



# Wind Power Variability Model

## Part III – Validation of the model

Joaquín Mur-Amada, Ángel A. Bayod-Rújula  
Zaragoza University, Department of Electrical Engineering  
C/ María de Luna 3, Ed. Torres Quevedo  
50018 – Zaragoza (Spain)

phone: +34 976 761920, e-mail: joako@unizar.es and aabayod@unizar.es

**Abstract**— In this article, the model described in the first part of the series is applied to a system with three wind farms. The model is adjusted from one year data and some measures of the goodness of the fit are shown. In particular, the exponential distribution of the permanence time in states is contrasted with real data. The uncertainty of the transition probabilities and the estimation of uncommon events are also studied. Some theoretical properties obtained from the model are also checked with experimental data.

**Index Terms**— wind power, Markov, variability.

### I. INTRODUCTION

Six wind farms totaling 251,3 MW are connected to a PCC at transmission level (220 kV). However, only data from 2/3 of the wind generation was available for this study. Therefore, short-circuit impedances at PCC have been scaled proportionally to account that the other farms at PCC will probably have an output similar to the measured ones (per unit short circuit impedance at PCC is computed based on installed wind power instead of measured wind power). The equivalent layout is shown in Fig. 1.

The effect of wind power variability is investigated on the voltage and the number of tap changes at the main transformer. The only available data is the active and reactive power output of the wind farm at connection buses. In case reactive power is not available or it is a parameter to be optimized, it can be derived from a wind farm model.

### II. MARKOV MODEL OBTAINED FROM CONVENTIONAL CLUSTERING

Fig. 2 shows the array of all the bivariate scatterplots between active power of the tree farms, along with a univariate histogram for each active power. The classification of each

point is codified with different colors in the scatterplots. Table I show the values of the centroids of each group (i.e., their mass center). Due to the fact that the wind farm power output is highly correlated, only eight states have been used in the clustering algorithm obtaining 0.011 p.u. average classification error. This example shows that if there is a high degree of correlation between variables, clustering can decrease notably the number of states to be considered (compare 8 to  $4^3=64$  states). In [1], an example of power classification 14 wind farms from an area of about 100 km of diameter is shown.

The input of the clustering process can be only the active power or also the reactive power of the wind farms. Generally, the inclusion of the wind farm reactive power  $Q$  does not decrease the performance of the clustering process since  $P$  and  $Q$  are usually highly related. For a fixed power factor regulation, the reactive power  $Q$  can be computed from the active power  $P$  with acceptable precision.

If  $Q$  is controlled according to network parameters or a scheduled planning, a suitable approach is to model such relationships directly in the power flow run. Occasionally,  $Q$  can depend greatly on unmeasured parameters or unknown control policies and it must be statistically characterized. In those

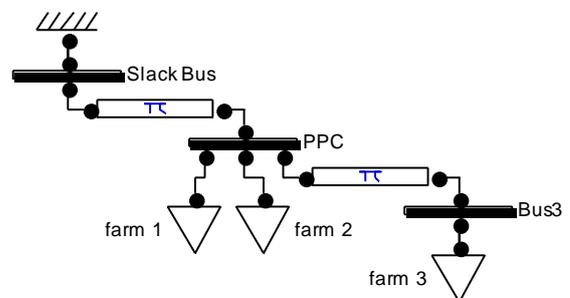


Fig. 1: PSAT model for the tree wind farms, modeled as PQ nodes connected to the PCC

situations both P and Q are the inputs of the clustering process at the cost of increasing the number of groups to maintain the classification error.

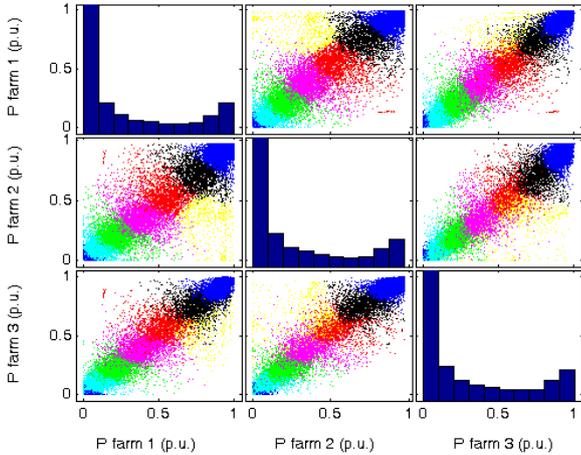


Fig. 2: Scatterplots between active farm power P and their histogram (classes are shown in blue, green, magenta, red, black and dark blue).

