

Article

Initial Results of an Extensive, Long-term Study of the Forecasting of Voltage Sags[†]

Michele De Santis ¹, Leonardo Di Stasio ^{2,*}, Christian Noce ³, Paola Verde ² and Pietro Varilone ²

¹ Department of Ingegneria, Università Niccolò Cusano di Roma, 00166 Roma, Italy; michele.desantis@unicusano.it

² Department of Ingegneria elettrica e dell'informazione "Maurizio Scarano", Università di Cassino e dell'informazione, 03043 Cassino, Italy; verde@unicas.it (P.V.); varilone@unicas.it (P.V.)

³ E-distribuzione S.p.A., 00198 Roma, Italy; christian.noce@e-distribuzione.com

* Correspondence: leonardo.distasio@unicas.it

[†] This paper is an extended version of our paper published in the Proceedings of the 55th International Universities Power Engineering Conference (UPEC), Torino, Italy, September 1–4 2020, pp. 1–6, doi:10.1109/UPEC49904.2020.9209816.

Abstract: This paper presents the preliminary results of our research activity aimed at forecasting the number of voltage sags in distribution networks. The final goal of the research is to develop proper algorithms that the network operators could use to forecast how many voltage sags will occur at a given site. The availability of four years of measurements at Italian Medium Voltage (MV) networks allowed the statistical analyses of the sample voltage sags without performing model-based simulations of the electric systems in short-circuit conditions. The challenge we faced was to overcome the barrier of the extremely long measurement times that are considered mandatory to obtain a forecast with adequate confidence. The method we have presented uses the random variable *time to next event* to characterize the statistics of the voltage sags instead of the variable *number of sags*, which usually is expressed on an annual basis. The choice of this variable allows the use of a large data set, even if only a few years of measurements are available. The statistical characterization of the measured voltage sags by the variable *time to next event* requires preliminary data-conditioning steps, since the voltage sags that are measured can be divided in two main categories, i.e., rare voltage sags and clusters of voltage sags. Only the rare voltage sags meet the conditions of a Poisson process, and they can be used to forecast the performance that can be expected in the future. However, the clusters do not have the characteristics of memoryless events because they are sequential, time-dependent phenomena the occurrences of which are due to exogenic factors, such as rain, lightning strikes, wind, and other adverse weather conditions. In this paper, we show that filtering the clusters out from all the measured sags is crucial for making successful forecast. In addition, we show that a filter, equal for all of the nodes of the system, represents the origin of the most important critical aspects in the successive steps of the forecasting method. In the paper, we also provide a means of tracking the main problems that are encountered. The initial results encouraged the future development of new efficient techniques of filtering on a site-by-site basis to eliminate the clusters.

Keywords: power quality; forecasting method; voltage sags

Citation: De Santis, M.; Di Stasio, L.; Noce, C.; Verde, P.; Varilone, P. Initial Results of an Extensive, Long-term Study of the Forecasting of Voltage Sags. *Energies* **2021**, *14*, 1264. <https://doi.org/10.3390/en14051264>

Academic Editor: José Matas

Received: 8 January 2021

Accepted: 17 February 2021

Published: 25 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The presence of conducted disturbances in the voltages and currents contributes to the problems of power quality (PQ). As defined in EN50160 [1], the disturbances can be grouped into two main categories, i.e., variations and events. Typical variations are the harmonic and inter-harmonic distortions, unbalances, and fluctuations. Typical events are interruptions, voltage sags, and voltage swells; the interruptions fall into the wide

topic of continuity of supply, which traditionally is approached in a separate manner from the other PQ disturbances.

Among the PQ disturbances, voltage sags are the dominant occurrences that cause detrimental effects to a large variety of equipment and apparatuses. The main effect of a voltage sag is the tripping of devices with several adverse consequences on the system in which the devices are located. For example, automated manufacturing lines and financial and bank services can experience complete stoppages of their ability to function, and, as a result, the owners of the plants that have automated manufacturing lines, in particular, may incur huge costs as a result. The literature of reference dealt with the problems of voltage sags with several studies aimed at characterizing the phenomenon [2], modelling systems in the presence of the short circuits that cause the sags [3,4], proposing adequate indices for quantifying their severity [5–7], deriving methods and tools for compensating them [8–10], and estimating their related costs [11–14]. A significant portion of the related literature is based on the simulation of the electric system in presence of voltage sags, and the main methods are the fault position method (FPM) and the critical distance method (CDM) [3,15–19]; both methods require an understanding of the topology and typology of the system and of all its components in order to develop adequate models in short-circuit conditions.

All of the available simulation methods are used for estimating the average performance of the nodes of the electric system and/or of the entire system in terms of the expected number of voltage sags, which usually is expressed as yearly frequency. For this objective, the methods assume that the average frequency of short circuits caused by the failure of any component of the electric system is known, and it is modelled by its fault rate, λ , which usually is expressed on a yearly basis. In fact, being aware of the expected performance is one of the most attractive themes for the system operators, i.e., both distribution system operators (DSOs) and transmission system operators (TSOs), who are interested in understanding the future features of their system well enough to implement adequate maintenance and/or effective sag compensation to prevent service interruptions. The importance of this subject has been reinforced significantly by the regulatory activities of the national energy authorities that started with several actions in essentially every European country [20,21]. In Italy, the national Authority of Energy, Networks, and Environment (ARERA, in Italian, Autorità di Regolazione per Energia Reti e Ambiente, formerly AEEGSI, in Italian, Autorità per l'energia Elettrica il Gas e il Sistema Idrico) requires every DSO to register all of the voltage sags at each busbar at medium voltage (MV) of every High Voltage/Medium Voltage (HV/MV) substation [22,23].

The DSO, starting from 2014, installed the measurement devices in respect to CEI EN 61000-4-30 [24] as depicted in Figure 1 where two measuring systems (MSs) are installed (MS1 and MS2) at the two MV busbars. It is evident in Figure 1 that there is a switch between the busbars, which is in open state in the normal operating conditions. The presence of the switch, normally in open state, between the MV busbars at every HV/MV substations determined a number of MSs larger than the number of HV/MV substations. In the real distribution system of regional dimensions studied in this paper, we had 152 MSs for 83 HV/MV substations of which 17 substations are equipped with one HV/MV transformer (one MS), 63 substations with two HV/MV transformers (two MSs for each substation) and three substations with three HV/MV transformers (three MSs for each substation).

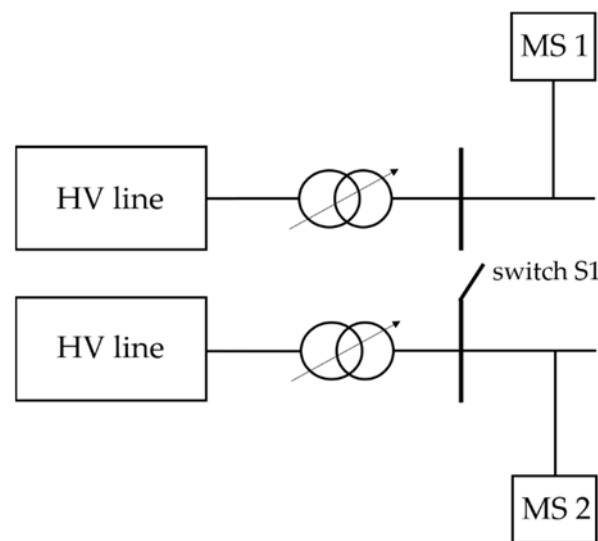


Figure 1. Simplified scheme of an example of an HV/MV (High Voltage/Medium Voltage) substation equipped with two measuring systems, i.e., MS1 (Measurement System 1) and MS2 (Measurement System 2), installed at the MV (Medium Voltage) busbars.

In this paper, we addressed the subject forecasting the performance of a real system by using the measurements obtained by the MSs installed at the HV/MV substations within an Italian Region due to the regulation of ARERA. This study was conducted without performing any simulations, and only the available data were used to conduct the statistical analyses. This paper is an extended version of our earlier paper [25], in which the results of the preliminary studies were presented. These preliminary studies were focused mainly on the evaluation of the statistical techniques that were used to assess the input data that came from the field. The successive phases of the study will analyse and compare other techniques, beginning with those considered in this paper, for applying those that proved to be the most suitable and efficient for completely forecasting the expected performance of real systems. The overriding goal of this study is to give to the network operators proper algorithms that will allow them to forecast how many voltage sags will occur and the sites at which they are likely to occur. Also, this extended version of the paper [25] allows identifying the details of the main critical aspects that are encountered, and it proposes adequate solutions, which are in progress.

Comparing the method used in this paper with the model-based methods in the literature [2,4], it is worth noting that the data measured in the field by other methods only provide the statistics of the failure rates of the components. These data were combined successively with the results of the simulations performed by FPM, CDM, or other analytical models. Instead, the data used in our method are just the voltage sags measured in real systems, and these data are used directly to forecast how many voltage sags will occur at which sites in the future, usually during the incoming year. In addition, in this paper, we analyzed very large datasets that came from the measurements taken during 4 consecutive years at 152 sites in a real distribution system. Such data are not readily available in the relevant literature.

As stated earlier, the expected performance in terms of voltage sags in the literature was assessed by simulating the system in a short-circuit condition. To the best of the authors' knowledge, no study to date has used voltage sags measured in real systems for forecasting future performance. The main barrier to this approach was the time required to obtain statistical descriptions of the random variable *number of voltage sags* with adequate confidence. It always was stated that measurements were needed over a period of several years, even one or more decades. In effect, voltage sags are rare events, and consequently a prohibitive amount of time would be required to obtain a sufficiently large number of field measurements to forecast with adequate confidence the variable *number*

of voltage sags. However, this impediment can be overcome by choosing a different random variable for describing the voltage sags from a *number of voltage sags*.

As first proposed in [26], the variable *time to next event*, *ttne*, allows the statistical description of the phenomenon, just as the random variable *number of voltage sags* does, but with the great advantage of obtaining a larger set of measured data than the set obtained by means of the variable *number of voltage sags*. The random variable, *ttne*, is defined as the time that elapses between the end of a voltage sag to the beginning of the next voltage sag.

We used the random variable, *ttne*, to characterize the voltage sags measured during a period of 4 years (2014–2017), in an actual MV distribution network. The goal was to statistically characterize the first 3 years (2014–2016) and use that information to forecast the fourth year (2017). The statistical characterization of the first 3 years required the initial handling of the measured data, and the main steps were the time aggregation for collecting the multiple sags, the detection and removal of the voltage sags that appeared as clusters, the validation of the voltage sags that remained after the clusters were removed, and the use of the data that were obtained to make the final forecast.

In this paper, which is an extended version of an earlier paper [25], we presented the preliminary results of the study, indicating, with more details than were presented in [25], the main critical aspects encountered and some solutions to overcome them.

The paper is organized as follows: Section 2 describes in detail the characteristics of the measured data and the first handling of the data with the goal of reducing the quantity of data. Section 3 presents the approach that was used to describe the statistical characteristics of the voltage sags measured in real systems. Section 4 illustrates the procedure for the assessment of the distributional assumptions about the data. Section 5 shows the results of the forecasting activity, the critical aspects, and some possible solutions.

2. Main Characteristics of the Measured Data and Time Aggregation for Multiple Sags

In this paper, the same regional system presented in [25] was considered. The system has 83 HV/MV substations with 152 MV busbars that represent the points of measurement of the voltage sags (we discarded three sites of the network since they registered in all the considered years a number of voltage sags less than 50; therefore, we considered 149 sites). The measured voltage sags occurred in 4 consecutive years, i.e., from 1 January 2014 to 12 December 2017. The data were available due to the cooperation of E-Distribuzione S.p.A., the largest DSO (Distribution System Operator) in Italy, which, in respect to the resolutions of ARERA [22,23], installed MSs in all of the HV/MV substations it managed. In particular, E-Distribuzione S.p.A (Rome, Italy), like every other DSO in Italy, was required to provide a summary table relating to voltage sags that occurred every year. The registered data were composed of several parameters of the measured voltage sags in every MV busbar of every HV/MV substation. All of the data were collected in a format table [27], such as Table 1. With reference to Table 1, the data are:

- the sequential number of the sag (#);
- the code that identifies the HV/MV substation (ID Sub);
- the code that identifies the MV busbar (ID Bus);
- the phase involved (RS and/or ST and/or TR);
- the start time of the sag in terms of day, month, year, hour, minute, second with a least a resolution of hundredths of a second (T_{occ});
- the duration of the sag with at least resolution of hundredths of a second (D);
- the nominal voltage in kV (V_n);
- the residual voltage both in kV and in percentage of V_n (V_r);
- the intervention of the distance protection (DP);
- the intervention of the maximum current protection of the transformer ($MaxI_{Tr}$) (the type of voltage sag if true (T), false (F), not available (NA) The classification

true-false-not available refers to the possible effects of the voltage transformer saturation used in the MSs, which can result in false voltage sags. The resolution of ARERA [22] required DSOs to install a proper algorithm in the MSs for identifying the true voltage sags, distinguishing them from false sags and from those whose identification is not unique (indicated in the table with not available). The algorithm is based on the measurement of the asymmetry of the voltage waveform during the sag [28].

- the system, HV or MV, where the sag was originated (Origin).

Table 1. Format of recorded data relating to voltage sags.

