MODELLISTICA E CALCOLO SCIENTIFICO

MODELING AND SCIENTIFIC COMPUTING

MOX–Report No. 6/2008

Statistical analyses of exceptional events: the Italian experiences

Elena Fumagalli, Luca Lo Schiavo,
Anna Maria Paganoni, Piercesare Secchi

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 29 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

Statistical analyses of exceptional events: the Italian experience

Elena Fumagalli[a], Luca Lo Schiavo[b], Anna Maria Paganoni[c], and Piercesare Secchi[c]

22nd February 2008

[a] Department of Economics and Industrial Engineering
Politecnico di Milano, Milano, Italy
[b] Italian Regulatory Authority for Electricity and Gas,
Quality and Consumers Affairs, Milan, Italy
[c] MOX– Modellistica e Calcolo Scientifico
Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy
elena.fumagalli@polimi.it
lloschiavo@autorita.energia.it
anna.paganoni@polimi.it
piercesare.secchi@polimi.it

Keywords: Continuity of supply, exceptional events, performance-based regulation, reliability, statistical methodology.

Abstract

In the analysis of reliability performance of distribution utilities as well as in continuity of supply regulation, criteria are needed for separating normal operation data from exceptional events. In recent years a number of statistical methodologies has been proposed for this purpose. We present here the new methodology that was adopted by the Italian regulatory authority at the beginning of 2008. The decision is supported by a statistical analysis of the number of faults on the MV and on the LV networks, for each 6-hour time interval in a three year time span, for different provinces and distribution companies. The new methodology is employed in the reward and penalty mechanisms that regulate the SAIDI, SAIFI and MAIFI indicators and, with some original provisions, also in the Guaranteed Standard on maximum restoration times.

1 Introduction

Reliability performance of distribution utilities has received considerable attention in recent years. The introduction of continuity of supply regulation in several European countries and an increased awareness by the customers, are the key factors in this process. Analyses of continuity of supply indicators are fundamental for setting regulatory targets, monitoring utility performance, and disseminating information to the public [1, 2].

One of the main problems in these analyses is how to identify events that are exceptional with respect to normal performance. The exclusion of these extreme cases from the data set enable utilities, regulators and the public to observe more meaningful trends in 'normal operation' performance, that would be, otherwise, hard to capture. In addition, regulatory instruments employed in quality regulation usually penalise and/or reward utilities on the basis of expected performance. It is therefore crucial to understand clearly when failure in meeting regulatory targets is due to the utility behavior or to events that are outside the utility's control. Moreover, even if some events, such as extreme weather conditions, are unavoidable, regulators have become more and more interested in controlling the efficiency and effectiveness of utility restoration schemes under such conditions.

Traditional criteria for separating continuity of supply data into normal operation data and exceptional data are based on definitions of exceptional events, given in terms of number of customers interrupted, duration of the interruption, weather conditions, extent of the mechanical damage to the distribution system, and combinations of these factors. Criteria of this sort, however, are not always sufficiently unambiguous or objective in the implementation phase [3, 4]. For similar reasons they have been criticized also in the literature [5, 6].

A few recent contributions attempted to overcome these difficulties, with the use of statistical methodologies [7, 6, 4, 8]. A statistical approach, in fact, is expected to present significant advantages because of a reduction in ambiguities and an increase in fairness. Nevertheless, statistical analyses of exceptional events can be performed in very different ways, depending on the choice (often the availability) of the quality indicator, on the spatial and temporal units of such measure, and on the statistical methodology employed. In addition, the choice of the threshold that will separate normal from exceptional events allows for a fair amount of discretion.

In this paper, after reviewing the existing statistical approaches (Sections 2 and 3), we present the new methodology that was adopted by the Italian regulatory authority at the beginning of 2008. Section 4 describes the statistical basis for this decision and Section 5 discusses the regulatory aspects. Section 6 concludes and indicates directions for further work.

