RECENT DEVELOPMENTS IN THE COMPUTATION OF THE ROCOF OF MULTI-STATE SYSTEMS AND ITS APPLICATIONS

Guglielmo D'Amico[1] and Fulvio Gismondi[2]

•

[1]Department of Economic Studies, University G. d'Annunzio of Chieti-Pescara, Italy
g.damico@unich.it
[2]Department of Economic and Business Science, "Guglielmo Marconi" University, Italy
f.gismondi@unimarconi.it

Abstract

This paper reviews several theoretical works on the computation of the Rate of Occurrence of Failure (ROCOF) for general multi-state random systems, focusing on recent generalizations. The discussion begins by defining the ROCOF for a Markov process and discussing the main results achieved in the literature, then moves towards the richer framework represented by semi-Markov systems. The paper discusses complications that arise when extending the ROCOF to higher orders so that a measure of the association of failures in time can be obtained. The work then analyzes possible modifications in terms of a conditional version of the ROCOF, which is of special interest in applications. The findings are illustrated by a numerical example from reliability, and the broad applicability is demonstrated by a discussion of different applications in other domains.

Keywords: Markov processes, semi-Markov processes, reliability, applications

1. Introduction

Several studies deal with the proposal of new measures of performance for a random system and their computation in applied problems; see e.g. [1]. Among the available indicators, the Rate of Occurrence of Failure (ROCOF) is one of the most frequently used in understanding a system's performance over time. Once a system's failure is defined, the ROCOF is the derivative of the expected number of failures with respect to the time variable. Systems with an increasing path of the ROCOF are expected to deteriorate as time goes on. Contrarily, if the ROCOF shows a decreasing shape, then the system is expected to improve its quality of functioning over time. This seems to be a simple concept, at least in its intuition, but computation and analysis pose relevant questions that have been solved at different moments during the last half century.

In the seventies, some research articles dealing with systems that have already reached the steady state appeared in the literature and showed how to compute the frequency of system failures and their durations [2,3,4].
A few years later, Shi Ding-hua [5] developed a new method for calculating the ROCOF of a system described by a finite-state continuous time-homogeneous Markov chain. The author also discussed the case of a special high-dimensional Markov process equipped with supplementary variables. In the middle of the 1990s, further contributions were given in a couple of research articles by Yeh Lam [6,7]. In the first of his contributions, the author considers a system described by a continuous-time

Markov chain of higher dimension after having introduced additional variables. A formula for evaluating the ROCOF was derived, and an application for a two-component parallel system was presented. In his second contribution, the author enlarged the stage by considering Markovian systems with a denumerable state space.

Markov processes are very frequently adopted for system reliability analysis. Unfortunately, in many circumstances, they are not suitable, either for practical or theoretical reasons. Hence, the need to use more general models, and the semi-Markov ones are a valuable alternative; see e.g. [8,9].

The question of how to compute the ROCOF for a semi-Markov process attracted the attention of Ouhbi, and Limnios [10]. In that work, the authors derived a formula for evaluating the ROCOF for semi-Markov systems and also proposed a statistical estimator of this indicator. A Similar analysis was executed by Georgiadis et al. [11] for the semi-Markov chain, i.e., for semi-Markov models in discrete time.

A new idea was advanced by D'Amico [12] with the concept of ROCOF of order n (shortly denoted by n -ROCOF) for Markov processes. This new indicator coincides with the ROCOF when $n = 1$ and expresses a measure of clustering in the time of failure events. For example, for $n = 2$ it expresses a measure of association in time of a couple of failures at any couple of times (t_1, t_2) with $t_1 < t_2$. After having defined the n-ROCOF, the author derived an explicit formula in terms of the matrix generator and initial probability distribution. Next, a nonparametric estimator was advanced, and its asymptotic properties were determined. The n-ROCOF was applied to the modeling of financial credit ratings, where a conditional version of it was shown to be particularly useful. A few years later, Votsi [13] exploited a conditional version of the ROCOF for semi-Markov chains.

Finally, in [14], the analysis of the n-ROCOF for semi-Markov processes with finite state space is executed in such a way that the previous quoted articles were generalized. The authors determined an explicit formula for the n-ROCOF under a general random starting mechanism, considering any possible state and duration of permanence in it. This was done using a mixed continuous-discrete initial probability distribution function. A set of hypotheses on the model parameters was advanced so that the derivation of an explicit formula expressing the n-ROCOF was obtained. The results were sufficiently general to be of interest not only in the reliability theory field but for every general system and could be applied every time that a partition of the state space can be introduced into working and not-working states.

After giving a thorough analysis of the theoretical findings pertaining to the computation of the ROCOF, we show a numerical example from the dependability area and go on to describe some unusual applications in various fields, spanning from financial mathematics to wind energy generation. Considering that several of these applications have never been mentioned before, they also offer a suggestion for detailed future research.

The paper proceeds as follows: Section 2 describes the basic reliability problem and some of the most important reliability indicators, ROCOF and n-ROCOF included. Section 3 considers a Markov process for the probabilistic description of the system and shows formulas for the ROCOF and n-ROCOF. Section 4 provides a summary of the results related to the extension to the semi-Markov framework, showing the latest results in the literature. Section 5 discusses a numerical example for a Markov system and demonstrates the practical usefulness of the considered measures. Moreover, different possible applications from real life problems are detailed. The discussion concludes in Section 6, which reviews the content of the paper and provides general conclusions.

2. Basic description of the reliability problem and main indicators

The basic reliability problem can be described assuming that the performance of the system can be identified with one element of a finite set $E = \{1,2, \dots, m\}$ called state space. Frequently, an ordering relation on the set E is considered in such a way that higher ranks $j \in E$ correspond to a higher system's performance. The state space E is partitioned into two disjoint subsets W and F such that:

$$E = W \cup F, \qquad W \cap F = \emptyset,\ W \neq \emptyset,\ F \neq \emptyset.$$

The subset W contains all the elements of E denoting acceptable working levels of the systems; instead the subset F contains all the states of E in which the system is not performing in a satisfactory way or has a fault. Sometimes the state space is divided into three subsets, denoting the working states, the changeable states, and the failure states. The changeable states denote a working system if and only if, before entering the changeable subset, the system was working and will continue to work after leaving it; see [15].