#	ID Sub	ID Bus	RS	ST	TR	T _{occ}	D	V _n	V _r	V _f	DP	MaxI _{TR}	Type	Origin
1	Id_Sub Code	Id_Bus Code	X	X	X	(gg/mm/aaaa hh:mm:ss.cc)	(hh:mm:ss.cc)	kV	kV	%	Yes No	Yes No	T F NA	HV MV
.
M

It is important to show that the phase aggregation was applied by the DSO before the registration of the data. In fact, as indicated in [24], the voltage sags on multiple phases were recorded as only one sag with a residual voltage V_r , corresponding to the lowest value among the phases, and a duration D , corresponding to the highest value among the phases. In addition, the DSO must indicate in which system (HV or MV) the fault occurred that caused that measured sag. The method for establishing the system in which the voltage sag originated was proposed by ARERA in [22]; it uses only the measures on the voltages during the sags without any additional measures, for example, on the currents or on the voltages before the sags. Other methods have been proposed in the literature [29,30] that can ascertain the origin by analyzing only the registered data of the voltages during the sag, thereby overcoming the problems of the results obtained with the method [24] for the presence of distributed generator units and the different positions of the on-load tap changing transformer.

The 39,137 voltage sags that were registered from 2014 to 2016 were used to forecast the sags in 2017. On all of these sags, the first step consisted of exploring the multiple sags, i.e., the sags that occurred in succession in a very short time.

The data from the field, in fact, proved that the measured sags can be divided into two main categories, i.e., the rare voltage sags and the multiple sags. The multiple sags appear as grouped events with the intervals of the adjacent voltage sags concentrated in $1\sim 10^5$ s [5,6,31]. The sags registered in the Italian real MV networks, which were analyzed in this paper, also showed times between events that were less than 2 s.

The first action on the multiple voltage sags was the aggregation of time, which meant that all of the events grouped within a given time interval were considered as a single event, which is the first of the group. As discussed in [6], the main reasons to aggregate these multiple sags are:

- 1) the effect on the operation of equipment and processes;
- 2) the possibility of incorrect registration;
- 3) the common cause that originated them.

Regarding the effect on the operation of equipment and processes, if the first voltage sag of the group will trip the equipment, causing the process to stop, the second and the successive sags will find the equipment still down. Consequently, it is reasonable to aggregate the events, separated from each other by a time that is less than the recovery time of the equipment and/or process them as a single event.

Regarding the possibility of the incorrect registration of multiple sags, the algorithm used in the measurement systems may register the end of a sag before the event is over. In fact, the voltage in real cases could recover above the threshold of sag-ending for some cycles and, afterwards, drop below the threshold again.

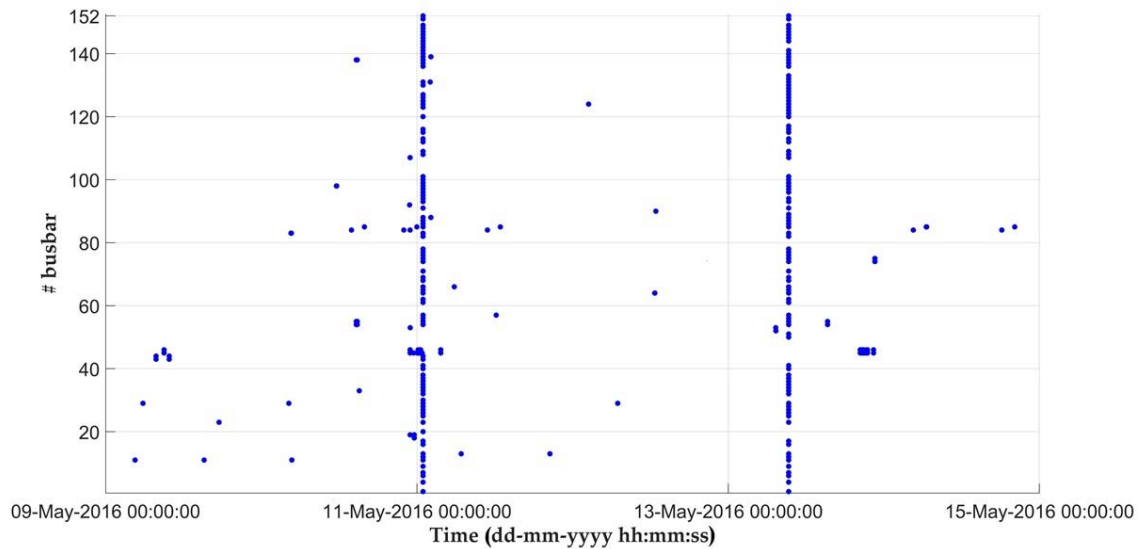


Figure 3. Scatter plot of the data registered during the year 2016 before time aggregation.

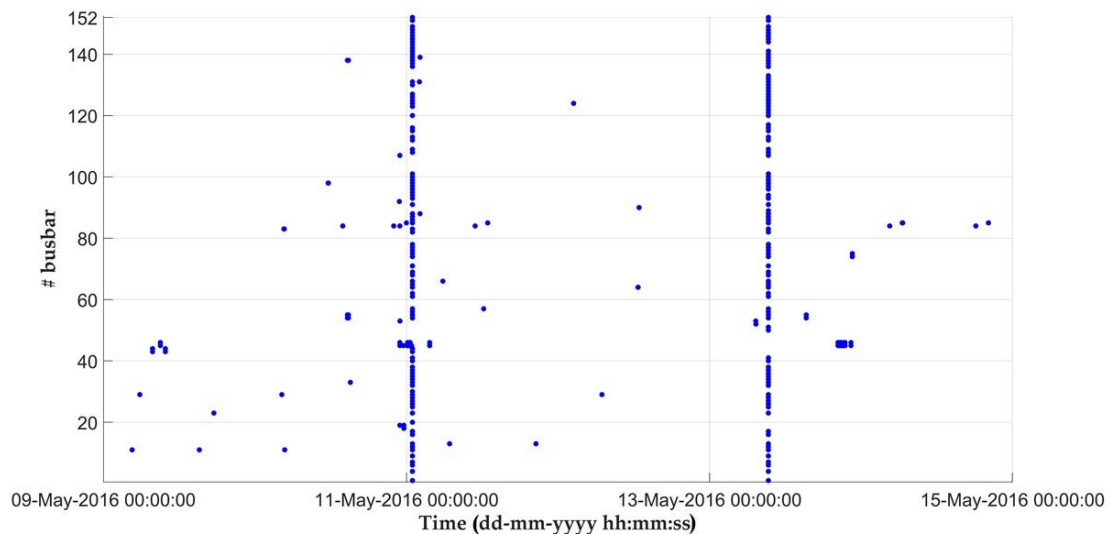


Figure 4. Scatter plot of the data registered during the year 2016 after time aggregation.

Figures 5 and 6 present the zooms of Figure 3 and Figure 4, respectively, on only one day (13 May 2016) applied to only two sites, i.e., sites #45, and #46. The zooms allowed the simultaneous sags to be ascertained, and they were measured at these after 06:00 a.m. These sags were present before the time aggregation (as indicated in Figure 5 by the arrows), and they also were maintained after the time aggregation (as indicated by the arrows in Figure 6). Further, the zooms indicated that the distribution of the sag was not modified over time by the time aggregation. In fact, comparing the sags in Figure 5 with those in Figure 6, it was evident that the sags were distributed in both figures as rare sags (blue circles) and clusters (black circles).

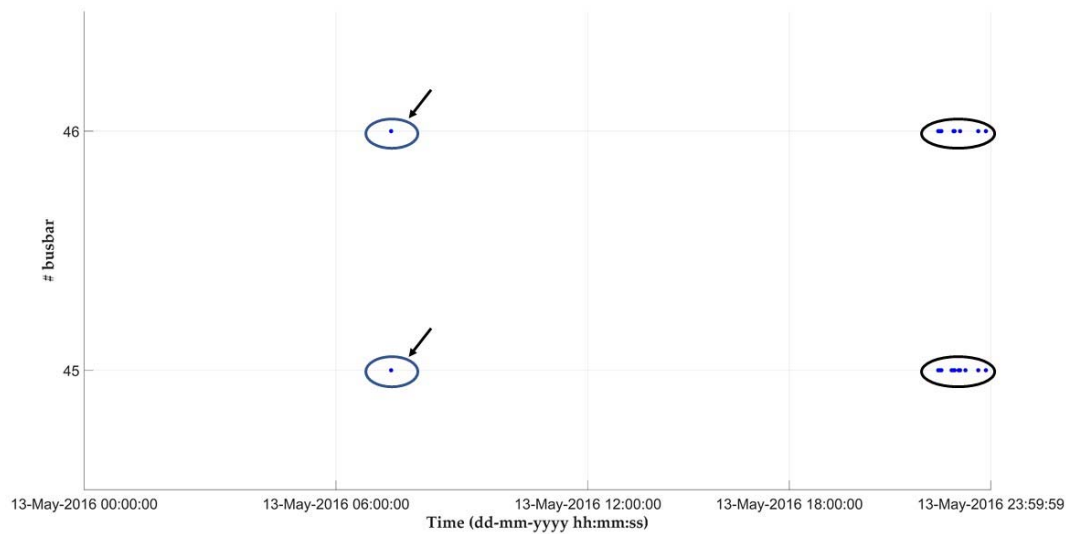


Figure 5. Scatter plot of voltage sags measured on 13 May 2016 at the sites #45 and #46 before the time aggregation. The blue circles indicate the rare sags; the black circles indicate clusters; the arrows indicate simultaneous sags.

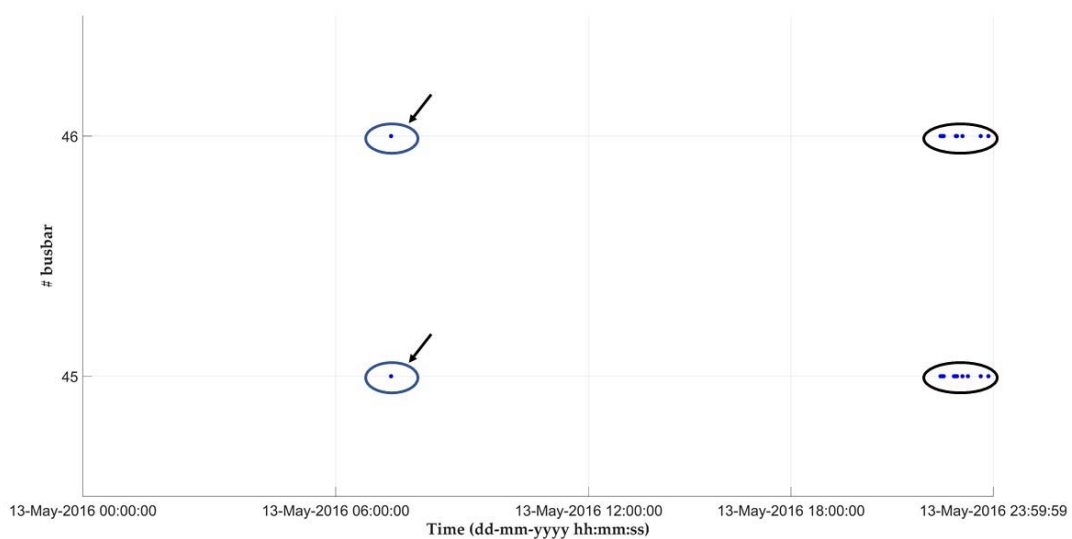


Figure 6. Scatter plot of voltage sags measured on 13 May 2016 at the sites #45 and #46 after the time aggregation. The blue circles indicate the rare sags; the black circles indicate the clusters; the arrows indicate the simultaneous sags.

Further verifications of the acceptability of the effects of the time aggregation were performed also on the plane (Amplitude, Duration) (the amplitude of a voltage sag is the minimum RMS (Root Mean Square) of the voltage during the sag; it is also called residual voltage) representing the measured sags from 2014 to 2016. Figures 7 and 8 show these plots for the sites #20, #44, #74, and #75 before and after the time aggregation, respectively. The distribution of the sags in the plane (Amplitude, Duration) appeared unchanged after the time aggregation preserving the sag features in this plane.