2 Statistical methodologies: US and UK approaches

2.1 Beta Method

After careful consideration of several alternatives, the IEEE Group on System Design[1] created, in 2004, a statistical methodology called the Beta Method, that allows segmentation of reliability data into normal and exceptional categories [6]. The main purpose of the methodology is to enable utilities and regulators to study reliability performance that is observed during normal operating periods (i.e. to identify trends in this performance). To this end, the large statistical effects of major events need to be separated from normal operation data. Other desirable properties of the methodology are: fairness to all utilities, regardless of their size, consistency for different volumes of available data, and simplicity in application [9].

The Beta Method for identifying abnormal events is based on the idea of Major Event Days (MEDs). An MED is a day in which the daily SAIDI (System Average Interruption Duration Index) exceeds a threshold value, referred to as T_{MED}. For the calculation of T_{MED}, the methodology recommends to use five years of daily SAIDI data (excluding days with SAIDI equals to zero) and to take the natural logarithm (*ln*) of each value in the data set. The threshold is then calculated using the equation:

$$T_{MED} = e^{(\alpha + 2.5\beta)}$$

where
α is the average of the *ln* of the daily SAIDI;
β is the standard deviation of the *ln* of the daily SAIDI.

Any day with a daily SAIDI greater than the threshold value T_{MED} is designated as an MED. Data for these days should be removed when calculating normal reliability performance.

Note that the Beta Method assumes that the five-year distribution of daily SAIDI, for any utility, is log-normal. If this is true, the 2.5 multiplier finds an average of 2.3 MEDs per year. The method and the multiplier value were evaluated by a number of utilities and "found to correlate reasonably well to current major event identification results for those utilities" [IEEE, 2004, page 26]. However, the actual data of an utility are not expected to conform precisely to the log-normal distribution. In practice, it is not uncommon to find higher or lower numbers of MEDs per year [6].

[1] Part of the Distribution Subcommittee that reports to the IEEE Power Engineering Society (PES) Transmission and Distribution Technical Committee.

2.2 UK approach

A different motivation lead Ofgem, the UK regulatory authority, to examine the nature of exceptional events. Ofgem conducted this study within the work on developing penalty payment arrangements under the Guaranteed Standard on supply restoration. This GS requires a Distribution Network Operator (DNO) to compensate consumers for supply interruptions longer than a fixed number of hours, also in case of severe weather conditions [10].

In order to separate normal from severe weather conditions (and to classify the latter in a scale of severity) Ofgem looked both at the daily number of incidents on the distribution network and at the daily Customer Minutes Lost (CML). The number of incidents was found a more indicative measure of performance: historical CML figures are subject to additional volatility when compared to incident numbers. Balancing between statistical robustness, simplicity and acceptability, the threshold was set at [7]:

$$\textit{Average daily number of incidents at higher}^2\ \textit{voltages} \cdot 7$$

This threshold, later modified to *Average* $\cdot$ *8* uses a simple multiplication factor applied to the average daily figure [11]. For regulatory purposes, the average for each DNO is calculated using 5 years of daily data. Although there might be issues regarding its statistical robustness, this approach has the advantage of being simple to administer and calculate.

For the regulatory period 2005-2011, the GS on maximum restoration time is split in two standards, one covering *normal* weather conditions (≤ 8 times the daily average number of faults) and one covering *severe* weather conditions [10]. Under normal weather conditions a compensation is paid after 18 hours of interruptions. A scale of severity differentiates maximum restoration times in case of:

- Medium weather events (> 8 times and ≤ 13 times the daily average number of faults): *24 hours*;
- Large weather events (≥ 13 times the daily average number of faults): *48 hours*;
- Very large events (any severe event where $\geq 35\%$ of exposed[3] customers are affected - but less than 60%): $48\,hours \cdot \left(\frac{Number\,of\,customers\,affected}{35\%\,of\,exposed\,customers}\right)^2$.

The regulation includes specifications that are important from the perspective of practical implementation of the GS. First, a DNO's financial exposure to penalty

[2]Higher voltage means any nominal voltage of more than 1 kV up to and including 132 kV [10].