The system evolves in time and changes its state migrating from one state i to another state j. The most natural way to study this evolution is by using a stochastic process $Z = \{Z(t),\ t \geq 0\}$. Hence, $Z(t)$ denotes the state occupied by the system at time t and if $Z(t) \in W$ the system is working while if $Z(t) \in F$ the system is not working.
Specific indicators are used to measure the overall quality of the system; among them, we remember:

- the *availability function*, which is defined by

$$A_i(t) \coloneqq P[Z(t) \in W | Z(0) = i].$$

It expresses the probability that the system ranked i at time *0* will be operational at time t independently of the possible behavior before this time.

- The *reliability function* which is defined by

$$R_i(t) \coloneqq P[Z(n) \in W,\ \forall n \in [0, t] | Z(0) = i].$$

This indicator consists of the probability that a system ranked i at time *0* will not experience a fault (a visit to the subset F) from time *0* up to time *t*. A generalization of the reliability function considers interval reliability indicators [16], recent results are available in [17,18], and the sequential reliability function [19].

Denote by $N_f(t)$ the number of failures of the system until time t, i.e. the number of passages from a state of W to one of F. Then,

- the *ROCOF* at time t for a random system, denoted by $ro(t)$, is defined by

$$ro(t) = \lim_{\Delta t \to 0} \frac{E\left[N_f(t + \Delta t) - N_f(t)\right]}{\Delta t}. \tag{1}$$

The ROCOF gives information on whether there are a lot of failures or only a few within a time, and it has a simple probabilistic interpretation that for deteriorating systems shows an increasing behavior and for improving systems, it is decreasing in time.

When studying the reliability of a reparable system, it is of great interest to also measure the relative positioning of tuples of failures. For this reason, the n-ROCOF was defined D'Amico [12]:

- the *n-ROCOF* at times $\boldsymbol{t}_1^n = (t_1, t_2, \dots, t_n)$ with $t_i < t_{i+1}$ for a random system is defined by

$$ro(\boldsymbol{t}_1^n) = \lim_{(\Delta t_i \to 0)_{i=1}^n} \frac{E\left[dN_f(t_1) \cdot dN_f(t_2) \cdot \dots \cdot dN_f(t_n)\right]}{\Delta t_1 \Delta t_2 \cdot \dots \cdot \Delta t_n}. \tag{2}$$

Clearly, for $n = 1$ the n-ROCOF coincides with the ROCOF of the system. For $n = 2$ we obtain an interesting particular case called the 2-ROCOF which is given by

$$ro(\boldsymbol{t}_1^2) = \lim_{(\Delta t_i \to 0)_{i=1}^2} \frac{E\left[dN_f(t_1)dN_f(t_2)\right]}{\Delta t_1 \Delta t_2}, \tag{3}$$

and expresses a measure of association of failure events in correspondence of a 2-dimensional vector of times $\boldsymbol{t}_1^2 = (t_1, t_2)$.

3. n-ROCOF for Markov processes

A class of models frequently used in the reliability field is that represented by Markov processes; see e.g. [20,21,22,23]. Here we briefly introduce them and show the formula for the n-ROCOF.

Let consider a continuous time Markov process $(Z(t),\ \ t \in \mathbb{R})$ with a finite state space $E = \{1,2, \dots, m\}$ and generator matrix $\boldsymbol{Q} = (q_{ij}), i,j \in E$ where $q_{ij} \geq 0$, $\forall i \neq j$ and $q_{ii} = -\sum_{j \neq i} q_{ij}$. Consider also an initial probability distribution over the states of the process at time zero denoted by the vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)$ where

$$\alpha_i = P(Z(0) = i).$$

Let $p_i(t) = P(Z(t) = i), \forall i \in E$ be the state probability at time t, then it results that

$$p_j(t) = \sum_{i \in E} \alpha_i \, p_{ij}(t),$$

where $p_{ij}(t) = P(Z(t) = j | Z(0) = i) = (e^{\boldsymbol{Q}t})_{ij}$.

Theorem [12] The n-ROCOF at times $\boldsymbol{t}_1^n = (t_1, t_2, \dots, t_n)$ with $t_i < t_{i+1}$ for a Markov jump process $(Z(t),\ \ t \in \mathbb{R})$ over a finite state space $E = \{1,2, \dots, m\}$ and generator matrix $\boldsymbol{Q} = (q_{ij}), i,j \in E$ is given by

$$ro(\boldsymbol{t}_1^n) = \sum_{w,f}^{n} \prod_{i=1}^{n} \alpha_{f_0} \cdot \left(e^{\boldsymbol{Q}(t_i - t_{i-1})}\right)_{f_{i-1} w_i} \cdot q_{w_i f_i} \,, \tag{4}$$

where $t_0 = 0$ and the symbol $\sum_{w,f}^{n}$ is an abbreviate notation for $\sum_{f_0 \in E} \sum_{w_i \in W, \forall i=1,\dots,n} \sum_{f_i \in F, \forall i=1,\dots,n}$.

Formula (4) contains interesting particular cases of which we give representation. The ROCOF is simply obtained by setting $n = 1$ in formula (4) with $t_0 = 0$. The result is:

$$ro(t_1) = \sum_{f_0 \in E} \sum_{w_1 \in W} \sum_{f_1 \in F} \alpha_{f_0} \cdot (e^{\boldsymbol{Q}\, t_1})_{f_0 w_1} \cdot q_{w_1 f_1}$$

$$= \sum_{w_1 \in W} \sum_{f_1 \in F} p_{w_1}(t_1) \cdot q_{w_1 f_1}, \tag{5}$$

which expresses exactly the formula established by Yeh [7].