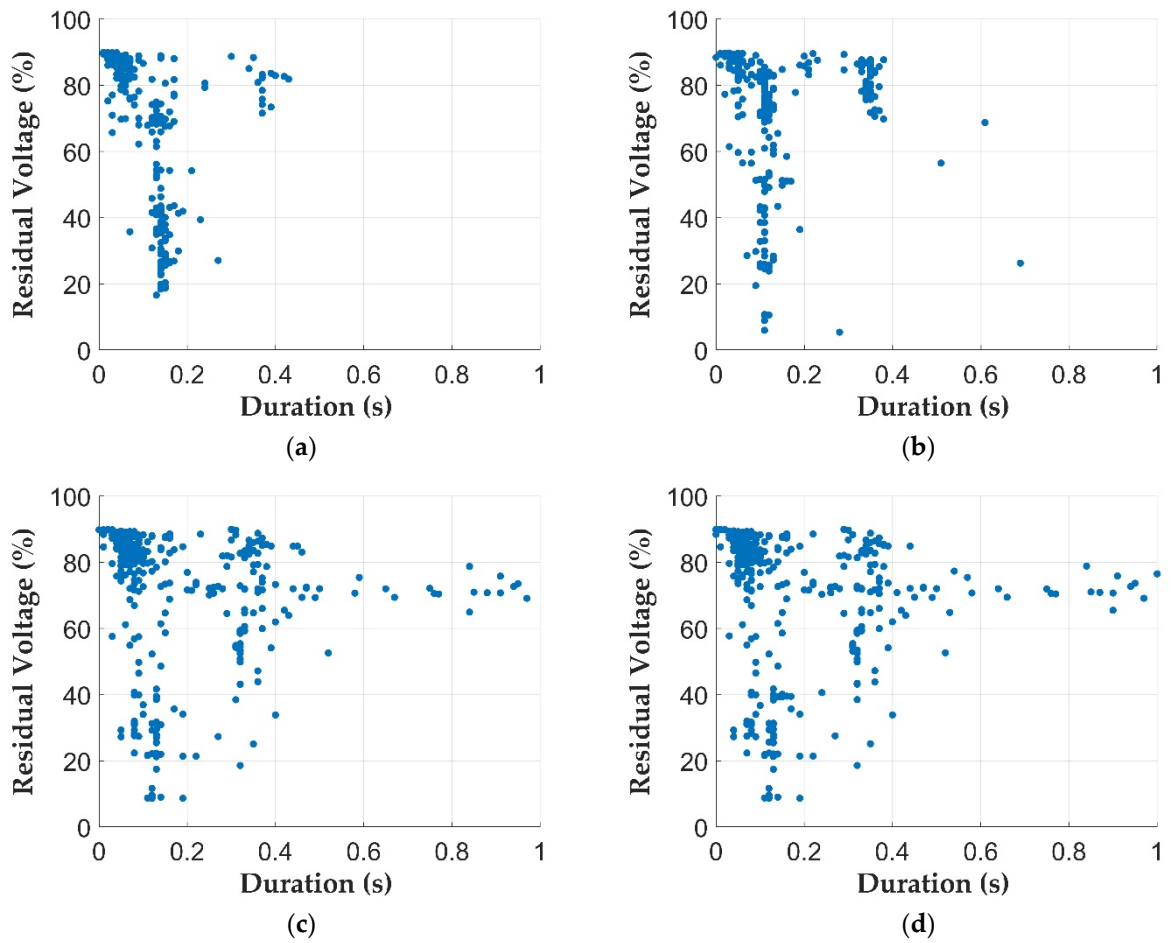
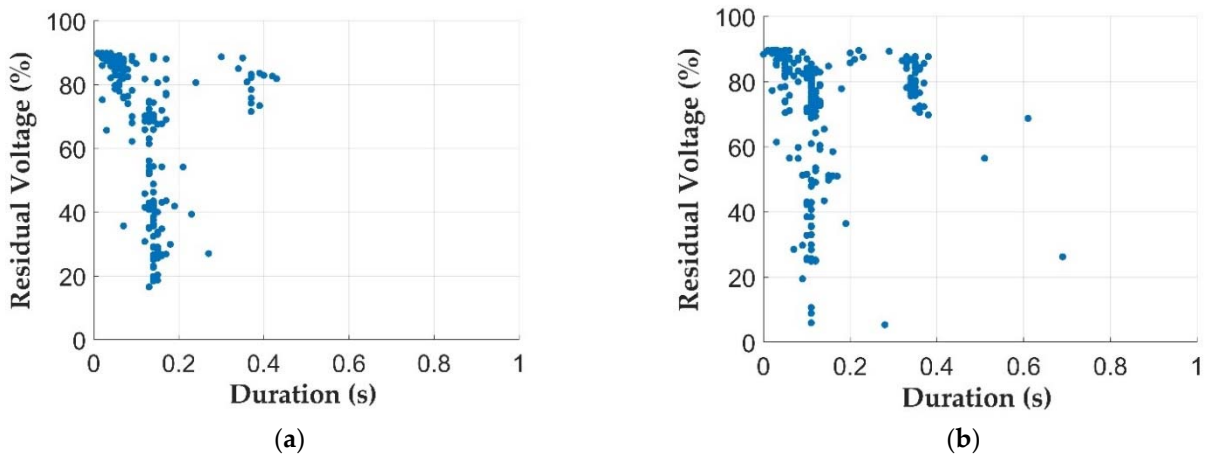


Figure 7. Scatter plots (Residual Voltage, Duration) of the voltage sags measured in 2014, 2015, and 2016 at the site #20: (a), site #44; (b), site #74; (c), and site #75: (d), before time aggregation.



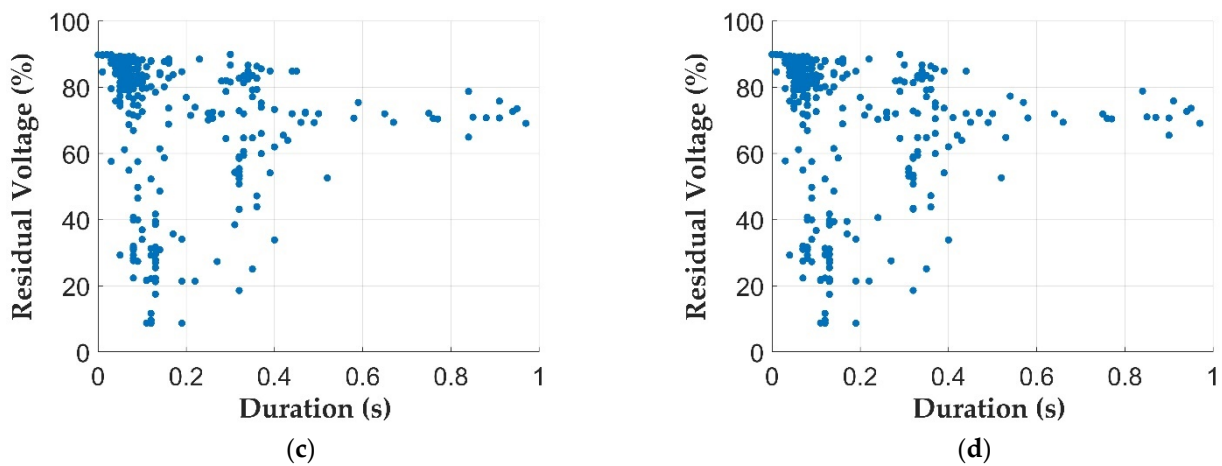


Figure 8. Scatter plot (Residual Voltage, Duration) of the voltage sags measured in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d), after time aggregation.

Figures 9 and 10, respectively, show the cumulative distribution, F , of the amplitude and of the duration before and after the time aggregation of the sags measured from 2014 to 2016 at sites #20, #44, #74, and #75. It is evident that the time aggregation had no substantial impact on the values of the cumulative distributions.

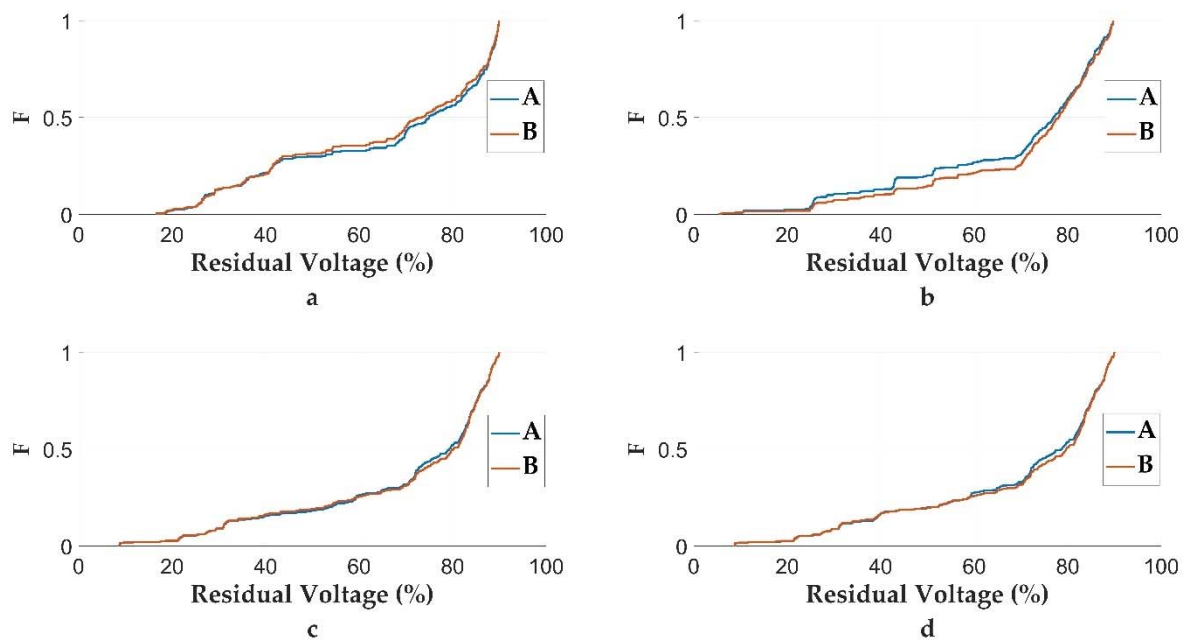


Figure 9. Cumulative distribution function, F , of the residual voltage of voltage sags measured in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c) and #75: (d), A before time aggregation and B after time aggregation.

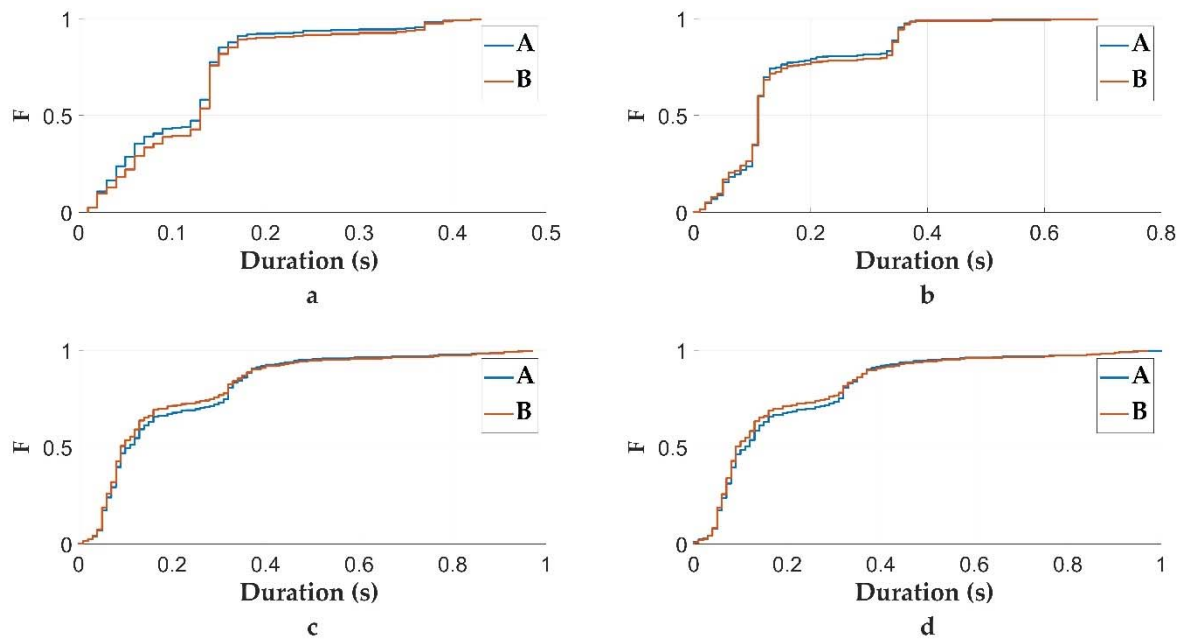


Figure 10. Cumulative distribution function, F , of the duration of voltage sags measured in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c) and #75: (d), before time aggregation A, and after time aggregation B.

To quantify any possible modification of the original data due to time aggregation, the sags measured in all of the sites of the regional system that was studied from 2014. to 2016 were classified by tables, such as those provided by [1]. Tables 2 and 3 furnished the number of sags in every cell expressed in percentage of the total number of sags, which was reported in the caption of the table. Tables 2 and 3 show also references to the susceptibility curves provided by [32]. The cells in grey are those above the curve of Class II, and the cells above the red line are those above the curve of Class III.

Table 2. Distribution of the voltage sags measured in the regional system before the aggregation of time; total number of sags = 39,137. The cells in grey are those above the curve of Class II, and the cells above the red line are those above the curve of Class III.

Residual Voltage V_r (%)	Duration (ms)				
	$10 \leq t \leq 200$	$200 < t \leq 500$	$500 < t \leq 1000$	$1000 < t \leq 5000$	$5000 < t \leq 60000$
$90 > V_r \geq 80$	40.20	10.17	1.17	1.03	0.18
$80 > V_r \geq 70$	15.94	4.74	0.35	0.35	0.01
$70 > V_r \geq 40$	16.69	2.81	0.37	0.22	0.04
$40 > V_r \geq 5$	4.36	0.52	0.12	0.02	0.01
$5 > V_r$	0.05	0.00	0.00	0.00	0.00

The comparison of Tables 2 and 3 reveals that the time aggregation did not significantly modify the classification of the sags in terms of their amplitude and duration.

Concerning the classification of the sags with respect to the susceptibility curves, the total percentage of sags above the limit of Class II (Class III) passed from 71.05% (89.94%) before the time aggregation, to 71.44% (90.43%) after the time aggregation with a maximum variation of 0.5%. Similar considerations also are valid for the deeper and longer sags, which are a small part of the totality of the sags, as often is the case in actual systems.

Table 3. Distribution of the voltage sags measured in the regional system after the time aggregation; total number of sags = 33,602. The cells in grey are those above the curve of Class II, and the cells above the red line are those above the curve of Class III.

