricity production of wind power plants. When modelling, the hourly time resolution was used. In total, 6 scenarios were run. The first one was a reference scenario, each day in a year was modelled (8760 time steps). It is assumed that it had the greatest accuracy, since no time aggregation was applied. In the second scenario a year was represented by 4 seasons (spring, summer, autumn and winter) and each season by typical work days and weekends (192 time slices). In the third one a year was repre-

sented by 12 months and each month by typical work days and weekends (576 time slices). Other 3 scenarios are like "reference", "4 seasons" and "12 months" scenarios, just with removed pumped storage hydro power plant (PSHPP) technology. These 3 additional scenarios were modelled in order to determine the relationship just between time aggregation and electricity production in wind power plants. Since PSHPP operation is also highly dependent on time aggregation, the effects of PSHPP on wind power plants had to be eliminated. Other technologies in this model are not dependent on time aggregation. Results for all scenarios were calculated for years 2016, 2020, 2030 and 2040. Calculations were made on how much electricity production from wind power plants in aggregated time scenarios differed from production in corresponding reference scenarios. See Table 2.

Year	Scenarios						
	4 seasons	12 months	4 seasons without PSHPP	12 months without PSHPP			
2016	-0.25%	0.01%	-0.25%	0.01%			
2020	13.82%	5.95%	10.54%	5.35%			
2030	16.43%	8.56%	7.76%	1.81%			
2040	6.21%	1.32%	7.05%	4.42%			

Table 2. Difference of electricity generated from wind power plants between aggregated time and corresponding reference scenarios

As expected, the "12 months" scenario had a greater accuracy than "4 seasons" when compared to the "reference" scenario. "4 seasons" scenario results differed from 6.21% to 16.43% and "12 months" scenario results differed from 1.32% to 8.56%. In year 2016, no new wind power plants were built, so results are practically the same as in the reference scenario. From Table 2 we also see that energy storage technologies have a significant impact on the accuracy of the aggregated time model. By eliminating pumped storage hydro power plant technology difference was reduced to 7.05% - 10.54% in the "4 seasons" scenario and to 1.81% - 5.35% in the "12 months" scenario. When selecting time aggregation for the model, an impact of energy storage technologies on model accuracy should be considered.

When using aggregated time in the model, accuracy is reduced depending on the level of aggregation, however optimization time is greatly improved. In Table 3, optimization time, number of equations and variables are given for all scenarios. The model with the "reference" scenario, which had 8760 time slices, was optimized in 4934.7 s. The "12 months" scenario had 15.2 times less time slices (576), but optimization time was 370.2 times (13.33 s.) faster. The "4 seasons" scenario had 3 times less time slices (192) than "12 months", but optimization was done in 2.12 s., 6.29 times less. There is non-linear relationship between time slices and optimization time.

CONCLUSIONS

In this paper a methodology is described on how to represent fluctuations electricity generation from wind power plants in economic models for energy planning, which have aggregated time. The methodology is based on calculations of how much time wind power plants produce certain output within a typical day of a season or some time period and computer algorithm, which generates wind power plants generation curves for these typical days.

The main advantages of this methodology are that generated wind power plants electricity production curves represent not only variability of wind, but also how much time wind power plants generate a certain output range during some time period for which the curve is generated. This output distribution is calculated from wind power plants actual electricity generation or from wind speeds data. Also, it is ensured that energy generated during a typical day corresponds to energy generated during the selected time period. Furthermore, this methodology is applicable for different levels of time aggregation. Since computer algorithm is used to generate these curves, it is not difficult to generate a large number of these curves.

The described methodology was tested with the MESSAGE model of an isolated Lithuanian electricity and district heat sector by comparing high time aggregation level scenarios "4 seasons" and "12 months" with the "reference"

	Scenarios							
Year	Reference	4 seasons	12 months	Reference without PSHPP	4 seasons without PSHPP	12 months without PSHPP		
No. of equations	773147	18015	53727	729227	17055	50847		
No. of variables	773104	17784	53112	737960	17008	50800		
Optimization time, s.	4934.7	2.12	13.33	2597.17	1.26	10.11		

Table 3. Optimization time and number of equations and variables of different scenario matrixes

scenario, which had all 365 days in a year represented. The "4 seasons" scenario results differed by 6.21% – 16.43% and "12 months" scenario results differed by 1.32% – 8.56%. However, the "4 seasons" scenario was optimized in only 2.12 s., "12 months" in 13.33 s., while "reference" in 4934.7 s. If a higher model accuracy is required, then a lower time aggregation level should be used. For example, a year can be represented by 12 months and each month by all typical week-days (typical Monday, typical Tuesday, ...). A model with such time aggregation would have 2016 time slices in a year, more than 4 times less than in the model with no time aggregation.

Time aggregation in the model could be used to reduce the number of equations and variables, which results in significantly faster optimization. However, there is a trade-off between time aggregation and accuracy of the model. Since relationship between time slices in a year and optimization time is non-linear, it is possible to choose aggregation level, at which the model would have sufficient accuracy.