Table I is the result of a conventional clustering algorithm whose inputs are the active P and reactive Q powers measured at the billing meter of three wind farms. If only active powers are used for the clustering process, the output is very similar because P and Q are highly related in this case.

TABLE I  
CENTROIDS OF P AND Q VALUES P.U. OF THREE WIND FARMS

#	P <sub>1</sub>	Q <sub>1</sub>	P <sub>2</sub>	Q <sub>2</sub>	P <sub>3</sub>	Q <sub>3</sub>	Freq.
1	0.0112	-0.0036	0.0119	-0.0100	0.0124	0.0009	37.9%
2	0.1092	-0.0061	0.1083	-0.0205	0.1156	0.0040	15.14%
3	0.2385	0.0024	0.2199	-0.0154	0.2402	0.0114	10.12%
4	0.4026	0.0206	0.3621	-0.0045	0.3995	0.0280	8.00%
5	0.5409	0.0435	0.5641	0.0203	0.5850	0.0587	5.97%
6	0.8152	0.1012	0.3269	-0.0035	0.6356	0.0702	1.77%
7	0.7624	0.0922	0.7238	0.0465	0.7711	0.1011	6.74%
8	0.9199	0.1354	0.8941	0.0781	0.9268	0.1437	14.37%

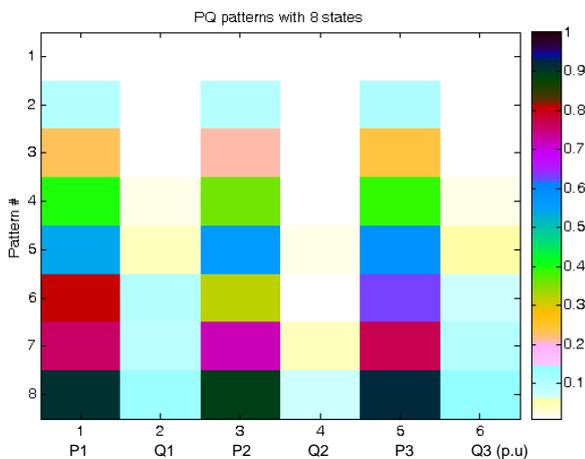


Fig. 3: Color graph of centroids of P and Q powers of table I.

Fig. 3 is a color graph representation of the centroids of the eight patterns obtained from the clustering process.

Fig. 4 is the histogram of the eight patterns obtained from the clustering process from data of a year measured every 15 minutes. The no-load and low load states are the more frequent (37.9% and 15.14%) followed by full load (14.37%). This is typical of low wind resource wind farms. The clustering algorithm has selected the pattern #6, which corresponds to high generation at the first one, low generation at the second wind farm and middle generation at the third one. The selection of this pattern decrease the total classification error, even though it represents only 1,77 % of time operation.

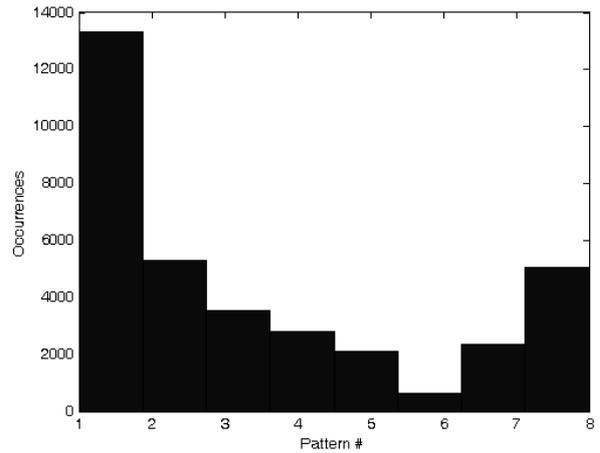


Fig. 4: Histogram of the states in data from a whole year operation.

The study of the transition rates for 15 minutes interval data reveals that the more stable states are the first and the last (full and no generation), whereas a transition to immediately upper or lower states are noticeable. However, transitions to non-adjacent states are very low except in states 6 and 7 (these states correspond to different wind directions, which are very steady in the zone).

The probability transition matrix  $\hat{P}$  is estimated from real data in Table II. Fig. 5 shows that transitions with immediately upper and lower states are relatively frequent (light gray), but jumps to far states are scarcely probable (shades of white). Note that the rows correspond to the initial state and the columns correspond to the state of the next interval. Thus, the probability of going from state 5 to state 7 in one step is stored in row number 5 and column number 7.