Residual Voltage V_r (%)	Duration (ms)				
	$10 \leq t \leq 200$	$200 < t \leq 500$	$500 < t \leq 1000$	$1000 < t \leq 5000$	$5000 < t \leq 60,000$
$90 > V_r \geq 80$	41.65	8.99	1.15	0.95	0.21
$80 > V_r \geq 70$	16.53	4.27	0.39	0.34	0.01
$70 > V_r \geq 40$	16.89	2.37	0.40	0.22	0.04
$40 > V_r \geq 5$	4.22	0.41	0.14	0.01	0.00
$5 > V_r$	0.06	0.00	0.00	0.00	0.00

3. The Use of *ttne* for the Statistical Analysis of Measured Voltage Sags: Clusters and Rare Sags

In the relevant literature, the only way to characterize the performance of the system in terms of the expected number of voltage sags, usually expressed as yearly frequency, was to resort to simulations based on the system modelling in short circuit conditions [3–7,15–19,27,33]. It was considered unfeasible to characterize the expected number of sags derived from measurements because decades of years were needed to obtain the statistical description of the number of voltage sags suitable for forecasting the future performance with adequate confidence [34].

The main reason was linked to the characteristics of rare events, e.g., voltage sags, together with the choice of the random variable *number of voltage sags* in a period of time, usually one year, which was assumed to be a Poisson variable [35].

Following the approach first proposed in [26], in this paper, as was done in [25], we used the random variable *time to next event* (*ttne*) rather than the random variable *number of sags per year*. The random variable *ttne* was defined as the time that elapsed between the end of an event and the start of the next event. This fundamental choice allowed a huge increase in the data available for describing each node, certainly more than only one datum constituted by the *number of sags* per year. Figure 11 illustrates a theoretical example to describe this aspect.

Let us assume that six voltage sags were measured in a generic site in a time period ΔT . Figure 11 shows the residual voltage, V_s , of these six voltage sags along the time axis; in Figure 11, the time that elapsed between every couple of voltage sags also is represented by the values of the variable *ttne*, i.e., from *ttne1* to *ttne5*. If this site was characterized by the variable *number of sags*, the available value of this variable would be only one, equal to 6. Instead, if this site was characterized by *ttne*, the values available for this different random variable are 5, i.e., $N-1$, where N represents the number of voltage sags at that site.

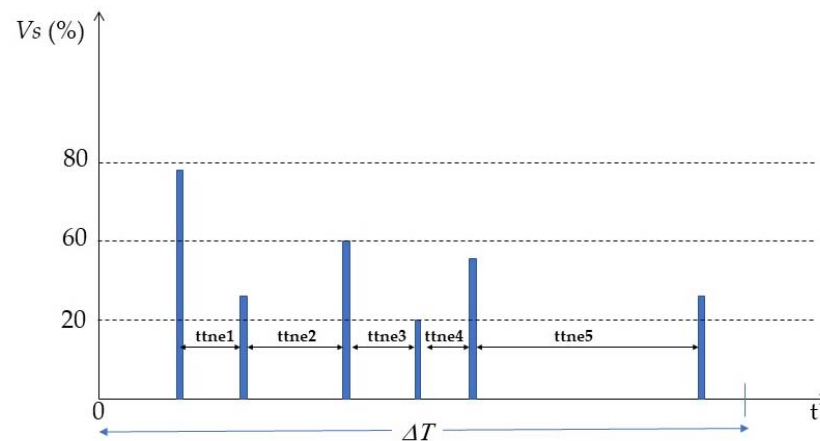
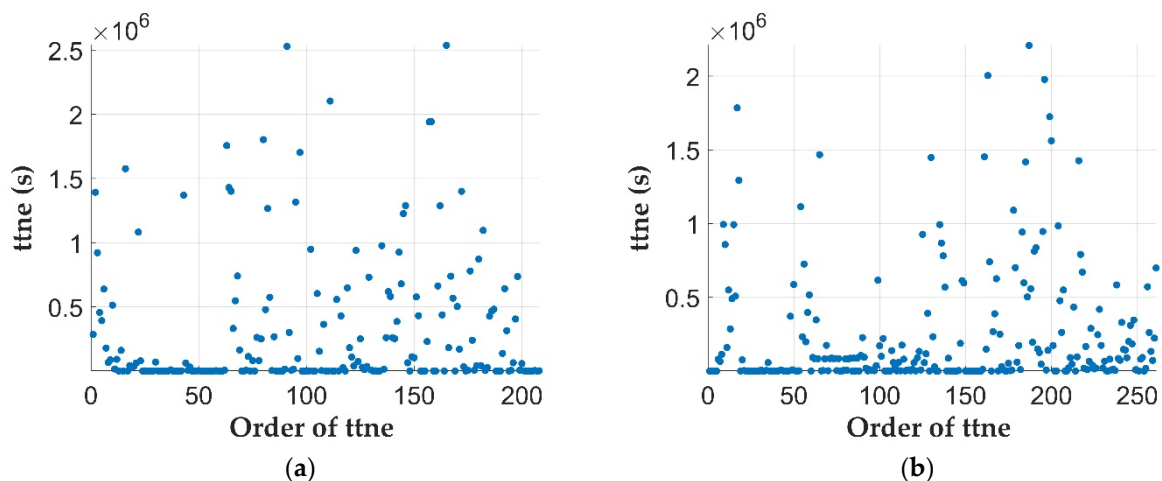


Figure 11. Theoretical example of the use of the random variable time to next event, $ttne$.

Since it was impossible to show the characteristics of the voltage sags measured at all of the sites of the system, we selected four sites as examples of those in which a sufficient number of voltage sags were measured. This criterium allowed better demonstrating the intermediate and final findings of the method. These sites, i.e., sites #20, #44, #74, and #75, are used as examples throughout the paper. Figure 12 shows the scatter plots of the $ttne$ registered at the example sites from 2014 to 2016. For these four sites, since the values of $ttne$ were plotted as functions of the variable order of $ttne$, which measured the order of the voltage sags that occurred in succession; for example, the $ttne$ that corresponds to the order of $ttne$ equal to 1 is the $ttne$ between the second sag and the first sag; the $ttne$ that corresponds to the order of $ttne$ equal to 2 is the $ttne$ between the third sag and the second sag, and so on. To help in reading the plots, the point $(300, 0.5 \times 10^6 \text{ s})$, highlighted with a red circle in Figure 12d), means that the time that elapsed between the 301st and the 300th voltage sag was $0.5 \times 10^6 \text{ s}$ (138.8 h, which is less than 6 days).



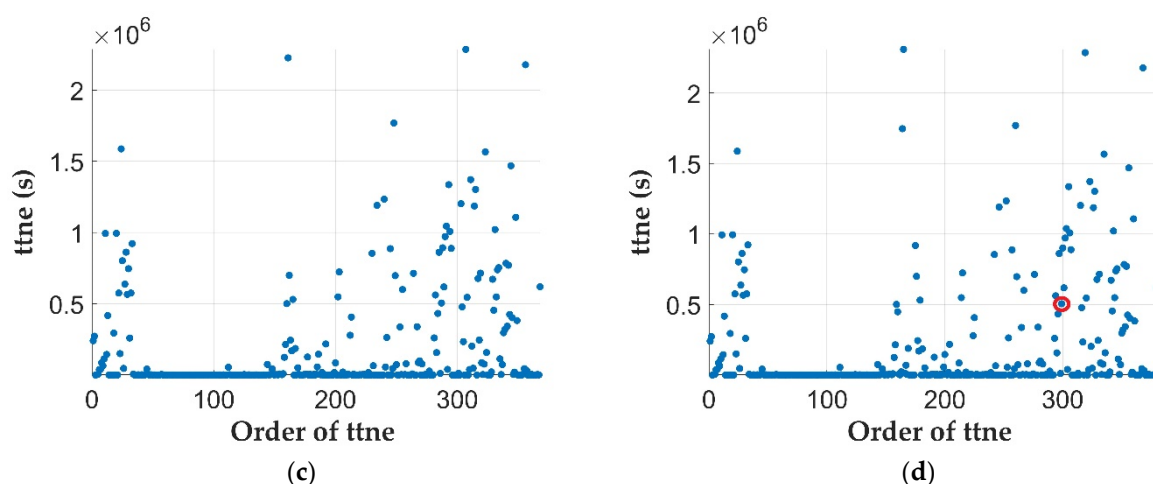


Figure 12. Scatter plot of ttn registered in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d). The red circle highlights that the time that elapsed between the 301st and the 300th voltage sag was 0.5×10^6 s (138.8 h, which is less than 6 days).

The plots of Figure 12 show that ttn had several values that were very close to the x axis; they always were greater than 2 s, even if the scale of the y axis did not allow distinguishing 2 s from zero. This important characteristic, observed for all the sites of the system in this study, proved that the measured sags were characterized by two categories, i.e., sags grouped close to each other, i.e., clusters, and sags dispersed over time, i.e., rare sags. Similar evidence also was observed in [26] with voltage sags measured in other electric systems.

Figure 13 shows the cumulative distribution functions, F , of the ttn computed for voltage sags registered from 2014 to 2016 at the #20, #44, #74, #75 sites. For the same busbars, Table 4 provides the value of a specific example point that cannot be ascertained in the plots of Figure 13 for the time scale of the x axis. In particular, Table 4 shows the $P_{50\%}$ percentile, i.e., the value of ttn such that the probability of having values equal to or less than ttn equals 50%.

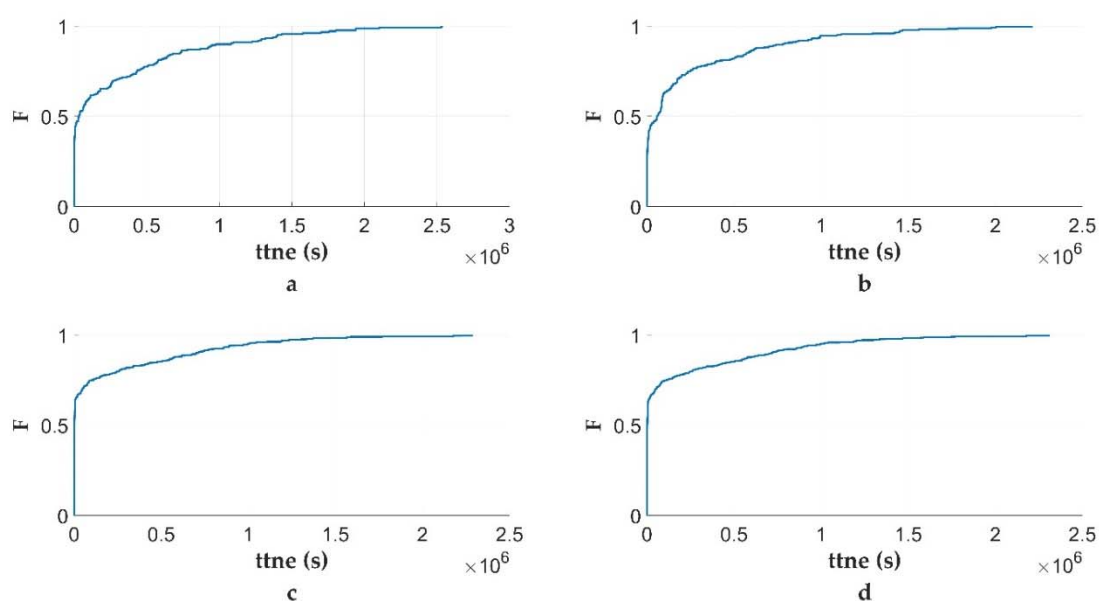


Figure 13. Cumulative distribution function F of the ttn measured in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d).

Table 4. Values of the percentile $P_{50\%}$ of ttne registered at busbars #20, #44, #74 and #75.

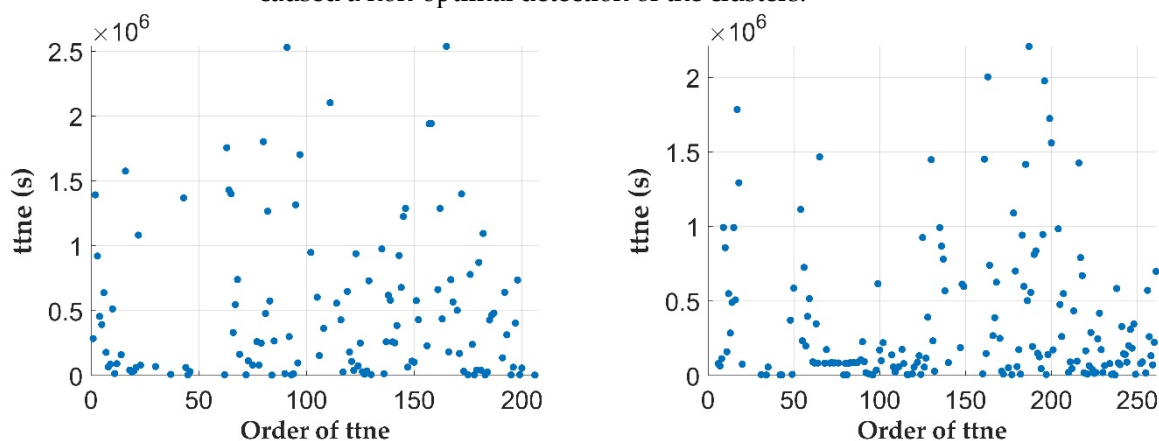
Site #	$P_{50\%}(s)$	$P_{50\%}(h)$
20	35,000	9.7
44	64,000	17.8
74	1836	0.5
75	1951	0.5

Figure 13 and Table 4 show that several clusters were present; in particular, the initial value of the cumulative probability, F , of the example sites was always greater than 40%, and the values were greater than 60% for sites #74 and #75; in addition, in any of the plots in Figure 12, about 80% of the ttne was not greater than 0.5×10^6 s (138.8 h, which is less than 6 days). These observations indicate that the largest part of the measured sags occurred as clusters.

The sags grouped in clusters do not have the statistical characteristics of a Poisson process, i.e., a model for a series of discrete events such that the arrival of an event is independent of the preceding event [36]. Instead, the clusters appear as time-dependent events the occurrence of which is due to exogenous factors, such as rain, lightning strikes, wind, and other adverse weather conditions. Consequently, as also was done in [25], the successive step consisted of detecting and removing the clusters from all of the measured data.