> Received 1 February 2018 Accepted 15 March 2018

References

- Wei Y.-M., Wu G., Fan Y., Liu L.-C. Progress in energy complex system modelling, *International Journal of Global Energy Issues*. 2006. Vol. 25. P. 109–128.
- United Nations Framework Convention on Climate Change. The Paris Agreement. http://unfccc.int/paris_agreement/items/9485.php.
- International Energy Agency. The Energy Technology Systems Analysis Program (ETSAP). TIMES. https://iea-etsap.org/index.php/etsap-tools/model-generators/times.
- 4. International Institute for Applied Systems Analysis. MESSAGE. http://www.iiasa.ac.at/

web/home/research/modelsData/MESSAGE/ MESSAGE.en.html.

- 5. Balmorel. http://balmorel.com.
- Energy-Economy-Environment Modelling Laboratory Research and Policy Analysis. The PRIMES Model. http://147.102.23.135/ index.php?option=com_content&view=category&id=35%3Aprimes&Itemid=80&layout=default&lang=en.
- European Commission. Energy modelling. https://ec.europa.eu/energy/en/data-analysis/ energy-modelling.
- Pina A., Silva C., Ferrão P. Modeling hourly electricity dynamics for policy making in long-term scenarios. *Energy Policy*. 2011. Vol. 39. No. 9. P. 4692–4702.
- Haydt G., Leal V., Pina A., Silva C. A. The relevance of the energy resource dynamics in the mid/long-term energy planning models. *Renewable Energy*. 2011. Vol. 36. No. 11. P. 3068– 3074.
- Pfenninger S. Dealing with multiple decades of hourly wind and PV time series in energy models: A comparison of methods to reduce time resolution and the planning implications of inter-annual variability. *Applied Energy*. 2017. Vol. 197. P. 1–13.
- Lythcke-Jørgensen C. E., Münster M., Ensinas A. V., Haglind F. A method for aggregating external operating conditions in multi-generation system optimization models. *Applied Ener*gy. 2016. Vol. 166. P. 59–75.
- 12. Litgrid. Dashboard. http://www.litgrid.eu/index. php/dashboard/630.
- 13. Patel M. R. *Wind and Solar Power Systems*. CRC Press, 1999.
- Microsoft. Getting Started with VBA in Office 2010. https://msdn.microsoft.com/en-us/ library/office/ee814735(v=office.14).aspx#odc_ office14_ta_AddingVBAProgrammingToYourOfficeToolkit_VBAProgrammingInOffice.

Eimantas Neniškis, Arvydas Galinis

ELEKTROS ENERGIJOS GENERAVIMO IŠ VĖJO ATVAIZDAVIMAS ENERGETIKOS ILGALAIKIO PLANAVIMO EKONOMINIUOSE MODELIUOSE

Santrauka

Ekonominiuose energetikos planavimo modeliuose nėra sudėtinga atvaizduoti elektros energijos generavima vėjo elektrinėse, nes jis nuolatos kinta, todėl galimybės pakankamai tiksliai prognozuoti vėjo greičio pokyčius daugiau nei kelioms dienoms į priekį. Dažnai modeliuojant elektros energijos generavimą yra tiesiog atkartojamas vienų metų elektros gamybos vėjo elektrinėse grafikas. Paprastai šis metodas taikomas analizuojant energetikos sistemos veiklą, kada visos 365 kiekvienų modeliuojamų metų dienos yra atvaizduojamos valandiniais laiko intervalais. Vis dėlto valandinė laiko rezoliucija nėra galima dideliuose modeliuose, kurių optimizavimui paprastai reikalingi dideli kompiuteriniai pajėgumai. Tokiu atveju yra reikalingas koks nors laiko agregavimas. Kita vertus, šiuo metu naudojamos metodikos, kuriose taikoma mažesnė nei valandinė laiko rezoliucija, tampa nebetinkamos dėl didėjančio kintančio elektros generavimo iš vėjo bendroje elektros gamyboje.

Šiame straipsnyje pateikiama metodika, kaip įvertinti vėjo elektrinių elektros energijos gamybos nepastovumą energetikos sektoriaus perspektyvinės raidos modeliuose. Metodika grindžiama vertinimu, kiek valandų per sezoną ar tam tikrame laiko periode per metus viena ar daugiau vėjo elektrinių generuoja tam tikrą galios diapazoną, taip pat generuojamų galių pasiskirstymo pasirinkto laikotarpio tipinėje dienoje modeliavimu ir elektros generavimo kreivių sudarymu. Modeliuojant vėjo elektrines modeliuose su agreguotu laiku pagal pateiktą metodiką sudarytos elektros gamybos kreivės leidžia geriau įvertinti gamybos netolygumus ir pastebimus dėsningumus, taip pat objektyviau nustatyti rezervinių galių poreikį. Negana to, užtikrinama, kad per tipinę dieną pagamintas elektros energijos kiekis atitiktų pagamintą per pasirinktą laiko periodą. Siūloma metodika padės geriau nustatyti racionalų elektrą generuojančių technologijų derinį, tiksliau įvertinti vėjo elektrinių ekonomiškumą ekonominiuose energetikos planavimo modeliuose per daug nedidinant ir taip didelių modelių dydžių.

Raktažodžiai: vėjas, nepastovumas, energetikos sistema, raida, ekonominis modelis