State 6 reflects the operation of farm #2 at unusual low power (only 1,77 % of occurrences from Fig. 4). The probability transition matrix shows that jumps between states 5 and 7 states are more probable than jumping to the adjacent state 6 (unusual low power at wind farm 2 and high power at farm 1).

The unobserved transitions from state  $i$  to  $j$  of Table I,  $F_{ij} = 0$ , can be due to real unfeasibility or to the limited available data (for example, infrequent transitions of sporadic states). An improved estimation can be employed if unobserved transitions are feasible even though they have not been observed because they are very rare events and data record, short.

If  $F_{ij} = 0$ , the transition probability  $\hat{p}_{ij}$  is bound to  $[0, 1 - (1 - \beta)^{1/F_i}]$  with confidence level  $\beta$  ( $F_{ij}$  is binomially distributed) [2]. Thus, the null elements of  $\hat{p}_{ij}$  could be substituted by a value in the interval  $[0, 1 - (1 - \beta)^{1/F_i}]$  and the rest of the elements rearranged to make  $\sum_{j=1}^n \hat{p}_{ij} = 1$ . However, the use of a random point in the interval introduces bias in the estimation (usually, pessimistic overestimation of extreme variability).

The unbiased estimation of  $\hat{p}_{ij}$  is zero for unobserved transitions, but its uncertainty is inversely proportional to the number of occurrences of state  $F_i$  because their transition probability  $\hat{p}_{ij}$  is bound to  $[0, 1 - (1 - \beta)^{1/F_i}] \approx [0, -F_i^{-1} \ln(1 - \beta)]$  with confidence level  $\beta$ .

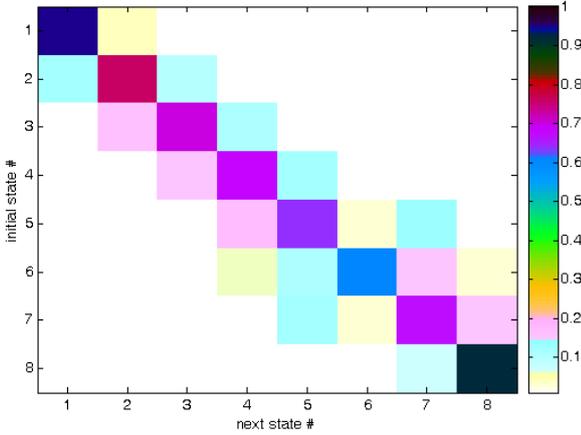


Fig. 5: Color map representation of transition matrix  $\hat{\mathbf{P}}$ .

TABLE II  
INITIAL PROBABILITY TRANSITION MATRIX  $\hat{\mathbf{P}}$

to \ From state	1	2	3	4	5	6	7	8
1	0.9510	0.0472	0.0015	0.0001	0.0002	0	0	0
2	0.1201	0.7646	0.1074	0.0064	0.0006	0.0006	0.0004	0
3	0.0025	0.1693	0.7004	0.1156	0.0068	0.0034	0.0014	0.0006
4	0.0011	0.0057	0.1565	0.6899	0.1216	0.0181	0.0057	0.0014
5	0.0005	0.0005	0.0110	0.1760	0.6419	0.0324	0.1359	0.0019
6	0	0.0016	0.0144	0.0674	0.1124	0.6116	0.1541	0.0385
7	0	0.0017	0.0008	0.0051	0.1284	0.0363	0.6755	0.1521
8	0	0	0	0.0004	0.0010	0.0044	0.0721	0.9221

Characteristic times of the system (eigenvalues of  $\hat{\mathbf{P}}$ ) and limiting distribution of states are continuous functions of matrix elements and the effect of almost zero elements are not important. But a rare transition can have very high cost associated in the optimization algorithm (for example, a sudden loss of all wind generation which can cause a blackout in an island). Therefore, a rare transition might dominate optimization.

There is a tradeoff between the number of classification states and the uncertainty of transition matrix. The use of a bigger number of states decreases classification error but increases uncertainty of infrequent transitions.

If there is a bottleneck or an important topological change when an atypical generation pattern occurs, then a big number of states is advisable because increasing the uncertainty in  $\hat{\mathbf{P}}$  is acceptable. If there are no bottlenecks or violations out of the ordinary and the main purpose of the study is the effects of exceptional events (for example, assessment of contingencies, optimization of the spinning reserve allocation,...), then a reduced set of states can be enough.