As previous stated, [26] is the only reference in the literature that statistically analyzed the voltage sags that were measured on a real system. Consequently, for sake of comparison with the literature, first, we used the same criterion on the ttne introduced in [26]. The criterion, presented as 1h-filter in the following, was applied in this paper in the same manner of [26], i.e., the same criterion at every site. It consisted of considering as clusters all of the voltage sags measured in succession with an hour or less between the measurements (i.e., $ttne \leq 1$ h). The basics of this criterion in [26] was described by the observation of the statistical characteristics of voltage sags measured at 4 monitored busses. The authors of [26] noticed that, for low values of ttne, the empirical distribution had a higher number of observations than it should have had considering the theoretical exponential distribution. This suggested that events of lower duration, i.e., observation values less than about 1 h, might be interdependent, i.e., related to each other, showing the existence of clusters. Then, this 1h-filter was applied to detect and remove the clusters at all busses, as we did also in this paper for sake of comparison.

Figures 14 and 15 show the scatter plots of the rare voltage sags and the clusters, respectively, obtained after the application of the 1h-filter on the voltage sags measured from 2014 to 2016 at sites #20, #44, #74, and #75, which are the same sites of the scatter plots in Figure 12. Comparing these plots with those reported in Figure 12, it is evident that a great part of the clusters was filtered out, even if some points in the scatter plots of the rare sags (Figure 14) still remain close to the x axis. This evidence represented the first critical aspect of this initial study. The 1 h-filter, applied to every site in the same way, caused a non-optimal detection of the clusters.



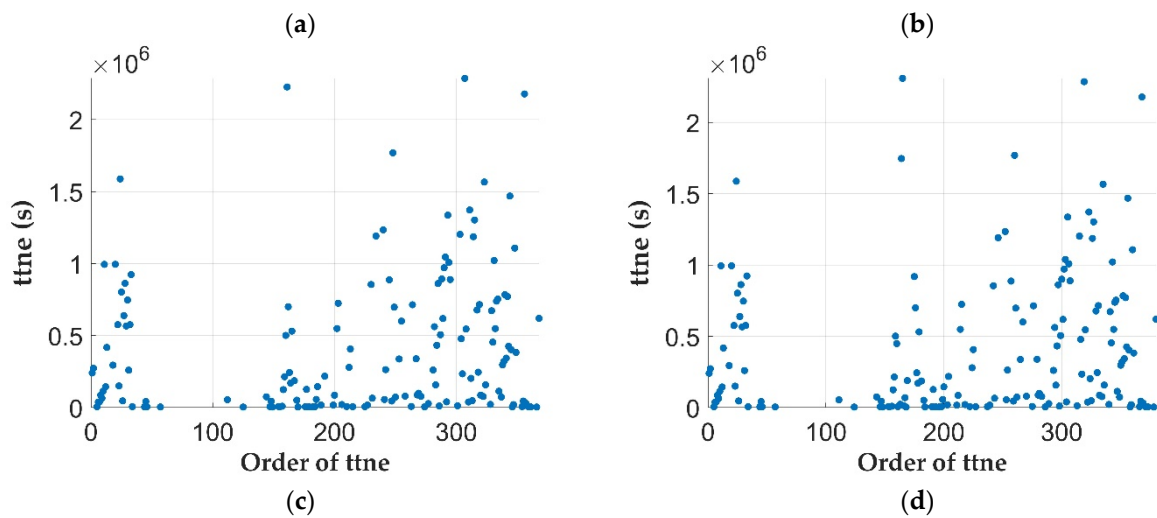


Figure 14. Scatter plot of ttime registered in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d), after the application of the 1h-filter—rare voltage sags.

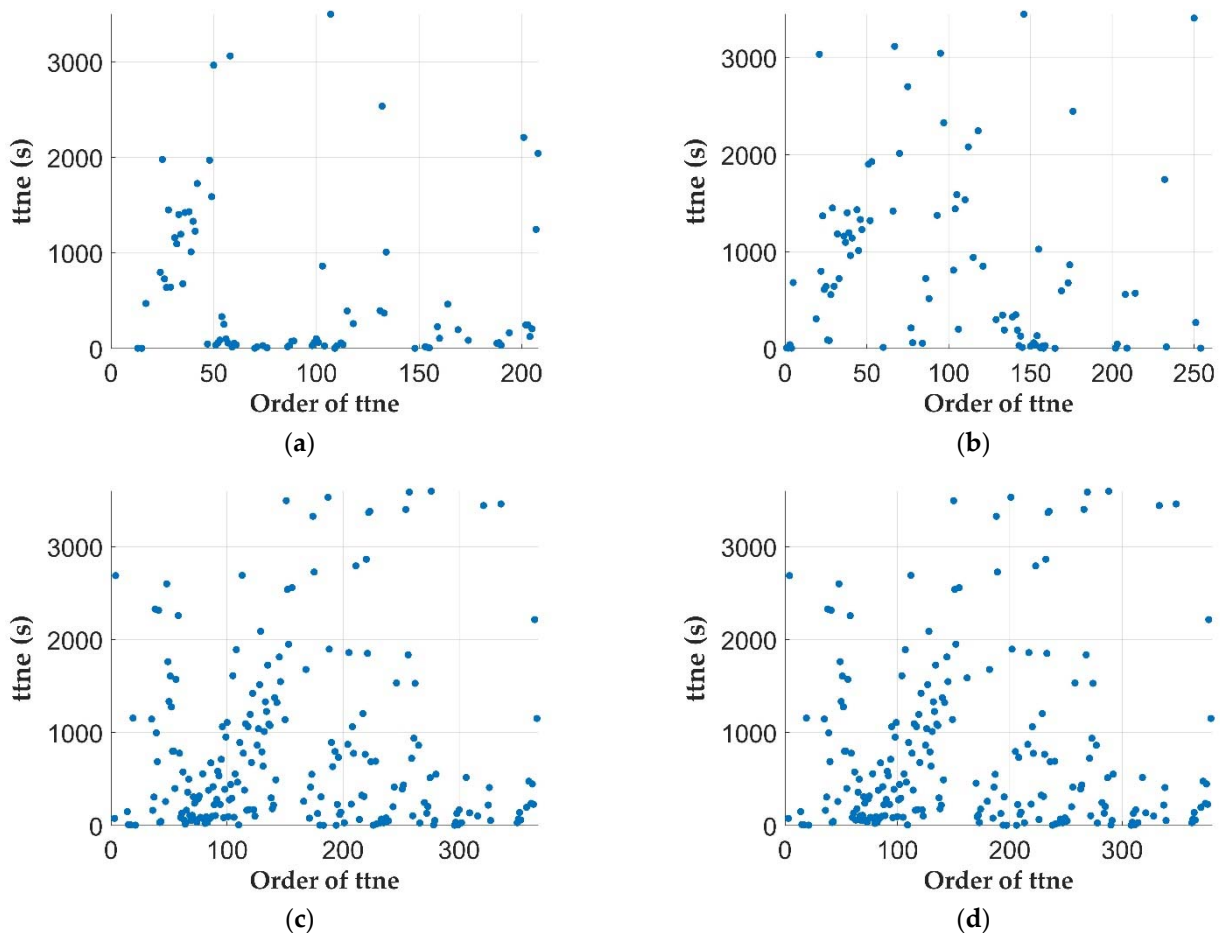


Figure 15. Scatter plot of ttime registered in 2014, 2015, and 2016 at the sites #20: (a), #44: (b), #74: (c), #75: (d), after the application of the 1h-filter—clusters.

It also was interesting to analyze the results of the filtering by the 1h-filter on the plane (Amplitude, Duration); Figures 16 and 17 show the clusters and the rare sags, respectively. Both the clusters (Figure 16) and the rare sags (Figure 17) presented values in all of the plane, proving that these two parameters, i.e., amplitude and duration, were not able by themselves to differentiate between clusters and rare voltage sags, at least those

obtained by the 1h-filter. In particular, the shallow and short sags probably originated due to phenomena other than faults in the network, such as switching of loads and/or motors. These phenomena are present in both the clusters (Figure 16) and the rare voltage sags (Figure 17), and consequently they occurred in succession with times that were shorter and longer than one hour.

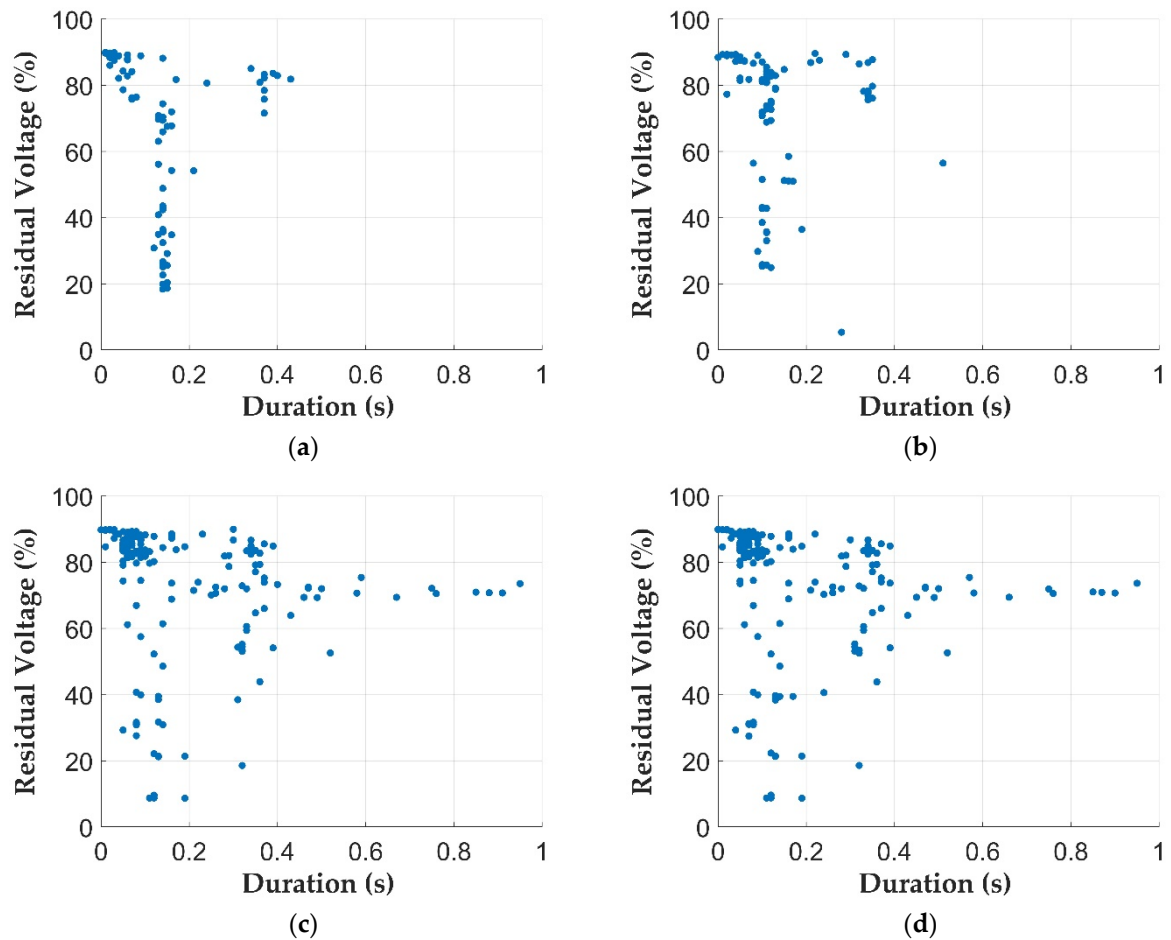
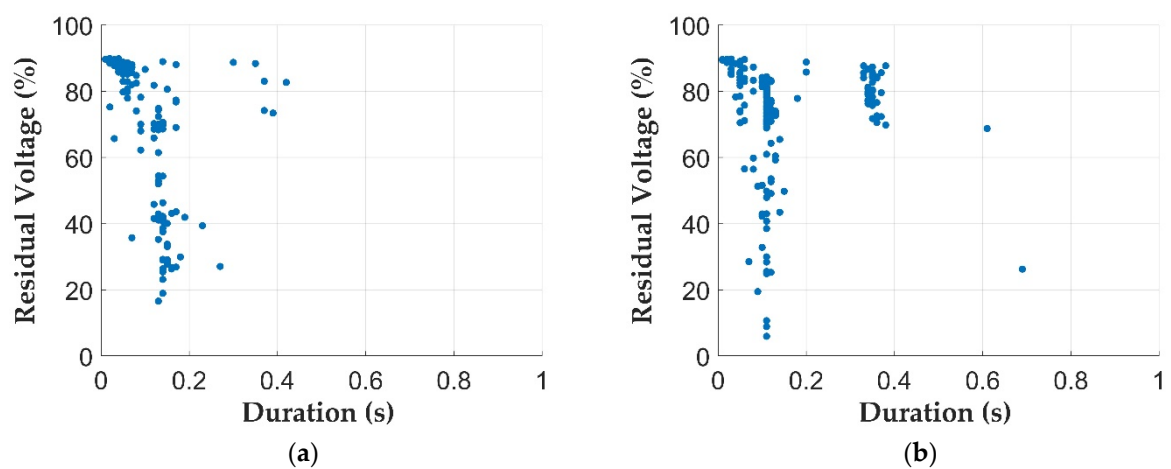


Figure 16. Scatter plot (residual voltage, duration) of the voltage sags registered in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d), after the application of the 1h-filter—clusters.