[3]Customers supplied by mixed or overhead voltage circuits (i.e. customers that may be affected by a severe weather event).

payments is capped at 2% of the price control revenue per annum. Any payments above this cap is passed through to customers. Second, under all weather conditions, Ofgem allows DNOs to consider a delay in the clock starting to count towards the trigger period for compensation if snow, flooding or ice directly prevent the utility from carrying out the work necessary to restore supply. Finally, a number of 'non-weather' exemptions apply. They include failures of the distribution system of another distributor, faults of the transmission system, cases when the government invokes emergency powers, and so on.

3 The Italian experience (2000-2007)

The Italian regulatory authority (AEEG) introduced, in the year 2000, a reward and penalty mechanism linking the electricity distribution tariff to the SAIDI indicator, for long, unplanned interruptions. This indicator is measured separately in more than 300 territorial districts, covering the entire national territory. A district is served by a single distribution company (in most cases, it represents a small part of the distribution territory), it is homogeneous in population density (districts are classified in high, medium and low density), and it is contained within the administrative boundaries of a single province. Financial incentives are calculated per district on an annual basis, as a function of the difference between a target-SAIDI and the actual-SAIDI.[4] The national distribution tariff, p_t in year t changes according to a modified price cap formula:

$$p_t = p_{t-1}(1 - RPI - X \pm Q)$$

where RPI is the retail price index, X is the efficiency gain and Q is the quality adjustment.

3.1 First regulatory period

In the first regulatory period (2000-2003), the regulation required companies to classify interruptions according to three categories: (i) *Force Majeure*, (ii) external causes, (iii) utility responsibility.[5] The actual-SAIDI used to calculate rewards and penalties did not include contributions from interruptions belonging to the first two categories.

AEEG accepted a *Force Majeure* attribution only if the exceptional nature of the event could be proven by technical or administrative evidence. For instance, a formal declaration of calamity made by the government or measures

[4]In 2008 the continuity indicators SAIFI and MAIFI were added to the incentive mechanism (see Section 5).

[5]*Force Majeure* included public authority (police, firemen) interventions, exceptional natural events leading to either a natural calamity declaration or to climatic conditions beyond the technical design parameters of the grid, and strikes. External causes included third party responsibilities (for instance, digging activities) and interruptions originated on the transmission grid (or on other interconnected systems).

of wind speed made by an independent weather center. In practical terms, this procedure turned out to be rather burdensome both for the companies, that were collecting the data, and for the regulatory authority, that was controlling the documentation provided. In addition, a few controversial cases, where the exceptional nature of the event was claimed by the companies, but could not be formally proven, generated a large amount of disputes [4].

3.2 Second regulatory period

In 2003, AEEG began to study a different procedure, that would identify exceptional events on the basis of a statistical methodology. Empirical evidence provided the idea that the daily CAIDI (Customer Average Interruption Duration Index) was a good indicator of difficult operational conditions and, hence, of a Major Event Day (one where longer-than-average restoration times were measured). According to this idea, AEEG introduced a Two-Step statistical method that was consistent with the territorial and temporal units used in continuity regulation: respectively, the district and the year. This model-free, non-parametric, approach aimed at identifying an extreme region in the SAIFI-SAIDI plane where MEDs belong to. For the sake of simplicity, the region boundaries were defined using thresholds defined by functions of the mean and of the standard deviation.

In particular, for each district and year the First Step extracts, from the yearly distribution of daily CAIDIs, *potential* MEDs (days with a large daily CAIDI - restoration time). The Second Step identifies, among potential MEDs, days with a large SAIDI value (called *computed* MEDs). In case no computed MEDs are found for the district, the day with the largest daily SAIDI, among potential MEDs, is classified as an *assigned* MED [4]. The actual-SAIDI used in the reward and penalty mechanisms did not include the minutes lost contributed by computed or assigned MEDs.

The Two Step method was employed, on a voluntary base, in the regulatory period 2004-2007. Most of the distribution utilities adopted it and the methodology performed as expected in the large majority of cases. A significant reduction in administrative work resulted, on both the regulator's and the utilities' side.