Another interesting case is represented by the 2-ROCOF which is simply obtained by setting $n = 2$ in formula (4) with $t_0 = 0$. The result is:

$$ro(\boldsymbol{t}_1^2) = \sum_{w,f} \prod_{i=1}^{2} \alpha_{f_0} \cdot \left(e^{\boldsymbol{Q}(t_i - t_{i-1})}\right)_{f_{i-1}w_i} \cdot q_{w_i f_i}$$

$$= \sum_{f_0 \in E} \sum_{w_1 \in W} \sum_{f_1 \in F} \sum_{w_2 \in W} \sum_{f_2 \in F} \alpha_{f_0} \cdot (e^{\boldsymbol{Q}\, t_1})_{f_0 w_1} \cdot q_{w_i f_i} \cdot \left(e^{\boldsymbol{Q}(t_2 - t_1)}\right)_{f_1 w_2} \cdot q_{w_2 f_2} \cdot \quad (6)$$

4. n-ROCOF for semi-Markov processes

Semi-Markov processes are a generalization of Markov processes, allowing for any kind of probability distribution function for the sojourn time in the state of the system. Contrarily, Markov processes require exponentially distributed sojourn times; this assumption can be inadequate in several application fields, reliability theory included; see e.g. [24,25].

The definition of a semi-Markov process needs some preliminary concepts to be introduced. We start by considering a bivariate random sequence $(J_n, T_n), n \in \mathbb{N}$. The random variable J_n denotes the state of the system at its n-th transition; this variable assumes values in the state space E. The random variable T_n denotes the time in which the system enters state J_n; this variable assumes any values in the set of positive real numbers. The time the system remains in the state J_{n-1} before entering state J_n is called sojourn time. It is denoted by $X_n = T_n - T_{n-1}$ having set $X_0 = 0$. The process (J_n, T_n) is called Markov Renewal Process (MRP) whenever it satisfies the next assumption:

$$P(J_{n+1} = j, X_n \le t | J_n, X_{n-1}, J_{n-1}, X_{n-2}, \dots) = P(J_{n+1} = j, X_n \le t | J_n).$$

The conditional joint probabilities of the MRP are denoted by

$$Q_{i,j}(t) = P(J_{n+1} = j, X_n \le t | J_n = i)$$

and the matrix $\boldsymbol{Q}(t) = \left(Q_{i,j}(t)\right)$ is called the semi-Markov kernel.

Let $N(t) = sup\{n: T_n \le t\}$ be the counting process of the number of transition up to the time t. Then the semi-Markov process can be defined by $Z(t) \coloneqq J_{N(t)}$.

Let assume that $\boldsymbol{Q}$ is absolutely continuous with respect to the Lebesgue measure on the set of positive real numbers and denote by $q_{i,j}(t) = \frac{Q_{i,j}(dt)}{dt}$ the corresponding Radon-Nikodym derivatives. Hence, we can consider the hazard rate functions according to the relation

$$\lambda_{ij}(t) = \begin{cases} \dfrac{q_{i,j}(t)}{1 - H_i(t)} & \text{if } p_{i,j} > 0 \text{ and } H_i(t) < 1 \\ 0 & \text{otherwise} \end{cases}$$

where $H_i(t) := P(X_n \le t | J_n = i) = \sum_{j \in E} Q_{i,j}(t)$ and $p_{i,j} := P(J_{n+1} = j | J_n = i) = \lim_{t \to +\infty} Q_{i,j}(t)$.

Let us introduce the backward recurrence time process $B(t) \coloneqq t - T_{N(t)}$. Now we can describe the set of three assumptions used in [14] to derive the formula for the n-ROCOF of a semi-Markov process.

Assumption A1: This first assumption explains a general random starting mechanism for semi-Markov processes. We first consider the vector $\boldsymbol{p} = (p_1, p_2, \dots, p_n)$ where

$$p_i = P(J_0 = i)$$

with $\sum_{i \in E} p_i = 1$. Moreover we specify a set of cumulative distribution function for the duration in the initial state i; i.e. $F_i(v_0) = P(B(0) \le v_0 | J_0 = i)$. Then we assume that

$$F_i(v_0) = \begin{cases} 0 & \text{if } v_0 < 0 \\ a_i & \text{if } v_0 = 0,\ 0 \le a_i \le 1 \\ G_i(v_0) \cdot (1 - a_i) + a_i & \text{if } v_0 > 0 \end{cases}$$

being $G_i(\cdot)$ an absolutely continuous cumulative distribution function with support in $(0,\infty)$ and corresponding density function $g_i(\cdot)$ having finite expectation.

Assumption A2: The semi-Markov process has uniformly bounded transition intensities, in formula:

$$\exists c \in \mathbb{R}_+ \text{ such that } \max_{i,j\in E} \sup_{y\in\mathbb{R}_+} \lambda_{ij}(y) \leq c \ .$$

Assumption A3: For each state $i \in E$ there exist a state $j \in E$ (depending on i) and a non-null subset $S_{ij}(y)$ of the real numbers such that $\lambda_{ij}(y) < c$ for all $y \in S_{ij}$.

The following main result gives the formula for the n-ROCOF of a semi-Markov process.