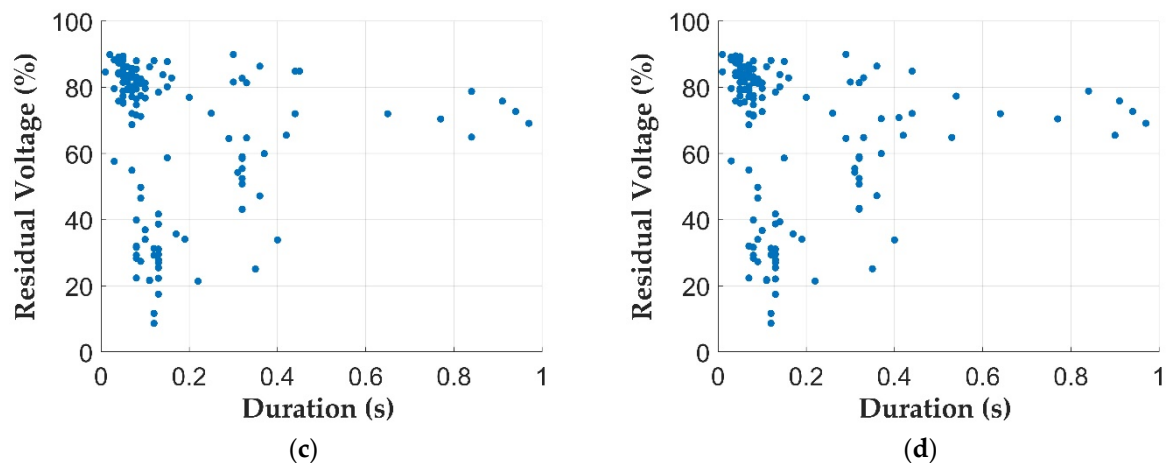


Figure 17. Scatter plot (duration of the residual voltage) of the voltage sags registered in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d), after the application of the 1h-filter—rare voltage sags.

Figure 18 shows the effect of the 1h-filter on the cumulative probability, F , of $ttne$ computed on the voltage sags measured from 2014 through 2016 at sites #20, #44, #74, and #75. Every plot of Figure 18 referred to a single site, and there are two curves, with curve A) referring to all of the sags before the application of the 1h-filter, and curve B) referring to the rare sags obtained after the application of the 1h-filter. It is evident that the initial value of all of the cumulative probabilities was appreciably decreased due to the filtering action. All of the initial values were less than 20%, and this meant that less than 20% of the rare sags were characterized by a minimum value of $ttne$ (the minimum value is not zero; it is at least 2 s for the time aggregation applied and discussed in Section 2).

The values of the 50th percentile, $P_{50\%}$, of the rare sags are presented in Table 5; Table 5 can be compared with Table 4, which referred to all of the voltage sags before the application of the 1h-filter. It is evident from the comparison of these tables that the action of the 1h-filter allowed a significant increase of $P_{50\%}$, with at least one order of magnitude and with minimum values passed from 0.5 h (Table 4 for sites #74 and #75) to 38.4 h (Table 5 for site #44).

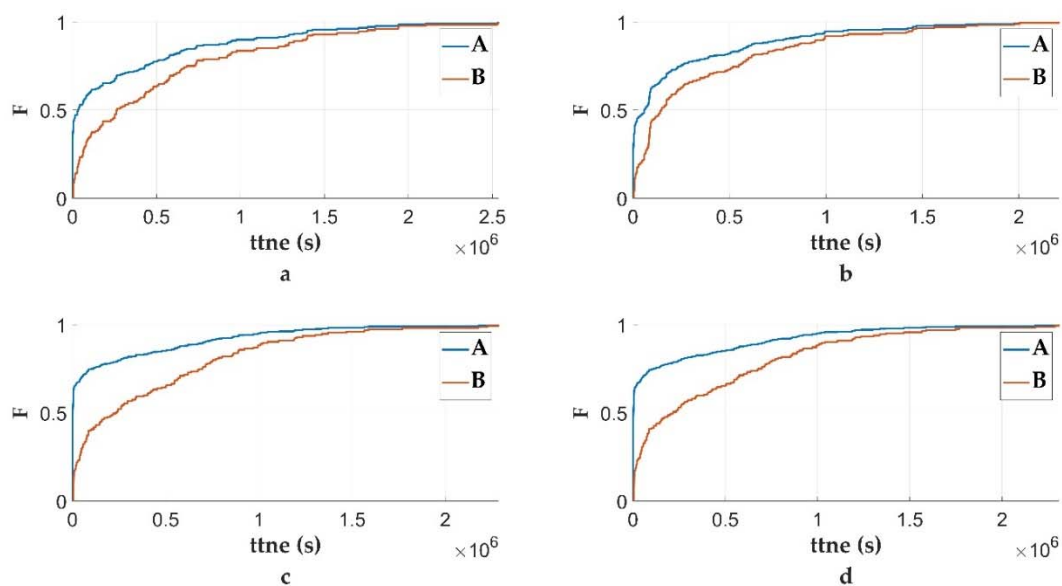


Figure 18. Cumulative distribution function of $ttne$ measured in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d), before the application of the 1h-filter A and after the application of the 1h-filter B.

Table 5. Values of the percentile $P_{50\%}$ of t_{tne} registered at the busbar #20, #44, #74, #75 after 1h-filter.

#Site	$P_{50\%}(s)$	$P_{50\%}(h)$
20	2.8×10^5	77.8
44	1.4×10^5	38.9
74	2.3×10^5	63.9
75	2.3×10^5	63.9

4. Assessing the Assumption about the Exponential Distribution of the Rare Voltage Sags

Once the rare voltage sags were separated from the clusters, as reported in all of the relevant literature on the voltage sags, we assumed that these rare voltage sags could be modelled as a Poisson process. In this case, t_{tne} can be described by the exponential density function, $f_{t_{tne}}$, which is given by:

$$f_{t_{tne}}(t) = \lambda e^{-\lambda t}, \quad (1)$$

where λ is the constant sag rate, i.e., the number of sags in a given time.

Thus, the number n of voltage sags in the monitoring time period was a Poisson random variable characterized by the probability, $P(n)$:

$$P(n = N) = \frac{\lambda^N e^{-\lambda}}{N!}, \text{ with } N \text{ integer number}, \quad (2)$$

where N is the number of events per unit time, λ is the mean value of (2) per time units, and consequently λ^{-1} is the mean value of the time until the next event, i.e., the mean value of the random variable, t_{tne} .

The assumed distribution (1) on the variable t_{tne} must be verified. We used the graphical method of the Q-Q (quantile-quantile) plots [37], as was proposed in [26].

The Q-Q plots are one of the graphical tools used to make distributional comparisons between an empirical distribution derived from a dataset and a theoretical distribution. Q-Q plots of this type also are called theoretical Q-Q plots or probability plots. A theoretical Q-Q plot allows a graphical comparison of the two probability distributions using the values of their quantiles. It consists of plotting the quantiles of the data against the corresponding quantiles of the theoretical distribution. If the two distributions are related linearly, the points in the Q-Q plot will lie approximately on a line with a grade of correlation measurable by the correlation factor, r .

In our case, the scope was to verify whether the distribution of t_{tne} respected the exponential distribution, so we plotted the quantiles of the measured t_{tne} against the quantiles of the theoretical exponential distribution. Since the range of the values of the measured t_{tne} was very large, a cubic-square transformation was used by introducing the auxiliary variable y_i , which was linked to t_{tne_i} by the following relationship:

$$y_i = \sqrt[3]{t_{tne_i}} \quad (3)$$

Figure 19 shows the Q-Q plots obtained for the rare voltage sags measured at sites #20, #44, #74, and #75 from 2014 to 2016; Table 6 reports the main data resulting from the verification.

It was interesting to see from Figure 19 that the distribution of the measured t_{tne} departed from the linearity at the extremes, i.e., the ends of the configuration curved downward on the right and upward on the left. This means that, relative to the middle of the data, the observations in the tails were closer to the center, the median, than they ought to be for a sample from an exponential distribution, so that we judged the measured t_{tne} as having shorter tails than the theoretical exponential distribution. This was very pronounced in sites #74 and #75 for the low values of the quantiles (left tails). This is a further critical aspect again linked to the filtering action. In fact, the thickening of the points in the left tails indicated that some clusters are still present in the data. This condition also was highlighted in [26] with voltage sags measured in different real networks.

To demonstrate the importance of the preliminary filtering, described in the Section 3, Figure 20 and Table 7 show the analogous results obtained by applying the Q-Q plots to the measured voltage sags in the same sites and in the same periods (2014–2016) without filtering the clusters.

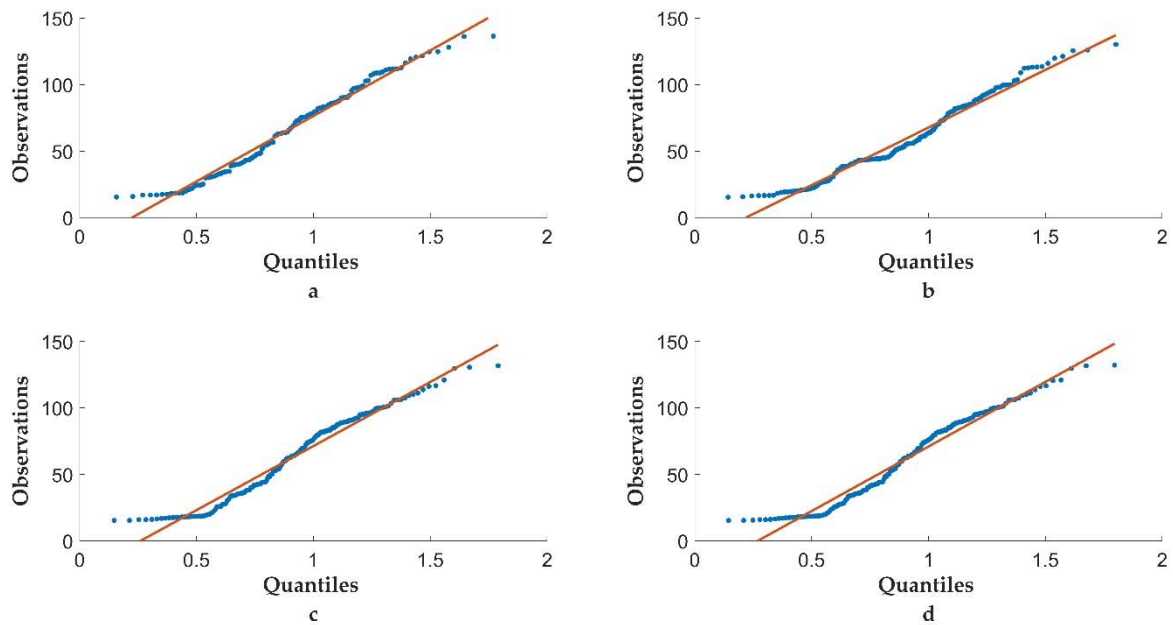


Figure 19. Exponential distribution Q-Q plots with cubic square transformation of the *t*_{tne} measured in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d), after the application of the 1h-filter.

Table 6. Results of the analysis of the Q-Q plots after the application of the 1h-filter.

Site #	Number of Rare Voltage Sags	Correlation Factor <i>r</i>
20	128	0.99
44	173	0.99
74	155	0.98
75	164	0.98

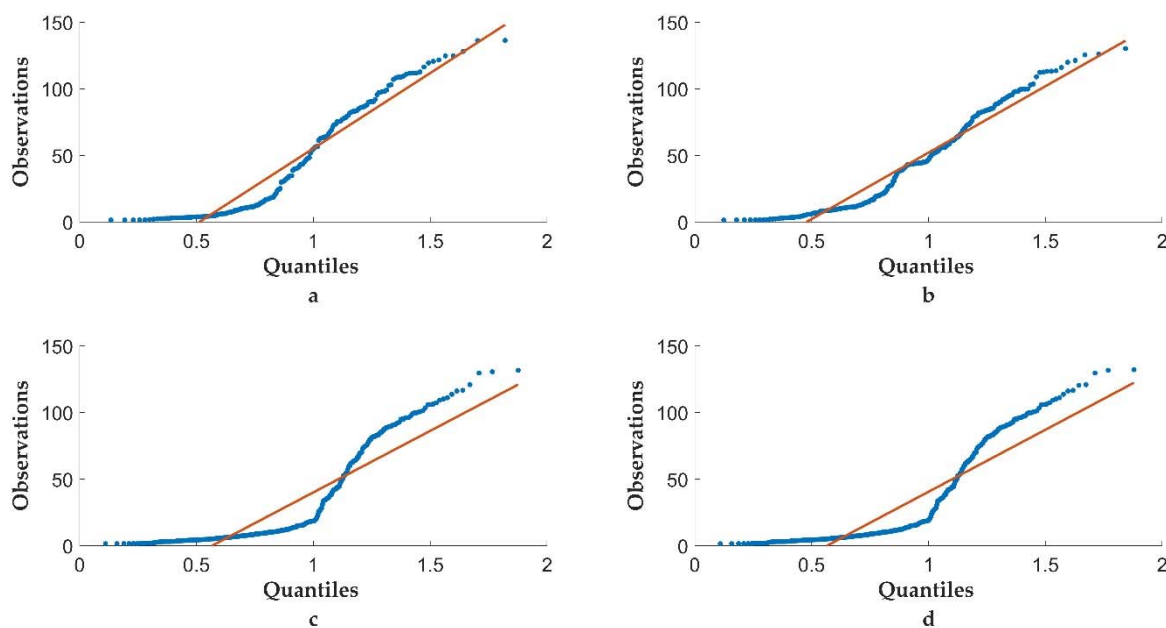


Figure 20. Exponential distribution Q-Q plot with cubic square transformation of the t_{tne} measured in 2014, 2015, and 2016 at sites #20: (a), #44: (b), #74: (c), #75: (d), before the application of the 1h-filter.