Nonetheless, a few drawbacks emerged. Consumers did not find satisfactory that the identification of MEDs was possible only at the end of the year. They also found confusing that the same event, when affecting different districts within a small geographical proximity, might result in exclusion of minutes lost in some, but not in others districts. Moreover, empirical evidence showed that the methodology was somehow rigid in the case of events spreading across two days. Finally, and most importantly, for a small number of districts the methodology was found unable to identify events that the companies would have classified as exceptional, on the basis of their knowledge and experience. A closer inspection of the data showed that these events did not pass the first of the two thresholds. The adequacy of the first threshold to extract all the potential MEDs was thus studied in detail.

The analysis revealed that for some districts and years the methodology found extremely large first thresholds [8]. These anomalous values masked potential MEDs, making the methodology less reliable as well as less equitable across different districts (the requirement for a potential *MED* was more stringent in some cases). In addition, limiting the number of days that moved to the Second step, large first thresholds created a prerequisite for increasing the number of assigned MEDs.

In turn, it was observed that, because percentiles are not influenced by extreme values as much as the mean and the standard deviation, they might be a more suitable tool for the purpose of identifying potential MEDs. Therefore, a Reviewed First step was proposed, that identified potential MEDs using the 75th percentile of the yearly distribution of daily CAIDIs [8].

Numerical analyses confirmed that using percentiles incorrect exclusions of potential MEDs could be avoided, as well as the need to administratively assign one MED per year. The study needed to be carried further in order to indicate the most adequate percentile to use in the regulation. However, during the consultation process for the third regulatory period additional elements emerged that lead to a new statistical analysis of exceptional events.

3.3 Towards the third regulatory period

At the beginning of 2007, a renovated interest emerged for exceptional events.

For eight years continuity of supply regulation in Italy had revolved mainly around the reward and penalty mechanism described above, with the key objective to improve districts average performances. For the third regulatory period, AEEG added to the scope of continuity regulation two specific objectives: (i) the protection of customers in case of very long and widespread interruptions, including those caused by exceptional events, (ii) the introduction of incentives for utilities to ensure prompt supply restoration under all circumstances (within the boundaries of ensuring safe working conditions for their personnel). To achieve these objectives, Guaranteed Standards on maximum restoration times were proposed, with the idea to differentiate penalty payments for normal and exceptional events.

In addition, an important refinement in the analysis of exceptional events became possible because of the availability of data on the number of faults on Medium Voltage (MV) and Low Voltage (LV) networks.[6] The number of faults is a good, technical indicator of network performance and one closely linked to the physical operation of the grid. These data had been included, since 2004, in the interruption registers kept by regulated utilities. Three years of data were thus available (2004-2005-2006).

Finally, the limitations of the previous model suggested a search for more appropriate spatial and temporal units for the identification of exceptional events.

[6]Medium voltage: between 1kV and 35 kV. Low voltage: below 1 kV.

As for the spatial unit, the proposal was to explore a geographical area larger than the district. The Italian distribution sector includes one very large utility, Enel Distribuzione (the ex-monopolist), that serves more than 80% of the consumers and a number of local companies. As for the latter, the analysis was carried out per company. As for Enel Distribuzione, a separation of the distribution territory in smaller units was necessary. Telecontrol centers are closely related to the technical structure of the network. Therefore, they appeared as the preferable choice. However, it was observed that they might be modified over time by the company to include different groups of consumers. For this reason, the company (for local utilities) and the province (for Enel Distribuzione) were preferred as the new spatial units of the analysis.[7] Note that local utilities have one telecontrol center each and their distribution territory is always confined within a single province.

Regarding the temporal unit, a significant drawback was identified in the above mentioned difficulty that was experienced with events spreading across two days.[8] Temperatures might fall considerably during the night in continental climates. This creates the conditions for observing most severe weather events between nightfall and sunset. The analysis of fixed 24-hour temporal units (from 0 to 24), because it divided the event between two days, was thus underestimating their impact on consumers. Several different temporal units, all smaller than the day, were taken into consideration. The preferred unit was then identified in intervals of 6 hours. Moreover, as specified in Section 5, the criterion employed in the regulatory decision entails an even greater flexibility, in order to account for the time during which the event rises and then descents.