Theorem [14] The n-ROCOF at times $\boldsymbol{t}_1^n = (t_1, t_2, \dots, t_n)$ with $t_i < t_{i+1}$ of a semi-Markov process $(Z(t),\ t \geq 0)$ over a finite state space $E = \{1,2,\dots,m\}$ and semi-Markov kernel $\boldsymbol{Q} = \big(Q_{ij}(t)\big), i,j \in E$ is given by

$$ro(\boldsymbol{t}_1^n) = \sum_{j_0=1}^{m} \sum_{\boldsymbol{w}\in \boldsymbol{W}^n, \boldsymbol{f}\in \boldsymbol{F}^n} a_{j_0} p_{j_0} \prod_{r=1}^{n} \int_0^{s_r} \psi'_{f_{r-1}w_r}(0;u_r) q_{w_r f_r}(s_r - u_r) du_r$$

$$+ \sum_{j_0=1}^{m} \sum_{\boldsymbol{w}\in \boldsymbol{W}^n, \boldsymbol{f}\in \boldsymbol{F}^n} (1 - a_{j_0}) p_{j_0} \left(\int_0^{\infty} \left(\int_{-v_0}^{s_1} \psi'_{j_0 w_1}(v_0;u_1) q_{w_1 f_1}(s_1 - u_1) du_1 \right) g_{j_0}(v_0) dv_0 \right)$$
$$\cdot \prod_{r=2}^{n} \int_0^{s_r} \psi'_{f_{r-1}w_r}(0;u_r) q_{w_r f_r}(s_r - u_r) du_r \ , \qquad (7)$$

where $s_r = t_r - t_{r-1}$, $s_1 = t_1$ and $\psi_{ij}(y;t) \coloneqq E_{(i,y)}\big[N_j(t)\big] = \sum_{n=0}^{\infty} Q_{ij}^{(n)}(y;t)$.

The quantity $Q_{ij}^{(n)}(y;t) = P(J_n = j, T_n \leq t | J_0 = i, B(0) = y)$.

We observe that if we consider for all $j_0 \in E$, $a_{j_0} = 1$ and $n = 1$, we have a null duration in the initial state and we obtain exactly the formula for the ROCOF as it was established by [10], i.e.

$$ro\big(t_1\big) = \sum_{j_0=1}^{m} \sum_{w\in W} \sum_{f\in F} p_{j_0} \int_0^{t_1} \psi'_{j_0 w}(0;u_1) q_{wf}(t_1 - u_1) du_1 \, . \qquad (8)$$

5. Applied problems

Applications of the n-ROCOF measures to the reliability field are clearly of great interest. Here we show a few results related to a numerical example based on Markov processes in subsection 5.1. Anyway, there are many other fields of application in which the same concepts are worth discussing. An example is the financial modeling of credit rating dynamics, which was extensively discussed in [12,26]. A new application to wind power production is discussed later in subsection 5.2 taking inspiration from the problem discussed in [27] and [28].

5.1 A numerical example

Let us consider a random system whose state space is given by the set $E = \{1,2,3,4,5,6,7,8\}$. Each number represents different levels of performance of the system, going from perfect functioning (state 1) to the worst one (state 8). The state space E is partitioned into two disjoint subsets $W = \{1,2,3\}$ and $F = \{4,5,6,7,8\}$.

We assume that the system evolves dynamically according to the generator matrix

$$\boldsymbol{Q} = \begin{pmatrix} -0.10 & 0.08 & 0.00 & 0.02 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.80 & -1.30 & 0.35 & 0.15 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.12 & -1.33 & 0.90 & 0.31 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.10 & 0.30 & -0.90 & 0.50 & 0.10 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.30 & 0.15 & -0.95 & 0.35 & 0.15 & 0.00 \\ 0.00 & 0.00 & 0.05 & 0.20 & 0.10 & -0.50 & 0.10 & 0.05 \\ 0.00 & 0.05 & 0.08 & 0.09 & 0.15 & 0.33 & -0.90 & 0.20 \\ 0.00 & 0.00 & 0.10 & 0.10 & 0.20 & 0.30 & 0.30 & -1.00 \end{pmatrix}.$$

Using equation (5), we compute the ROCOF of order 1 for the three working states. The results are shown in Figure 1. The continuous line refers to the ROCOF computed starting from state 1, i.e., using the initial probability distribution $\boldsymbol{\alpha} = (1,0,\dots,0)$; the dashed line refers to the ROCOF computed starting from state 2, i.e., using the initial probability distribution $\boldsymbol{\alpha} = (0,1,0,\dots,0)$; the dotted line refers to the ROCOF computed starting from state 3, i.e., using the initial probability distribution $\boldsymbol{\alpha} = (0,0,1,0\dots,0)$. The figure shows that independently of the initial state, the system is going to deteriorate as the ROCOF shows increasing paths. Nonetheless, the differences according to the initial state are remarkable and demonstrate a higher risk for state 3 and a lower risk for state 1.