Table 7. Results of the analysis of the Q-Q plots before the application of the 1h-filter.

Site #	Number of Total Voltage Sags (Rare Sags and Clusters)	Correlation Factor
		r
20	208	0.96
44	261	0.97
74	368	0.89
75	380	0.89

It is evident that the measured t_{tne} of all of the sags, including the clusters, moved away from the correlation lines, which also caused an appreciable reduction of the correlation factor in Table 7. The rough plateaus at the extremes of all the Q-Q plots in Figure 20 indicate that there were concentrations of measured t_{tne} around the first percentiles, which corresponded to clusters of sags and are not accounted for by the exponential distribution.

5. Forecast of the Average Expected Performance

The final goal was to forecast the mean number of events per year at each site, k , $N_{f,k}$. This value can be derived by correlating the average sag rate, λ_k , with the duration of one year by the relation:

$$N_{f,k} = \lambda_k \times P \quad (4)$$

where λ_k is the average sag rate expressed in number of sags per seconds at the site k and P is the duration of one year expressed in seconds (31.5×10^6 s).

λ_k in Equation (5) is the parameter of the exponential density function $f_{t_{tne}}$ given by relation (1). The Q-Q plots used in the previous section for assessing the exponential distributions of the t_{tne} of the rare sags can be used for estimating λ_k , in case of positive passing of the assessment like in the case of rare sags. The estimate of λ_k is the maximum likelihood estimate from the measured data [30], and it is obtained by measuring the slope m of the straight line resulting from the points in the Q-Q plot:

$$\lambda_k = m^3 \quad (5)$$

The value of m is computed by the least square method. The voltage sags measured in 2017 for the same regional system, $N_{m,k}$, were used to verify the mean number of voltage sags forecasted $N_{f,k}$ by the relation (5) in the preceding years of 2014, 2015, and 2016. The error of the forecast at each site k , $\epsilon_{f,k}$, was computed as the absolute value of the difference between the number of forecasted sags in 2017, $N_{f,k}$, and the number of measured sags in 2017, $N_{m,k}$, in per unit of the measured sags in 2017.

Table 8 shows the results for the considered sites, i.e., #20, #44, #74, #75.

Table 8. Main results of the forecasting activity after the application of the 1 h-filter.

Site #	λ_k^{-1}	$N_{f,k}$	$N_{m,k}$	$\epsilon_{f,k}$ (%)
20	955,316.20	33	32	3.12
44	652,144.06	48	50	4.00
74	895,082.38	35	97	63.92
75	915,980.25	34	98	65.31

For sites #20 and #44, the forecast error achieves satisfactory values, i.e., less than 5%, having forecasted a number of voltage sags in the year 2017 very close to those actually measured (at site #20, 33 voltage sags were forecasted and 32 voltage sags were measured). At site #44, 48 voltage sags were forecasted, and 50 voltage sags were measured. However, for sites #74 and #75, the performance of the forecast was unacceptable. In fact, at site #74, we forecasted 35 sags against 97 measured sags corresponding to a forecast error of 63.92%, and, at site #75, we forecasted 34 sags against 98 measured sags corresponding to a forecast error of 65.31%.

It was evident that the forecast presented a range from acceptable values lower than 5% to unacceptable values greater than 60%.

We started examining in deep detail the sags at sites #74 and #75 to detect the main origins of such a large error. The main aspect that emerged was that the voltage sags at these sites still presented voltage sags as clusters, after the application of the 1h-filter, discussed in the Section 3. The main critical aspect, in other words, was referred to the application of the same filter, that is the 1 h-filter, for all the considered sites.

The first solution to these unacceptable results of the forecast was to observe the clusters in every site. We ascertained that the maximum value of $ttne$, above which the clusters did not increase by more than 3%, was different in every site and was not necessarily equal to 1 h. The maximum values of $ttne$ searched with this criterium for sites #74 and #75 were 4.7 h and 5.3 h, respectively. By using these values for detecting and removing the clusters at these sites, we performed the validation of the rare sags by means of the Q-Q plots shown in Figure 21. Comparing these plots with those in Figures 20c and 20d an appreciable reduction of the number of points leaving the line at the left tail was evident. Further, the correlation factor, r , increased slightly from 0.98 in Table 6 to 0.99 for both of the sites that were considered. This proved that one of the critical aspects, which was discussed in the Section 4, was reduced by using values of the filtering time that were different from 1 h.

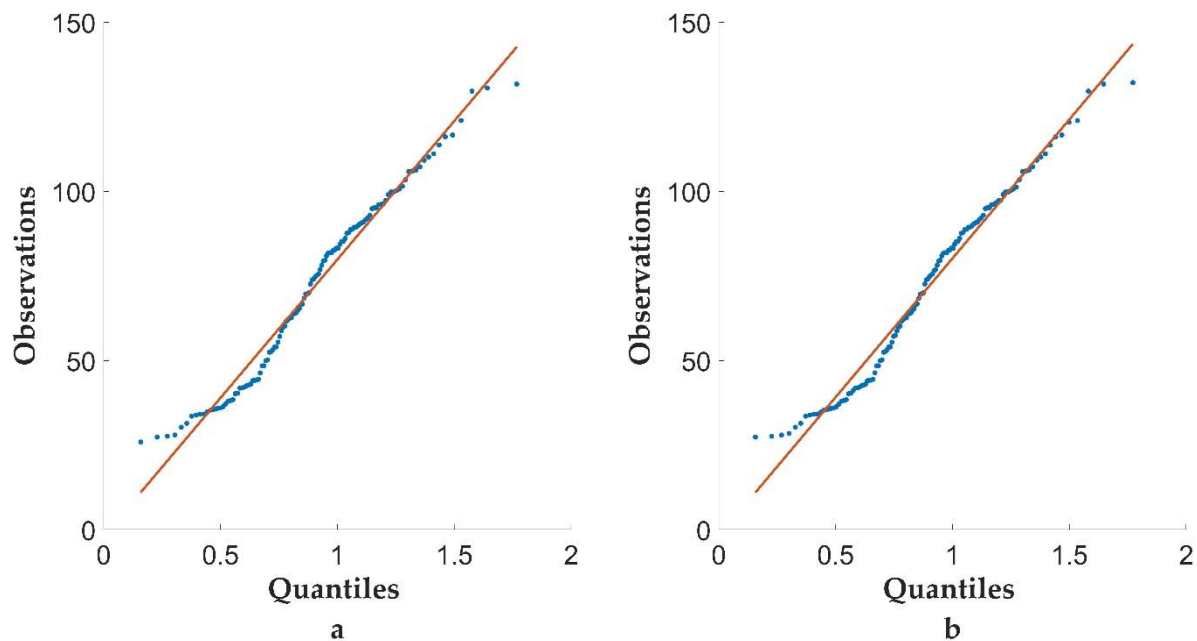


Figure 21. Exponential distribution Q-Q plot with cubic square transformation of the *ttne* measured in 2014, 2015, and 2016 at sites #74: (a), #75: (b) after the application of the filters (4.7 h for the site #74 and 5.3 h for the site #75).

By using the new values of the filters (4.8 h for site #74 and 5.3 h for site #75), we performed the forecast using the new dataset of the rare sags that resulted. We obtained a significant improvement in the forecast, and Table 9 shows the results.

Table 9. Main results of the forecasting activity after the application of the filters (4.7 h for the site #74 and 5.3 h for the site #75).

Site #	λ_k^{-1}	$N_{f,k}$	$N_{m,k}$	$\varepsilon_{f,k}$ (%)
74	5.7×10^5	58	58	0.00
75	5.5×10^5	57	57	0.00

The results shown in Table 9 show the great improvement in the forecast, which was characterized for both site #74 and site #75 by a large reduction of the error, $\varepsilon_{f,k}$, i.e., from 63.92% to 0.00% at site #74 and from 65.31% to 0.00% at site #75.

To prove that the improvements of the forecast were linked to the adequate filtering performed by detecting the *ttne* max at every site, at sites #74 and #75, we derived the cumulative probability distribution, *F*, and the probability density function, pdf, of *ttne* in two ways, the first of which consisted of computing the probability curves, *F* and pdf, on the samples of *ttne* obtained after the new time filters (4.7 h for site #74 and 5.3 h for site #75). The second way consisted of deriving the probability curves *F* and pdf using the average sag rate, i.e., the number of sags per second, λ_k , at sites #74 and #75 estimated by relation (4). Figures 22 and 23 compare the cumulative distributions *F* and the density distribution pdf, respectively, obtained in these two ways. It is evident that the distributions *F* and pdf of the data samples are very close to the ideal distributions that correspond to the estimated λ_k .

As a final remark, it is useful to give some details about the computational burden of the proposed method. The time required for the preliminary statistical process of 33,602 voltage sags measured over a period of three years at 152 sites of the considered networks is about 11 s. Further, 0.5 s is needed for the forecast at each site. The times of the computations refer to the numerical programs developed in MATLAB 2020 by an i7-8750H CPU processor operating @ 2.20 GHz.

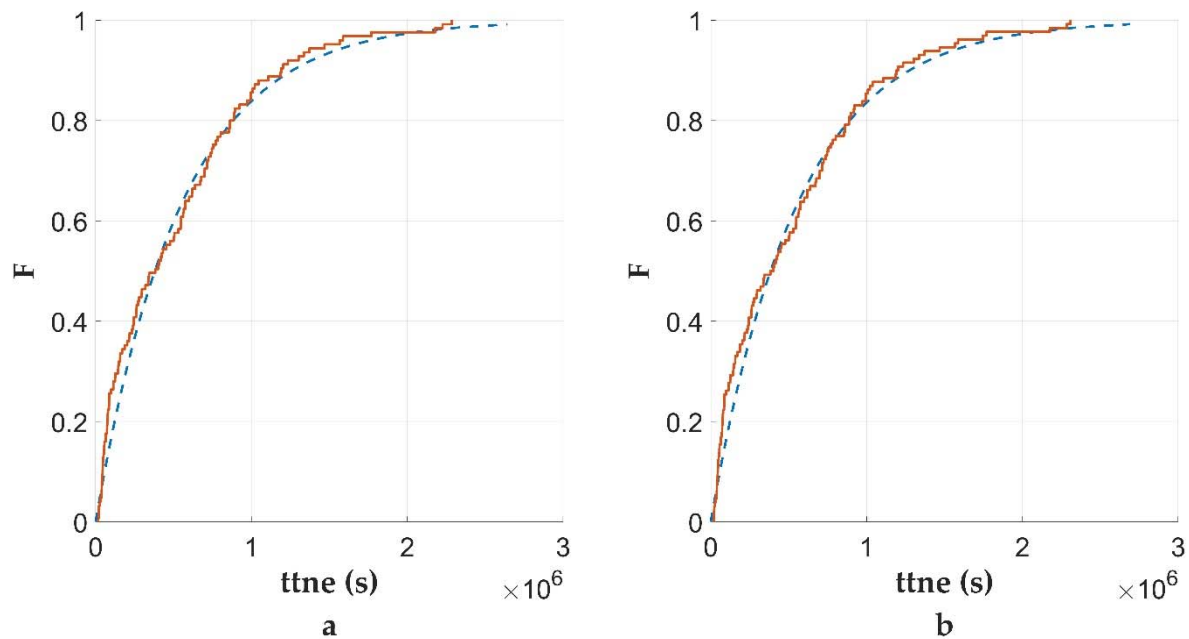


Figure 22. Cumulative distribution function, F , of the $ttne$ measured in 2014, 2015, and 2016 at sites #74: (a), #75: (b) after the application of the of the filters (4.7 h for the site #74 and 5.3 h for the site #75) red curve, and cumulative distribution function with distribution parameter equal to λ_k estimated by relation (4) (blue curve).

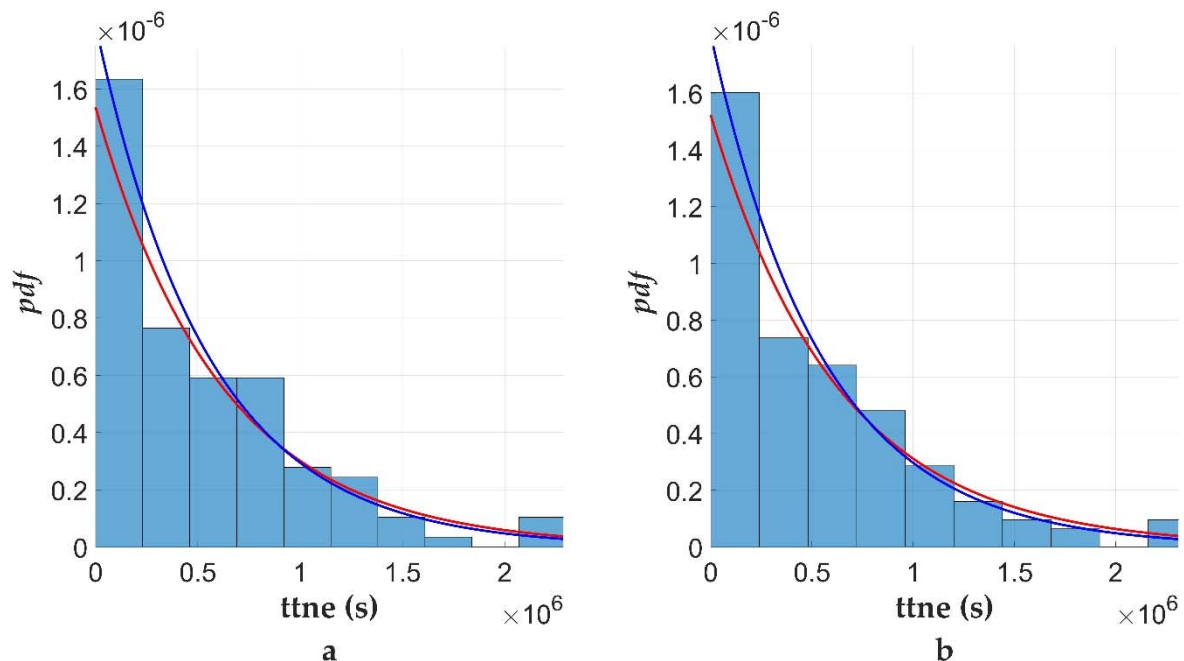


Figure 23. Probability density function fitted of the $ttne$ measured in 2014, 2015, and 2016 at sites #74 (a), #75 (b), before the application of the of the filters (4.7 h for the site #74 and 5.3 h for the site #75)—red curve, and probability density function estimated by relation (1) with distribution parameter equal to λ_k estimated by relation (4)—blue curve.