The unobserved transitions are in *italics* in Table II and Table III. Their bounds range from  $\hat{p}_{16} \in [0, 2 \cdot 10^{-4}]$  to  $\hat{p}_{61} \in [0, 48 \cdot 10^{-4}]$  for 95% confidence level. For a better estimation of such uncommon events, similar transitions can be joined (for example, estimate together the transition from low power – states 1 or 2– to high power – states 6, 7 or 8–). Therefore, transition probability from states {1, 2} to {6, 7, 8} would be assumed to be the same. The transitions which are similar can be inferred from the cluster dendrogram. Note that the numbers in *italics* are estimates of unobserved transitions based in available knowledge and some other elements have been adjusted for each row to sum 1.

TABLE III  
PROBABILITY TRANSITION MATRIX  $\hat{\mathbf{P}}$   
(ADJUSTED JOINING SIMILAR INFREQUENT TRANSITIONS)

to \ From state	1	2	3	4	5	6	7	8
1	0.9510	0.0472	0.0013	0.0001	0.0001	<i>0.0001</i>	<i>0.0001</i>	<i>0.0001</i>
2	0.1201	0.7643	0.1074	0.0064	0.0006	0.0006	0.0003	<i>0.0001</i>
3	0.0025	0.1693	0.7004	0.1156	0.0068	0.0034	0.0014	0.0006
4	0.0011	0.0057	0.1565	0.6899	0.1216	0.0181	0.0057	0.0014
5	0.0005	0.0005	0.0110	0.1760	0.6418	0.0324	0.1359	0.0019
6	<i>0.0001</i>	0.0015	0.0144	0.0674	0.1124	0.6116	0.1541	0.0385
7	<i>0.0003</i>	0.0015	0.0008	0.0051	0.1284	0.0363	0.6755	0.1521
8	<i>0.0001</i>	<i>0.0001</i>	<i>0.0001</i>	0.0002	0.0010	0.0044	0.0720	0.9221

### III. SYSTEM DYNAMICS AND EQUIVALENT STOCHASTIC DIFFERENTIAL EQUATIONS

The use of Discrete Time Markov Chains, (DTMC) implies that the permanence time in a state is distributed geometrically or as an exponential random variable if the time is considered continuous (Continuous Time Markov Chains, CTMC).

For DTMC, the probability of being in state  $i$  during  $k$  measuring intervals before changing to other state is a geometric random variable with parameter  $p = 1 - p_{i,i}$ , where  $p_{i,i}$  is the  $i$  diagonal element of the probability transition matrix  $\hat{\mathbf{P}}$  of Table III. The number of intervals  $k$  in a time interval  $t$  is  $k = t \cdot f$ , where  $f$  is the frequency of the recorded data.

The probability of permanence in state  $i$  more than  $k$  measuring intervals ( $k = t \cdot f$ ) is the complementary cumulative density function of a geometric random variable:

$$\Pr ( X[j] = i, j = 0, \dots, k \mid X[0] = i ) = p_{i,i}^k = p_{i,i}^{t \cdot f} = e^{-t/\tau_i} \quad (1)$$

$$\tau_i = \text{characteristic or decay time of state } i = \frac{-1}{f \ln(p_{i,i})}$$

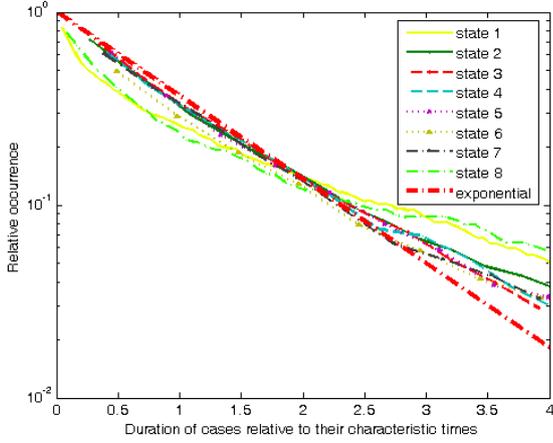


Fig. 6: Probability of permanence more than a given time in each state (complementary cumulative distribution function of permanency time) (x axis scaled to the characteristic time  $\tau_i = 1/\ln(p_{i,i})$ ).