In summary, in this new framework, for a given distribution utility, province and voltage level, a 6-hour interval was deemed to be exceptional when the number of faults registered in that interval was 'large' with respect to the number of faults normally registered. The detailed procedure for the identification of these Exceptional Intervals (EI) is described in the following Section.

4 Identification of Exceptional Intervals

The identification of EI moves from the statistical exploration of the distribution companies records of the number of electrical service faults in the three year time period 2004-2005-2006. Every day of the year has been divided in time intervals of 6 hours (0-6, 6-12, 12-18, 18-24). The data set available to AEEG consists of the observations, for each 6-hour time interval in the three year time span, of the number X of faults on the Medium-Voltage (MV) and on the Low-Voltage (LV) networks, for different provinces and distribution companies. The data set

[7]These units are called 'province and company combinations' in Section 4

[8]A similar problem was identified in [6]. In calculating the daily SAIDI, the Beta Method attributes interruption durations that extend into subsequent days to the day on which the interruption begins.This technique ties the customer-minutes of interruption to the instigating event.

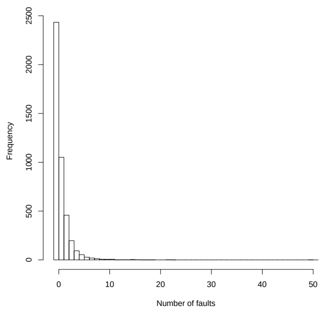

Figure 1: Histogram of number of faults on the MV network for a specific P&C, years 2004-2006.

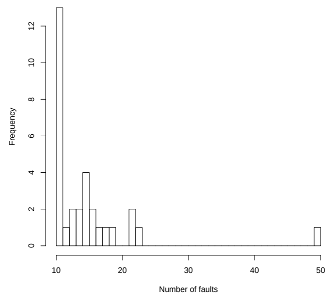

Figure 2: Detail of Fig. 1: 6-hour time periods with a number of faults greater than or equal to 10.

includes a total of 113 province and company combinations (P&C): for each of them the time series of 4384 consecutive observations of X is available both for the MV and the LV network.

The histogram of X on the MV network for a specific P&C is illustrated in Fig. 1. Note that for a large number of 6-hours time intervals, the number of faults is equal to 0, 1, or 2 whereas few time intervals in the three year time span record a large number of faults; this is evidenced in Fig. 2 that focuses on the right tail of the histogram of X. This clustered pattern is quite typical for all P&Cs and suggests that the distribution of X is in fact a mixture of distributions. By definition an Exceptional Interval generates a large number of faults; hence, time intervals with a number of faults belonging to a cluster of small values of X will be called Ordinary Intervals (OI) whereas those with a number of faults in a cluster of large values of X are termed Exceptional Intervals (EI).

Fig. 3 and 4 show the histogram of X on the LV network for the same P&C represented in Fig. 1 and 2: the clustered pattern is still evident.

Although the shape of the distribution of X on the MV and LV network is similar, there is at least one very important difference. The proportion of time intervals with 0 faults is always larger for the MV network: for all P&Cs examined the frequency of 6-hours time intervals with no faults of service is larger, sometimes much larger, for the MV network than for the LV network.

According to this explorative analysis, for every P&C the distribution of X on a given network is a mixture of distributions; in order to determine a threshold separating the values of X for OIs from those relative to EIs, i.e. the specific P&C *exceptionality threshold* on the given network, we proceed by estimating the member of the mixture generating the values of X for ordinary intervals. The exceptionality threshold is then fixed as the percentile $q_{1-\alpha}$ of this distribution, with α very small. Indeed, a value of X larger than $q_{1-\alpha}$ is very rarely generated by the distribution of the number of faults for an ordinary interval; therefore,

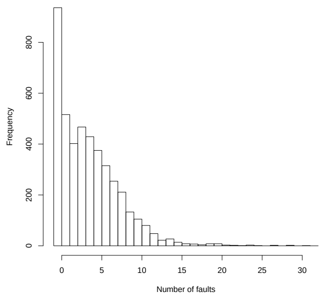

Figure 3: Histogram of number of faults on the LV network for a specific P&C, years 2004-2006.