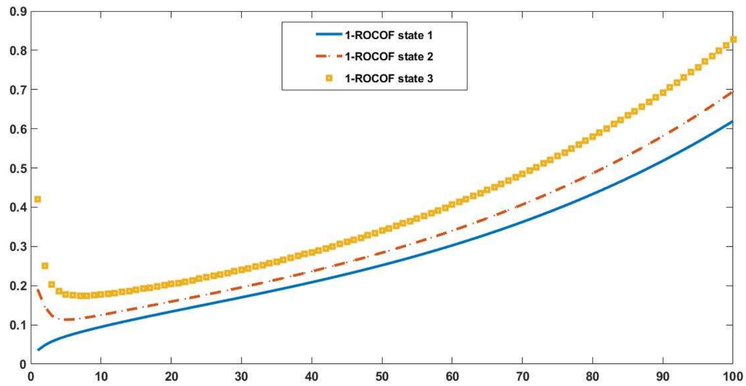

Figure 1. 1-*ROCOF for a Markov process*

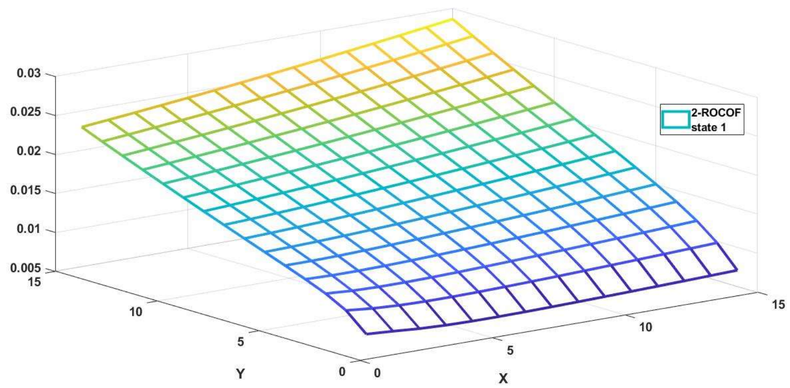

Figure 2. 2-*ROCOF for a Markov process starting from state 1*

Using equation (6), we compute the 2 −ROCOF starting from state 1. The result is graphically displayed in figure 2. On the X-axis, we report the time t_1, while on the Y-axis, we report the time $t_2 - t_1$. Hence, the point $(10, 5)$ on the XY-plane corresponds to the choice of $t_1 = 10$ and $t_2 = 15$. Any point on the surface represents the corresponding value of the $2-$ROCOF. High values of the surface show evidence for the association of failures at the corresponding times on the X and Y axes. The maximum values (for the times considered in the figure) are concentrated on large values of t_1 and $t_2 - t_1$.

We repeat the computations of the $2-$ROCOF changing the initial state. The case with state 2 as initial state and that for state 3 are considered in figures 3 and 4, respectively.

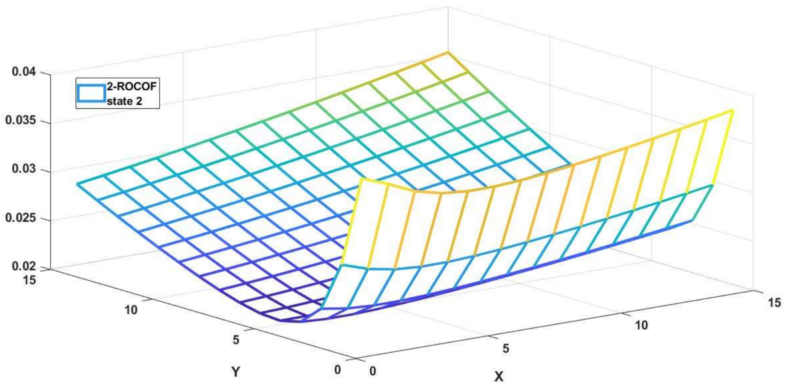

Figure 3. 2-*ROCOF for a Markov process starting from state 2*

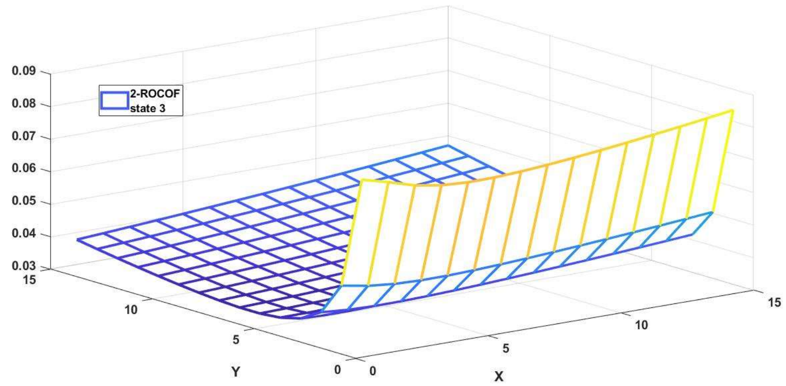

Figure 4. 2-*ROCOF for a Markov process starting from state 3*

As it is possible to see from these figures, their shapes are completely different from those of figure 2. In figures 3 and 4, the maximum values of association between couples of failures are for combinations of short times $t_2 - t_1$ independently from time t_1 (which shows a contained variability). This aspect is important because at short values of $t_2 - t_1$, the system may show trajectories of alternation between subsets $W \to F \to W \to F$, i.e., the presence of two close-in-time failures.

5.2 Wind power example

Wind power is one of the most important renewable energy sources. Because wind speed changes very sharply over time, the wind engineer must use mathematical models to predict the power output. Particular care should be given to abrupt interruptions of power production that may be caused by extreme wind speeds. Indeed, on the one hand, low wind is unable to move wind turbine's blades, which determines no energy production. On the other hand, extremely high wind speeds may cause damage to the turbine; hence, the wind engineer must switch it off to avoid structural breaking. The minimal wind speed necessary to activate the turbine is called the cut-in speed v_{ci}. The maximal wind speed that the turbine can handle is called the cut-out speed v_{co}. Finally, it is also important to consider the rated wind speed v_r, which represents the minimum wind speed value at which the turbine achieves its maximum power production, the so-called rated power. Using the power curve for the wind turbine under consideration, we may calculate the power output as a function of the wind speed. The relationship between wind speed $v(t)$ and wind power $Pow(t)$ at any time t is

$$Pow(t) = \begin{cases} 0 & \text{if} \quad v(t) \leq v_{ci} \\ \dfrac{P_r \cdot (v^3(t) - v_{ci}^3)}{v_r^3 - v_{ci}^3} & \text{if} \quad v_{ci} < v(t) < v_r \\ P_r & \text{if} \quad v_r < v(t) < v_{co} \\ 0 & \text{if} \quad v(t) > v_{co} \end{cases}$$

where P_r is the rated power. Essentially, for wind speed lower than the wind cut-in speed, there is no power production. For wind speed between the cut-in speed and the rated speed the output power is a cubic function of the wind speed. For values between the rated speed and the cut-off speed the turbine produces its rated power. Finally, for speed greater than the wind speed cut-off the turbine does not produce power.