The average permanency time  $\mu_i$  in state  $i$  can be computed using (2)

$$\begin{aligned} \mu_{i,i} &= \text{Expected Value}(\text{time staying in state } i) = \\ &= \frac{1}{f} \text{Expected Value}(k \text{ intervals staying in state } i) = \\ &= \frac{1}{f} E(k | X[k] \neq i; X[j] = i, j = 0, 1, \dots, k-1) = \frac{1}{(1-p_{i,i})f} \quad (2) \end{aligned}$$

Wind power measures are the average value during a time interval. But if instantaneous wind power is considered, wind power and time are continuous variables. Therefore, the mean value computed with continuous time is (2)

$$\begin{aligned} \mu_{i,i} &= \text{Expected Value}(\text{time staying in state } i) = \quad (3) \\ &= E(T | X[T] \neq i; X[t] = i, t < T) = \int_0^{\infty} t \frac{1}{\tau_i} \exp(-t/\tau_i) dt = \tau_i \end{aligned}$$

The characteristic time  $\tau_i$  can be seen as the average time spent in a state  $i$  before leaving it or, alternatively, as the time where the probability of remaining in state  $i$  at time  $t = \tau_i$  is  $1/e = 36,79\%$ .

$$\begin{aligned} \text{Probability}(\text{staying exactly } k \text{ intervals in state } i) &= \\ = \Pr(X[k] \neq i; X[j] = i, j = 0, \dots, k-1 | X[0] = i) &= \quad (4) \\ = p_{i,i}^{k-1} \cdot (1 - p_{i,i}) = p_{i,i}^k \cdot (1/p_{i,i} - 1) \end{aligned}$$

The expression (4) can be rewritten approximately in terms of characteristic time of the state  $\tau_i$

$$\begin{aligned} \text{Probability}(\text{staying exactly } k \text{ intervals in state } i) &= \\ = p_{i,i}^k \cdot (1/p_{i,i} - 1) = \exp\left(-\frac{k}{f \tau_i}\right) \left[ \exp\left(\frac{1}{f \tau_i}\right) - 1 \right] &= \quad (5) \\ = \exp\left\{-\frac{k}{f \tau_i} + \ln\left[\exp\left(\frac{1}{f \tau_i}\right) - 1\right]\right\} \end{aligned}$$

Therefore, the graph of the relative occurrences of staying  $k$  intervals in state  $i$  is a straight line in a semi-logarithmic plot with slope  $-1/\ln(p_{i,i})$  and intercept with vertical axis  $1/p_{i,i} - 1$ . In order to check the goodness of the fit, Fig. 7 shows the probability mass distribution of permanency time in each state with the horizontal axis is scaled by  $-1/\ln(p_{i,i})$  and with the relative occurrence scaled by  $1/p_{i,i} - 1$ . In Fig. 7, the  $\exp(-t)$  function (a straight dashed line) has been included to compare experimental data with theoretical distribution. Since states 1 and 8 have long characteristic times and experimental data is limited to one year, the occurrence of long periods with the same states is scarce and it shows a high variability (for example, the permanence during exactly 80 quarters of hour in state 1 can happen from 0 to 3 times in a year).

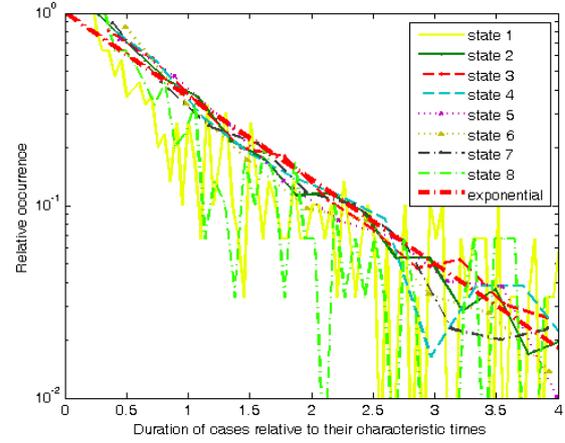


Fig. 7: Probability mass distribution of permanency time in each state (x axis scaled to  $1/\ln(p_{i,i})$  and y axis scaled to  $1/p_{i,i} - 1$ ).

$\hat{\mathbf{P}}$  contains much more information than just the distribution of time permanence in states. The rest of this section will look into the dynamics of the system. Note that states are not sharply defined and time is not discrete because wind farm power output is a continuous varying property.