6. Conclusions

This paper shows the first stage of the studies aimed at forecasting the expected number of sags at the sites of a real distribution system using the voltage sags measured in the field without performing a model-based analysis. The final goal of the research is to develop proper algorithms that the network operators could use to forecast how many voltage sags will occur at a given site. The measured sags were available for four years at

the sites of an actual regional system in Italy, and we used the data of the first three years (2014–2016) to forecast the expected number of sags at every site in the fourth year (2017). Since the actual voltage sags measured in 2017 were available, we were able to verify the forecasts. The forecasts, which used only three years of measurements, were possible with adequate confidence due to the choice of the random variable, *time to next event*, rather than the variable, *number of voltage sags*. In fact, as demonstrated in this paper, the dataset of values of the variable *time to next event* was much larger than that of the variable *number of voltage sags*.

The proposed method requires two main preliminary preparation steps for the data; the first step is to exclude the intermittent sags due to the cycles of the automatic closure of the protection systems; the second step is aimed to filter and discard the sags grouped in clusters. This second step is the most crucial action due to the impact it has on forecasting the performance. In this paper, we used a filter in time, i.e., a 1h-filter, for every site to detect and exclude the clusters of sags, as was proposed in the recent literature. The results obtained in the system we considered were not satisfactory at all of the sites. The most critical evidence was that, when we used a 1h-filter at every site, the clusters were not filtered completely at some sites, and this caused unacceptable forecast errors.

Starting from these initial results, we improved the accuracy of the forecast, measured by the reduction of the forecast error, and detected and discarded the clusters in a dynamic way along the sites. This new filtering action removed all of the clusters from the datasets that were used in the forecasts. The initial results encouraged the development of new, more efficient techniques of filtering, and we acted site by site to eliminate the clusters. The comparison among these techniques will allow us to ascertain the most efficient technique for detecting and discharging the clusters to be used in the entire procedure to forecast the voltage sags that will occur at every site.

Author Contributions: M.D.S., L.D.S., C.N., P.V. (Pietro Varilone) and P.V. (Paola Verde) conceived and designed the theoretical methodology and the numerical applications; M.D.S., L.D.S., P.V. (Pietro Varilone) and P.V. (Paola Verde) performed the numerical applications and analyzed the results; M.D.S., L. D.S., C.N., C.N., P.V. (Pietro Varilone) and P.V. (Paola Verde) wrote and revised the paper. All authors have read and agreed to the published version of the manuscript

Funding: This research was developed inside the grant “Attività di supporto scientifico sulle tematiche di *Power Quality* nelle reti di distribuzione”; in addition, it received the financial support from the Italian Ministry of University and Research (MIUR) through the Special Grant “Dipartimenti di eccellenza” Lex. N. 232/31.12.2016 (G.U. n. 297/21.12.2016 S.O. n. 57).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are not publicly available due to restrictions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. European Committee for Electrotechnical Standardization. *EN50160, Voltage Characteristics of Electricity Supplied by Public Electricity Networks*; CENELEC: Brussels, Belgium, 2016.
2. Bollen, M.H.J. *Understanding Power Quality Problems: Voltage Sags and Interruptions*; IEEE Press: New York, NY, USA, 1999.
3. Carpinelli, G.; Caramia, P.; Di Perna, C.; Varilone, P.; Verde, P. Complete matrix formulation of fault-position method for voltage-dip characterisation. *IET Gener. Transm. Distrib.* **2007**, *1*, 56, doi:10.1049/iet-gtd:20050544.
4. Caramia, P.; Carpinelli, G.; Di Perna, C.; Varilone, P.; Verde, P. Fast Probabilistic Assessment of Voltage Dips in Power Systems. In Proceedings of the 2006 International Conference on Probabilistic Methods Applied to Power Systems, Stockholm, Sweden, 11–15 June 2006; pp. 1–6.
5. Gong, X.; Yu, Y.; Wang, B.; Ye, F.; Zhou, N. A severity assessment method for multiple voltage sag. In Proceedings of the 2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 3–5 October 2017; pp. 135–140.
6. IEEE Guide for Voltage Sag Indices. In *IEEE Std 1564–2014*; IEEE: Piscataway, NJ, USA, 2014; pp. 1–59.

7. Joint Working Group Cigre/Cired C4.110. *Voltage Dip Immunity of Equipment and Installations*; Report 412; Cigré: Paris, France, 2010.
8. Sadigh, A.K.; Smedley, K.M. Review of voltage compensation methods in dynamic voltage restorer (DVR). In Proceedings of the 2012 IEEE Power and Energy Society General Meeting, San Diego, CA, USA, 22–26 July 2012, pp. 1–8.
9. Mottola, F.; Proto, D.; Varilone, P.; Verde, P. Planning of Distributed Energy Storage Systems in μ Grids Accounting for Voltage Dips. *Energies* **2020**, *13*, 401, doi:10.3390/en13020401.
10. Vilathgamuwa, D.; Perera, A.; Choi, S. Voltage sag compensation with energy optimized dynamic voltage restorer. *IEEE Trans. Power Deliv.* **2003**, *18*, 928–936, doi:10.1109/tpwrd.2003.813880.
11. Milanović, J.V.; Gupta, C.P. Probabilistic Assessment of Financial Losses due to Interruptions and Voltage Sags—Part I: The Methodology. *IEEE Trans. Power Deliv.* **2006**, *21*, 918–924, doi:10.1109/tpwrd.2006.870988.
12. Milanovic, J.V.; Gupta, C.P. Probabilistic Assessment of Financial Losses due to Interruptions and Voltage Sags—Part II: Practical Implementation. *IEEE Trans. Power Deliv.* **2006**, *21*, 925–932, doi:10.1109/tpwrd.2006.870987.
13. Didden, M.; Belmans, R.; D’Haeseleer, W. Cost-benefit analysis of voltage sag mitigation methods in fiber extrusion plants. *Eur. Trans. Electr. Power* **2003**, *13*, 73–77, doi:10.1002/etep.4450130202.
14. Di Fazio, A.; Duraccio, V.; Varilone, P.; Verde, P. Voltage sags in the automotive industry: Analysis and solutions. *Electr. Power Syst. Res.* **2014**, *110*, 25–30, doi:10.1016/j.epsr.2014.01.004.
15. Dos Santos, A.; De Barros, M.C. Voltage sag prediction for network planning. *Electr. Power Syst. Res.* **2016**, *140*, 976–983, doi:10.1016/j.epsr.2016.03.033.
16. Olguin, G.; Aedo, M.; Arias, M.; Ortiz, A. A Monte Carlo Simulation Approach to the Method of Fault Positions for Stochastic Assessment of Voltage Dips (Sags). In Proceedings of the 2005 IEEE/PES Transmission & Distribution Conference & Exposition: Asia and Pacific, Dalian, China, 18 August 2005; pp. 1–6.
17. Caramia, P.; Varilone, P.; Verde, P.; Vitale, L. Tools for Assessing the Robustness of Electrical System against Voltage Dips in terms of Amplitude, Duration and Frequency. *Renew. Energy Power Qual. J.* **2014**, *1*, 177–182, doi:10.24084/repqj12.275.
18. Lim, Y.; Strbac, G. Analytical approach to probabilistic prediction of voltage sags on transmission networks. *IEEE Proc. Gener. Transm. Distrib.* **2002**, *149*, 7–14, doi:10.1049/ip-gtd:20020008.
19. Espinosa Juarez, E.; Hernandez, A. An Analytical Approach for Stochastic Assessment of Balanced and Unbalanced Voltage Sags in Large Systems. *IEEE Trans. Power Deliv.* **2006**, *21*, 1493–1500, doi:10.1109/tpwrd.2005.860266.
20. Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on Common Rules for the Internal Market for Electricity and Amending Directive 2012/27/EU; Parlamento Europeo e il Consiglio dell’Unione Europea: Brussels, Belgium, 2019, Available online: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L0944> (accessed on 22 February 2021).
21. *Incentives Schemes for Regulating Distribution System Operators, Including for Innovation*; CEER: Brussels, Belgium, 2018. Available online: <https://www.ceer.eu/documents/104400/-/-/f04f3e11-6a20-ff42-7536-f8afd4c06ba4> (accessed on 22 February 2021).
22. AEEGSI. Resolution 29th December 2011, 198/2011. Available online: www.autorita.energia.it (accessed on 22 February 2021).
23. ARERA. Resolution 23rd December 2019, 566/2019/R/EEL. Available online: <https://www.arera.it/it/docs/19/566-19.html> (accessed on 22 February 2021).
24. CEI EN 61000-4-30: 2015. COR 2017. Electromagnetic Compatibility (EMC)-Part 4-30: Testing and Measurement Techniques-Power Quality Measurement Methods (IEC 61000-4-30:2015/COR1:2016). European Committee for Standards-Electrical.
25. Noce, C.; Di Stasio, L.; Varilone, P.; Verde, P.; De Santis, M. On the Forecast of the Voltage Sags: First Stages of Analysis on Real Systems. In Proceedings of the 2020 55th International Universities Power Engineering Conference (UPEC), Turin, Italy, 1–4 September 2020; pp. 1–6.
26. Dos Santos, A.; Rosa, T.; De Barros, M.T.C. Stochastic Characterization of Voltage Sag Occurrence Based on Field Data. *IEEE Trans. Power Deliv.* **2018**, *34*, 496–504, doi:10.1109/tpwrd.2018.2878997.
27. Caramia, P.; Di Mambro, E.; Varilone, P.; Verde, P. Impact of Distributed Generation on the Voltage Sag Performance of Transmission Systems. *Energies* **2017**, *10*, 959, doi:10.3390/en10070959.
28. Chiumeo, R.; De Nigris, M.; Garbero, L.; Gandolfi, C.; Tenti, L.; Carpaneto, E. Implementation of a New Method for an Improved Voltage Dips Evaluation by the Italian Power Quality Monitoring System in Presence of VT Saturation Effects. *Renew. Energy Power Qual. J.* **2010**, *1*, 519–522, doi:10.24084/repqj08.383.
29. De Santis, M.; Noce, C.; Varilone, P.; Verde, P. Analysis of the origin of measured voltage sags in interconnected networks. *Electr. Power Syst. Res.* **2018**, *154*, 391–400, doi:10.1016/j.epsr.2017.09.008.
30. Noce, C.; De Santis, M.; Di Stasio, L.; Varilone, P.; Verde, P. Detecting the Origin of the Voltage Sags Measured in the Smart Grids. In Proceedings of the 2019 International Conference on Clean Electrical Power (ICCEP), Otranto, Italy, 2–4 July 2019; pp. 129–135.
31. Liu, H.; Zeng, Q.; Liu, S.; Wu, Z.; Li, L.; Ma, R. Voltage sag analysis based on cluster analysis and correlation analysis. In IOP Conference Series: Materials Science and Engineering, In Proceedings of the 2020 The 6th International Conference on Electrical Engineering, Control and Robotics, Xiamen, China, 10–12 January 2020; IOP Publishing: Bristol UK, 2020; Volume 853, p. 12018.
32. IEC 61000-4-11:2020, Electromagnetic compatibility (EMC)-Part 4-11: Testing and measurement techniques-Voltage dips, short

- interruptions and voltage variations immunity tests for equipment with input current up to 16 A per phase. European Committee for Standards–Electrical. Available online: <https://webstore.iec.ch/publication/63503#additionalinfo> (accessed on 22 February 2021).
33. Di Mambro, E.; Caramia, P.; Varilone, P.; Verde, P. Voltage sag estimation of real transmission systems for faults along the lines. In Proceedings of the 2018 18th International Conference on Harmonics and Quality of Power (ICHQP), Ljubljana, Slovenia, 13–16 May 2018; pp. 1–6.
 34. Qader, M.; Bollen, M.; Allan, R. Stochastic prediction of voltage sags in a large transmission system. *IEEE Trans. Ind. Appl.* **1999**, *35*, 152–162, doi:10.1109/28.740859.
 35. Bollen, M.H.J.; Gu, I.Y.-H. *Signal Processing of Power Quality Disturbances*; Wiley: Hoboken, NJ, USA, 2006.
 36. Papoulis, A. *Probability, Random Variables, and Stochastic Processes*; McGraw-Hill: Japan, Tokyo, 1965.
 37. Chambers, J.M.; Cleveland, W.S.; Kleiner, B.; Tukey, P.A. *Graphical Methods for Data Analysis*; CRC Press: Boca Raton, FL, USA, 2018.