The ROCOF and its generalization could be fruitfully used in this applied field as we are going to show. First, we consider the first-order discrete-time Markov chain model proposed in [28] which was applied to a sample of hourly wind speed data collected by Malaysian Meteorological Station located at Mersing. The wind speed data range from 0 to $12 m/s$; hence the authors adopted a twelve state Markov chain model with $E = \{1,2, \dots ,12\}$ where the i-th state collects all wind speed measurements ranging between $(i-1)\ m/s$ and $i\ m/s$. The estimated transition probability matrix is taken from [28]:

$$\boldsymbol{P} = \begin{bmatrix} 0.371 & 0.407 & 0.174 & 0.036 & 0.009 & 0.002 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.166 & 0.446 & 0.312 & 0.059 & 0.012 & 0.004 & 0.000 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.051 & 0.243 & 0.504 & 0.163 & 0.028 & 0.008 & 0.002 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 \\ 0.017 & 0.083 & 0.303 & 0.391 & 0.160 & 0.035 & 0.008 & 0.002 & 0.001 & 0.000 & 0.000 & 0.000 \\ 0.010 & 0.035 & 0.099 & 0.277 & 0.382 & 0.157 & 0.031 & 0.007 & 0.001 & 0.001 & 0.000 & 0.000 \\ 0.006 & 0.021 & 0.043 & 0.108 & 0.295 & 0.343 & 0.146 & 0.031 & 0.004 & 0.003 & 0.000 & 0.000 \\ 0.005 & 0.016 & 0.027 & 0.047 & 0.110 & 0.302 & 0.324 & 0.142 & 0.021 & 0.004 & 0.002 & 0.000 \\ 0.006 & 0.016 & 0.030 & 0.033 & 0.055 & 0.127 & 0.365 & 0.239 & 0.105 & 0.022 & 0.002 & 0.000 \\ 0.009 & 0.019 & 0.014 & 0.018 & 0.042 & 0.065 & 0.140 & 0.326 & 0.269 & 0.079 & 0.014 & 0.005 \\ 0.014 & 0.054 & 0.055 & 0.014 & 0.027 & 0.028 & 0.041 & 0.205 & 0.288 & 0.164 & 0.083 & 0.027 \\ 0000 & 0.000 & 0.000 & 0.040 & 0.000 & 0.000 & 0.080 & 0.120 & 0.160 & 0.240 & 0.280 & 0.080 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.000 & 0.200 & 0.000 & 0.200 & 0.600 & 0.000 \end{bmatrix}$$

Now consider a commercial wind turbine with a cut-in speed of $4\ m/s$. This means that when the wind speed process is in one of the first four states of the Markov chain, there is no power production. The state space E is partitioned into two disjoint subsets W and F according to the following:

$$W = \{5,6, \dots ,12\} \quad F = \{1,2,3,4\}.$$

In order to apply the measures discussed in the previous sections, we need to transform the discrete dynamic expressed by the transition matrix estimated by [28] into a continuous-time one by finding a generator matrix that satisfactorily matches, in some sense, the discrete process. This is a well-known and still open problem in the theory of Markov chains that is called the embedding problem. A detailed discussion is provided in [29], and further results and applications are provided in [30]. We consider here a simple strategy to get a generator matrix rendering results "close" to the observed hourly transition probability matrix $\boldsymbol{P}$**.** We observe that the transition probability function for a continuous-time Markov process satisfy the relation

$$\boldsymbol{P}(t) = e^{\boldsymbol{Q}t} = \sum_{n=0}^{\infty} \frac{(\boldsymbol{Q}t)^n}{n!}.$$

Therefore, given a probability matrix $\boldsymbol{P}$, we can try to find a generator matrix $\boldsymbol{Q}$ such that

$$\boldsymbol{P} \approx \frac{(\boldsymbol{Q}t)^0}{0!} + \frac{(\boldsymbol{Q}t)^1}{1!} = \boldsymbol{I} + \boldsymbol{Q}t.$$

From this relation we recover $\boldsymbol{Q} = \frac{1}{t} \cdot (\boldsymbol{P} - \boldsymbol{I})$. and we obtain the initial guess $\boldsymbol{Q}^* = (\boldsymbol{P} - \boldsymbol{I})$ by setting $t = 1$ to denote one hour. Hence, we solve the following optimization problem:

$$\min_{\boldsymbol{Q} \in \boldsymbol{\Phi}} \| \boldsymbol{P} - e^{\boldsymbol{Q}} \|,$$

which consists of finding, within the set of generator matrices $\boldsymbol{\Phi}$, the one that minimizes the previous matrix norm. In our application, we consider the minimization of the Frobenius matrix norm, and we use the software Matlab to solve this optimization problem with an initial guess $\boldsymbol{Q}^*$. The result is the following generator matrix:

$$\boldsymbol{Q} = \begin{bmatrix}
-3.295 & 0.035 & 0.006 & 0.002 & 0.000 & 0.001 & 0.001 & 0.000 & 0.309 & 1.999 & 0.942 & 0.000 \\
0.020 & -2.891 & 0.000 & 0.004 & 0.000 & 0.000 & 0.000 & 0.000 & 0.277 & 1.633 & 0.939 & 0.018 \\
0.033 & 0.028 & -2.362 & 0.001 & 0.000 & 0.000 & 0.000 & 0.000 & 0.085 & 0.776 & 1.239 & 0.200 \\
0.268 & 0.059 & 0.014 & -1.975 & 0.014 & 0.000 & 0.000 & 0.000 & 0.028 & 0.137 & 0.504 & 0.951 \\
0.936 & 0.261 & 0.051 & 0.011 & -1.965 & 0.001 & 0.000 & 0.000 & 0.017 & 0.060 & 0.166 & 0.462 \\
0.492 & 0.572 & 0.243 & 0.052 & 0.007 & -1.663 & 0.000 & 0.000 & 0.010 & 0.035 & 0.072 & 0.180 \\
0.183 & 0.503 & 0.540 & 0.237 & 0.035 & 0.007 & -1.746 & 0.080 & 0.009 & 0.027 & 0.046 & 0.079 \\
0.091 & 0.212 & 0.608 & 0.398 & 0.175 & 0.037 & 0.004 & -1.758 & 0.100 & 0.027 & 0.051 & 0.055 \\
0.070 & 0.108 & 0.535 & 1.091 & 0.728 & 0.002 & 0.003 & 0.003 & -2.542 & 0.001 & 0.000 & 0.001 \\
0.045 & 0.047 & 0.068 & 0.341 & 0.480 & 0.273 & 0.138 & 0.045 & 0.024 & -1.877 & 0.392 & 0.024 \\
0000 & 0.000 & 0.133 & 0.200 & 0.266 & 0.455 & 0.522 & 0.133 & 0.000 & 0.001 & -1.777 & 0.067 \\
0.001 & 0.000 & 0.000 & 0.532 & 0.000 & 0.740 & 1.550 & 0.000 & 0.001 & 0.006 & 0.002 & -2.832
\end{bmatrix}$$