In the discrete case, a DMTC corresponds to the forward  $m$  order difference equation (6) with the initial probability distribution  $\mathbf{x}(0) = [x_1(0), x_2(0), \dots, x_m(0)]$  as initial condition.  $\mathbf{P}$  is the one-step transition probability matrix [3].

$$\mathbf{x}(k+1) = \mathbf{x}(k)\mathbf{P} \quad (6)$$

Therefore, the probability distribution  $k$  instants later can be computed as:

$$\mathbf{x}(k) = \mathbf{x}(0)\mathbf{P}^k \quad (7)$$

The stochastic matrix  $\mathbf{P}$  has a dominant eigenvalue  $\lambda_1 = 1$  and it is irreducible, recurrent and acyclic in wind farm characterization. Since  $\hat{\mathbf{P}}$  is the matrix  $\mathbf{P}$  estimated from data, the probability of  $\hat{\mathbf{P}}$  having two or more eigenvalues exactly the same tends to zero for increasing sets of data (eigenvalues are continuous functions of matrix elements). Therefore,  $\hat{\mathbf{P}}$  is diagonalizable in practice.

The solution (7) can be easily expressed in the coordinates specified by the left eigenvectors  $\mathbf{Y}$  of  $\hat{\mathbf{P}} = \mathbf{Y}^{-1} \text{diag}(\hat{\lambda}_i)\mathbf{Y}$ , where  $\text{diag}(\hat{\lambda}_i)$  is the diagonal matrix containing the eigenval-

ues  $\hat{\lambda}_i$  of  $\hat{\mathbf{P}}$ . The first eigenvalue is  $\hat{\lambda}_1 = 1$  and the rest are smaller.

If the state probabilities are expressed in the canonical coordinates specified by basis  $\mathbf{Y}$ :  $\mathbf{v}(k) = \mathbf{x}(k) \mathbf{Y}$ , then the dynamics of the system are much simpler:

$$\mathbf{v}(k) = \mathbf{v}(0) \left[ \text{diag}(\hat{\lambda}_i) \right]^k = [v_1(0), v_2(0), \dots, v_m(0)] \text{diag}(\hat{\lambda}_i^k) \quad (8)$$

In other words, the dynamics in these coordinates can be expressed in independent scalar equations corresponding to first order systems of characteristic time  $\hat{\tau}_i$ :

$$v_i(k) = v_i(0) \hat{\lambda}_i^k = v_i(0) \hat{\lambda}_i^{t/f} = v_i(0) \cdot \exp(-t / \hat{\tau}_i) \quad (9)$$

$$\hat{\tau}_i = \frac{-1}{f \ln \hat{\lambda}_i} \quad (10)$$

The first row of  $\mathbf{Y}$  is the limiting probability,  $\pi_1$  or just  $\pi$  (eigenvector corresponding to unity eigenvalue). When the initial probability distribution is  $\pi_1$ , the distribution doesn't change in time ( $\hat{\tau}_i = 1/0 = \infty$ ). The rest of distributions  $\pi_i$  decay with characteristic time  $\hat{\tau}_i$  (at  $t = 3\hat{\tau}_i$ , the probability distribution  $\pi_i$  has faded away to 5% of the initial value).

The forward equations of a CMTC correspond to the Chapman-Kolmogorov Nth order differential equation (11).  $\mathbf{Q}$  is the generator matrix of the CMTC.  $\mathbf{Q}$  can be estimated as  $\hat{\mathbf{Q}} = f \ln(\hat{\mathbf{P}}) \sim f(\hat{\mathbf{P}} - \mathbf{I})$ ,  $\ln$  is the matrix natural logarithm,  $\mathbf{I}$  is the identity matrix and  $f$  is the frequency of recorded data.

$$\frac{d}{dt} \mathbf{x}(t) = \mathbf{x}(t) \mathbf{Q} \quad (11)$$

Its solution is  $\mathbf{x}(t) = \mathbf{x}(0) e^{\mathbf{Q}t} = \mathbf{x}(0) \mathbf{P}^{f \cdot t}$ , where  $e^{\mathbf{Q}t}$  is the exponential of matrix  $\mathbf{Q}$   $t$ . If  $\hat{\mathbf{P}}$  is diagonalizable,  $\hat{\mathbf{Q}}$  is also diagonalizable and it has eigenvalues  $f \ln \hat{\lambda}_i = 1/\hat{\tau}_i$  and the same eigenvectors than  $\hat{\mathbf{P}}$ . Therefore, the equations of continuous time dynamics in the canonical coordinates of  $\hat{\mathbf{Q}}$  (12) are equivalent to (9), the discrete case [4].

$$v_i(t) = v_i(0) \cdot \exp(-t / \hat{\tau}_i) = v_i(0) \cdot \exp(-t / \hat{\tau}_i) \quad (12)$$

Therefore, DMTC and CMTC are equivalent. The computational burden is smaller for DMTC and discrete data is better suited for  $\hat{\mathbf{P}}$  estimation. CMTC gives deeper insight on system dynamics, mimics better its continuous behavior and it can be used to derive easier some properties of the system. Moreover, CMTC approach is more familiar for control engineers.