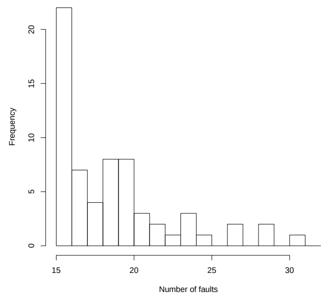

Figure 4: Detail of Fig. 3: 6-hour time periods with a number of faults greater than or equal to 15.

an observed value of X larger than $q_{1-\alpha}$ is more likely generated by an EI than by an OI.

The procedure for the elicitation of the exceptionality threshold is detailed in the next four steps, where model assumptions and estimation procedures are also specified. Each P&C is labeled with a number j, $j = 1, 2, ..., 113$, and X_j denotes the number of faults relative to the P&C labeled j.

i) *Identification of the OI-cluster.*

For $j = 1, 2, ..., 113$, the observed values of X_j are clustered with a k-means algorithm [12]. In fact, k is taken to be equal to 2 and the algorithm is initialized with the observed values of $\min(X_j)$ and $\max(X_j)$ in order to isolate the cluster consisting of the few time intervals with high values for X_j. A robustness analysis for the number of clusters using the average silhouette width [13] supported the optimal choice of $k = 2$. Indeed Fig. 5 shows the boxplots of the average silhouette width when $k = 2$ for the distribution of X_j, for $j = 1, ..., 113$, on the MV and LV network respectively; values of this index between 0.71 and 1 indicate that a strong structure has been found, values between 0.51 and 0.7 indicate that a reasonable structure has been found, whereas values less than 0.5 say that the found structure is weak and could be artificial. The two-cluster pattern for the distribution of number of faults seems highly reasonable for all P&Cs on both networks MV and LV. A scatterplot of the threshold g_j separating the two clusters for all P&Cs appears in Fig. 6. The OI-cluster is that of the observed values for X_j less than or equal to g_j; a time interval with a number of faults in the OI-cluster is assumed to be an ordinary interval.

ii) *Estimation of the distribution generating the OI-cluster.*

For all P&Cs and on both networks, the distribution generating number

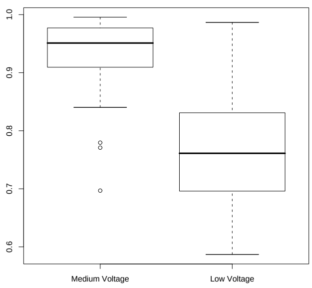

Figure 5: Average silhouette width: 2 groups.

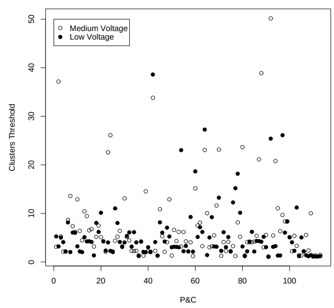

Figure 6: Thresholds defining the OI-clusters.

of faults in the OI-cluster is taken to be a geometric; indeed, for each $j = 1, ..., 113$, the observed values of X_j in the OI-cluster are considered as a sample from a geometric distribution whose parameter is estimated by maximum likelihood.

The geometric is in fact a simplistic model and it does not always adequately fit data belonging to OI-clusters; it has been assumed mainly for the purpose of having a unique model for the distribution generating the OI-cluster for all P&Cs and on both networks. Among simple parametric models for discrete data concentrated on the first few integers the geometric seemed to be the one that overall fitted best the 113*2 data samples. Moreover for a random variable Y with geometric distribution of parameter $p \in (0, 1)$ (i.e. $P(Y = k) = p(1 - p)^k$ for $k = 0, 1, ...$) the mean value is $\mu = (1 - p)/p$ and, if $q_{1-\alpha}$ indicates the $1 - \alpha$ percentile of the distribution of Y, then

$$q_{1-\alpha} = \left[\frac{log(\alpha)}{log(\mu/1 + \mu)} \right]$$

where $[x]$ indicates the integer part of x. For moderately large values of μ, the function within brackets can be approximated by a multiple of μ. Recalling that the percentile $q_{1-\alpha}$ of the estimated distribution generating data in the OI-cluster identifies the exceptionality threshold, the choice of a geometric model is thus coherent with Ofgem choice of a threshold equal to a multiple of the average number of faults in a given time interval.