Now, we can compute the ROCOF using the formula presented in the previous section. To highlight the potentiality of the continuous-time framework, we compute the indicators on a 5-minute time scale. This can be done simply by considering the hourly-based generator matrix $\boldsymbol{Q}$ and dividing it by a factor of 1/12.

In figure 5, we report the 1-ROCOF corresponding to three choices of the initial distribution over the states of the system. Specifically, the continuous blue line represents the indicator computed starting from state 5, which denotes a wind speed of $5m/s$. The dotted red line denotes the 1-ROCOF behavior starting with a wind speed of $9m/s$. Finally, the dashed yellow line stands for the 1-ROCOF with an initial wind speed of $1m/s$. It is possible to note that for short times, the 1-ROCOF is monotone with respect to the initial speed. Thus, being in a state with strong wind implies a higher chance of moving into one of the failure states where no power production occurs. The overall behavior becomes irrelevant to the initial state around time 20, which corresponds to $20 \cdot 5\text{min} = 100\text{min}$ where the system shows the achievement of a stationary value of the $1 - ROCOF$ equal to 0.0536.

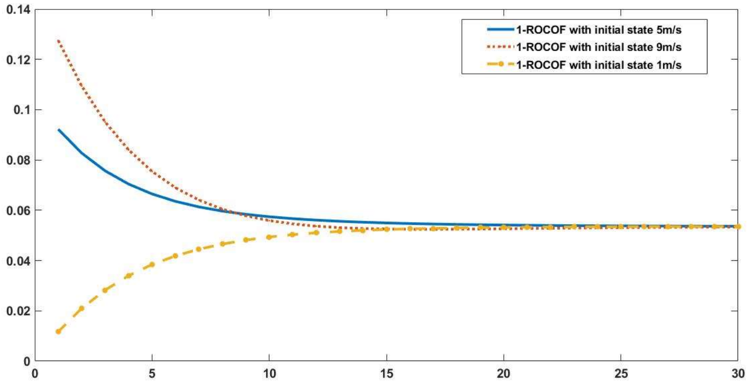

Figure 5. 1-*ROCOF for the Markov process of wind speed*

For completeness, we proceed by computing the 2-ROCOF for the wind speed Markov process. In figure 6, we display the indicator when the initial state is equal to $5m/s$ (left panel) and when the initial state is equal to $9m/s$ (right panel). The panels have similar surfaces. Both indicate a maximum value of the 2-ROCOF for high values of the time t_1 and low values of the time $t_2 - t_1$. Hence, the maximum chance for a couple of transitions from working to failure states is for combinations of times as ($t_1 = 20, t_2 = 21$). The indicator is increasing with respect to time t_1 and decreasing with respect to time $t_2 - t_1$. The 2-ROCOF assumes higher values for the initial wind speed of $9m/s$ as compared to the initial wind speed of $5m/s$ case. In this way, the reliability engineer has a clear idea of when the association between two failure events is high or low. This information can also be used to measure the riskiness of a wind park investment in terms of the intermittency of power production.

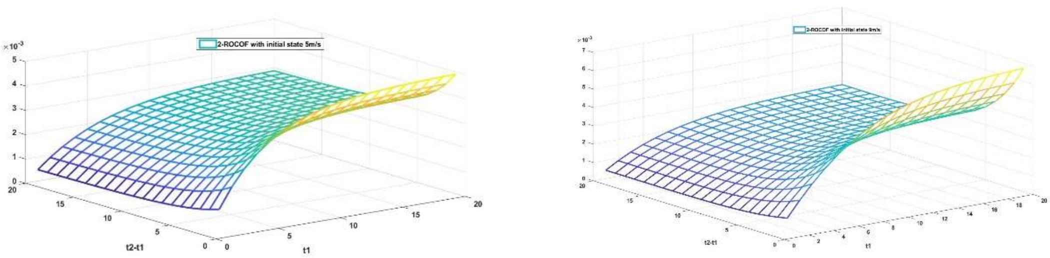

Figure 6. 2-*ROCOF for the Markov process of wind speed for two initial wind speed values*

IV. Discussion

Understanding the rate of occurrence of failures in a random system is of great relevance, both theoretically and practically. This paper offers some general information after commenting on a selection of recent major studies in the field.

- The definition of a recent measure called the n -ROCOF is reported, and a broad interpretation as a measure of clustering in time of failures is given.

- The computation of the n -ROCOF under the hypothesis of a continuous-time finite state space Markov chain is explained in detail. The results are also presented in the more general framework represented by semi-Markov processes.

- A numerical example clarifies the results and shows different shapes of the indicators that are flexible enough to represent a great variety of real system behavior. A new application of these concepts is provided in the field of wind engineering and reveals interesting aspects that need an accurate investigation in a specific research article.