Note that other numerical approaches (different from eigenvalue calculus) can be computational more efficient in some applications [5].

#### IV. PERMANENCE TIME IN A STATE

The use of Markov Chains implies that the permanence time in a state is distributed geometrically if time is discretized or, equivalently, as an exponential random variable if the time is a continuous variable.

Fig. 8 confirms the assumption that a Markov Chain can approximately model the behavior of the system since the distribution of permanency time in one state approximately corresponds to a geometric random variable (for DMTC) or exponential random variable (for CMTC). This can be checked in

Fig. 8, where the probabilities of remaining in the same state versus time shows an exponential relationship (i.e. distribution is approximately a straight line in a semi-log plot whose slope corresponds to the inverse of the standard characteristic time).

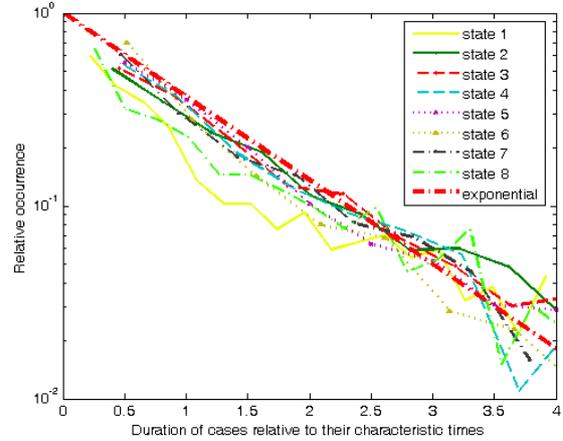


Fig. 8: Probability mass distribution of permanency time in each state (normalized scaling time by state characteristic time).

The main discrepancies are at the distribution tail in rare long-lasting periods at full or no generation (states 1 or 8) due to very stable meteorological situations. For example, period of almost 7 days in state 1 (calm) have occurred in one year data. These outliers caused states 1 (calm) and 8 (full generation) have overestimated characteristic time.

The geometrical distribution is a special case of negative binomial distribution with parameter  $r = 1$ . A negative binomial distribution has been adjusted and parameter  $r$  ranges from 0,5 to 3,5, depending on the state number (the 95% confidence interval did not include  $r = 1$ ). The negative binomial distribution is an alternative to the geometric distribution when the occurrence frequency (or transition probability) varies in time.

The deviation of permanence time from geometrical distribution is due to:

- The state transition rates depend physically on meteorological conditions and on wind farm availability. For example, meteorological stable conditions at calm or high winds can eventually last long periods.
- The “hard” classification of a measurement into a solo state increases observed transitions in the state, especially if the measurement is near two cluster borders. This can explain that permanence times of states 2 to 7 decrease steeper than exponential model ( $r$  ranges from 2 to 3).
- The estimation of system characteristic times with formulas (4) to (1) is an oversimplification. In fact, the Markov Chains are “centrifugal” in the sense that when the system is at partial generation, the system tends to evolve to the first (calm) or last (full power) states. This can explain why the first and last states show a slower decrease of permanence times compared to exponential (geometrical) distribution.
- The permanence time in a state is a concept easy to visualize but it does not correspond to the physical behavior: the

power output of the farms is continuously evolving. Fig. 8 shows the time that power output is bound to cluster area, but system dynamics is more complex.

- The concept of permanence time is not straightforward in fuzzy clustering. The membership level of an observation to a state can increase or decrease from an instant to the next. Therefore, the system can be thought to stay in the state with probability equal to the minimum of the pre and post membership levels. Thus, a consistent stochastic measure of time permanence must be defined (for example, the time interval where the membership is above a threshold level can be considered the permanence time with probability equal to the mean value of the membership to that state).
- An adequate test to check if Markov Chain is a suitable model must employ full dynamics, not only the permanence in a state that is physically evolving.