References

[1] Lisnianski, A., I. Frenkel, and Karagrigoriou, A. eds. Recent advances in multi-state systems reliability: Theory and applications, Springer, 2017.

[2] Singh, C. and Billinton, R. (1974). A new method to determine the failure frequency of a complex system, *IEEE Transactions on Reliability*, 23(4):231–234.

[3] Khan, N.M., K. Rajamani, and Banerjee, S.K. (1997). A direct method to calculate the frequency and duration of failures for large networks, *IEEE Transactions on Reliability*, 26(5):318–321.

[4] Iyer, R.K. and Downs, T. (1978). A moment approach to evaluation and optimization of complex system reliability, *IEEE Transactions on Reliability*, 27(3):226–229.

[5] Shi D. (1985). A new method for calculating the mean failure numbers of a repairable system during (0; t], *Acta Mathematicae Applicatae Sinica*, 8(1):101-110.

[6] Yeh, L. (1995). Calculating the rate of occurrence of failures for continuous-time Markov chains with application to a two-component parallel system, *Journal of the Operational Research Society*, 46(4):528-536.

[7] Yeh, L. (1997). The Rate of Occurrence of Failures. *Journal of Applied Probability,* 34(1):234-247.

[8] Grabski, F. *Semi-Markov processes: applications in system reliability and maintenance*. Elsevier, 2014.

[9] Vassiliou, P-CG, and Papadopoulou, A.A. (1992). Non-homogeneous semi-Markov systems and maintainability of the state sizes. *Journal of Applied Probability*, 519-534.

[10] Ouhbi B. and Limnios, N. (2002). The rate of occurrence of failures for semi-Markov processes and estimation. *Statistics and Probability Letters*, 59:245-255.

[11] Georgiadis, S., N. Limnios, and Votsi, I. (2013). Reliability and Probability of First Occurred Failure for Discrete-Time Semi-Markov Systems, *Applied Reliability Engineering and Risk Analysis: Probabilistic Models and Statistical Inference* 167-179.

[12] D'Amico G., (2015). Rate of Occurrence of Failures (ROCOF) of Higher-Order for Markov Processes: Analysis, Inference and Application to Financial Credit Ratings, *Methodology and Computing in Applied Probability*, 17:929-949.

[13] Votsi, I. (2019). Conditional failure occurrence rates for semi-markov chains, *Journal of Applied Statistics*, 46(15):2722-2743.

[14] D'Amico, G. and Petroni, F. (2023). ROCOF of higher order for semi-Markov processes, *Applied Mathematics and Computation*, 441:127719.

[15] Yi, H., Cui, L., Shen, J., and Li, Y. (2018). Stochastic properties and reliability measures of discrete-time semi-Markovian systems, *Reliability Engineering & System Safety*, 176:162-173.

[16] Csenki, A. (1994). On the interval reliability of systems modelled by finite semi-Markov processes, *Microelectron. Reliab.*, 34:1319–1335

[17] Georgiadis, S. and Limnios, N. (2014). Interval reliability for semi-Markov systems in discrete time, *Journal de la Société Française de Statistique,* 155(3):152-166.

[18] D'Amico, G., Manca, R., Petroni, F., and Selvamuthu, D. (2021). On the computation of some interval reliability indicators for semi-Markov systems, *Mathematics*, 9(5):575.

[19] Barbu, V.S., G. D'Amico, and Gkelsinis, T. (2021). Sequential interval reliability for discrete-time homogeneous semi-Markov repairable systems, *Mathematics*, 9(16): 1997.

[20] Billinton, R., and Bollinger, K.E. (1968). Transmission system reliability evaluation using Markov processes, *IEEE Transactions on power apparatus and systems*, 2:538-547.

[21] Jain, Vidyottama, Raina Raj, and Dharmaraja, S. (2022). Performability analysis of a M M A P[2]/P H [2]/S model with PH retrial times, *Communications in Statistics-Theory and Methods*, 1-20.

[22] S., Dharmaraja, Sivama, A. H., Raj, R., and Vishnevskyb, V. (2023). Study of reliability of the on-tether subsystem of a tethered high-altitude unmanned telecommunication platform. *Reliability: Theory & Applications*, *18*(1 (72)):172-178.

[23] Yadav, A.D., N. Nandal, and Malik, S.C. (2023). Markov approach for reliability and availability analysis of a four unit repairable system, *Reliability: Theory & Applications* 18.1(72): 193-205.

[24] Limnios, N. and G. Oprisan. *Semi-Markov processes and reliability*. Springer Science & Business Media, (2001).

[25] Dimitrov, Boyan, Vladimir Rykov, and Sahib Esa. Three Approaches in the Study of Recurrent Markovian and Semi-Markovian Processes. *Distributed Computer and Communication Networks: 23rd International Conference, DCCN 2020, Moscow, Russia, September 14–18, 2020, Revised Selected Papers 23*. Springer International Publishing, 2020.

[26] D'Amico, G., Regnault, P., Scocchera, S., and Storchi, L. (2018). A continuous-time inequality measure applied to financial risk: the case of the European Union, *International Journal of Financial Studies*, 6(3):62.

[27] Vergine, S., Álvarez-Arroyo, C., D'Amico, G., Escaño, J. M., and Alvarado-Barrios, L. (2022). Optimal management of a hybrid and isolated microgrid in a random setting. *Energy Reports*, *8*:9402-9419.

[28] Shamshad, A., Bawadi, M. A., Hussin, W. W., Majid, T. A., and Sanusi, S. A. M. (2005). First and second order Markov chain models for synthetic generation of wind speed time series. *Energy*, *30*(5):693-708.

[29] Bladt, M., and Sørensen, M. (2005). Statistical inference for discretely observed Markov jump processes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *67*(3):395-410.

[30] D'Amico, G., and Regnault, P. (2018). Dynamic measurement of poverty: Modeling and estimation. *Sankhya B*, *80*(2):305-340.