TABLE IV  
AVERAGE PERMANENCE IN STATES (IN HOURS)

State	Maximum likelihood estimation	Lower 95% confidence interval	Upper 95% confidence interval	Equation (2) $\frac{1}{(1-p_{i,i})f}$	Equation $\hat{t}_{i,i}^{***}$
1	5.1066	4.7362	5.5224	5.1066	5.3839
2	1.0621	1.0056	1.1235	1.0621	1.1834
3	0.8345	0.7866	0.8870	0.8345	0.9544
4	0.8065	0.7555	0.8628	0.8062	0.9276
5	0.6981	0.6507	0.7508	0.6981	0.8122
6	0.6436	0.5696	0.7331	0.6436	0.7357
7	0.7705	0.7188	0.8280	0.7705	0.8838
8	3.2029	2.9084	3.5448	3.2112	3.4453

An alternative to check model accuracy valid for “hard” clustering is to compare the theoretical and observed distribution starting from each state for various time spans (the error measure can be the mean squared difference of theoretical and observed histograms). When fuzzy clustering is used, each time the system is in a combination of states up to a certain degree and there is not possible to compute histograms in the usual way.

A measure of fit goodness that works even with fuzzy clustering is to compare the transition matrix for a time span of  $k$  measuring intervals computed from the one-step transition matrix  $\hat{\mathbf{P}}$  power  $k$  times and estimating a new matrix  $\hat{\mathbf{P}}_k$  based on transitions from initial states to states  $k$  intervals forward. The measure error can be  $\sqrt{|\hat{\mathbf{P}}_k - (\hat{\mathbf{P}})^k|/s}$

However, the system stays occasionally long time in full or no generation due to stable meteorological conditions which are maintained for long time. Moreover, if the wind farm is unavailable for long time, it can distort the distribution of no generation. This is the reason that the actual permanence times are somewhat different from the times computed from transition matrix.

Fig. 9 shows the residence times, where it can be seen that zero and full generation are the most stables states. In fact, the permanency time in such estate is computed as it were a Poisson process (permanency time exponentially statistically distributed). This is characteristic of the Markov chain approach used to characterize wind variability.

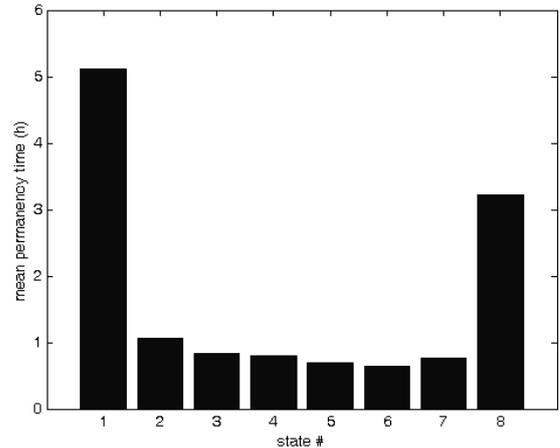


Fig. 9: Average permanency time in each state (in hours) from (2).

## V. CONCLUSIONS

The solution to a Markov Decision Process can be expressed as a policy (PI), which gives the action to take for a given wind farm state, regardless of prior history. Once a Markov decision Process is combined with a policy, this fixes the action of the control for each state and the resulting combination behaves like a Markov chain [6, 7]. The design and sizing of ancillary services can also be optimized.

## VI. ACKNOWLEDGEMENTS

The authors are grateful to Molinos de Aragón (Grupo SAMCA) for providing the data to test the model.

## REFERENCES

- [1] J. Mur, A. A. Bayod "Characterization of wind farm energy production in a zone by artificial neuronal networks", ICREP '03, Vigo 2003. Available at [www.icrepq.com](http://www.icrepq.com)
- [2] B. L. Nelson, "Stochastic Modeling. Analysis and Simulation", 1995, Mc-Graw Hill.
- [3] S. M. Ross, "Introduction to Probability Models", 9<sup>th</sup> Ed., Elsevier 2006.
- [4] Vidyadhar G. Kulkarni "Modeling and analysis of stochastic system", Chapman & Hall / CRC 1995.
- [5] W. J. Stewart, "Introduction to the Numerical Solution of Markov Chains", Princeton University Press, 1994.
- [6] J. Filar, K. Vrieze, "Competitive Markov Decision Processes", 1997, Springer.
- [7] M.L. Putterman, "Markov Decision Processes. Discrete Stochastic Dynamic Programming", John Wiley, New York, 1994.