iii) *Identification of the exceptionality threshold.*

For each P&C and on a given network, the exceptionality threshold is identified by the $1 - \alpha$ percentile of the geometric distribution fitted on the OI-cluster; a time interval is declared EI, for the specific P&C and on the given network, if the number of faults in the interval is greater than $q_{1-\alpha}$. Of course, different values of $\alpha \in (0, 1)$ generate different thresholds.

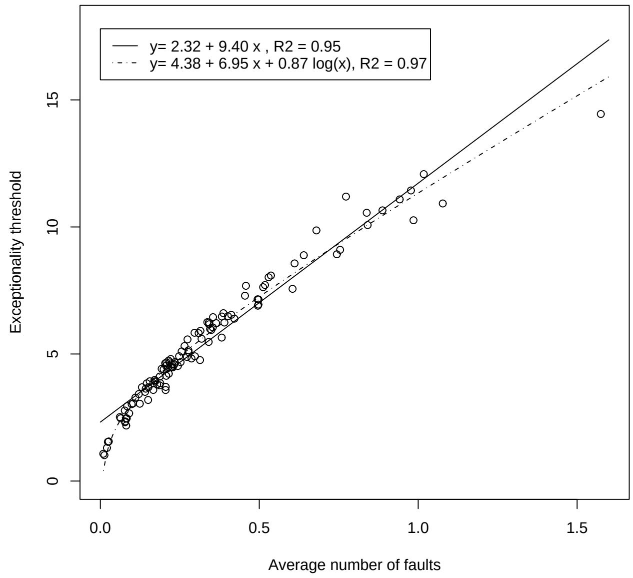

Figure 7: Exceptionality thresholds versus average number of faults on the MT network in a 6-hour time interval for all 113 P&Cs (years 2004-2006)

In order to have a robust grip on the meaning of $q_{1-\alpha}$, consider a sequential i.i.d. sampling from the distribution fitting the OI-cluster; the expected number of trials before observing a value greater than $q_{1-\alpha}$ is $N = 1/\alpha$. Since in fact each observation is relative to a time interval of length 6 hours, this corresponds to an expected time $t = 6/(\alpha * 8760)$ measured in years. Hence $\alpha = 6/(t * 8760)$, and different values of $q_{1-\alpha}$ were compared for $t = 20;\ 30;\ 50$; or 100 years.

iv) *Elicitation of the exceptionality threshold*

For ease of implementation, AEEG suggested to elicit a simple computational algorithm for identifying the exceptionality threshold as a function of the average number m of faults in a 6-hours time interval as observed in the three year time period 2004-2006.

Fig. 7 shows the scatterplot of the values of m and of the exceptionality threshold $q_{1-\alpha}$ on the MV network, for all 113 P&Cs, when α is set according to $t = 20$ years. Multiple regression suggests to fit the model $q_{1-\alpha} = \beta_0 + \beta_1 * m + \beta_2 * log(m)$; a parsimonious model linear in m was however deemed more appropriate by AEEG. Notice that the the fitted values of $q_{1-\alpha}$ on the simpler model $q_{1-\alpha} = \beta_0 + \beta_1 * m$ overestimate the values of the percentiles for very small and very large values of the predictor m. See Fig. 8 and 9 for the scatterplots of residual for the two fitted

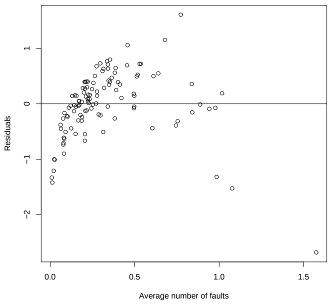

Figure 8: Residuals for the simple linear regression in Fig. 7.

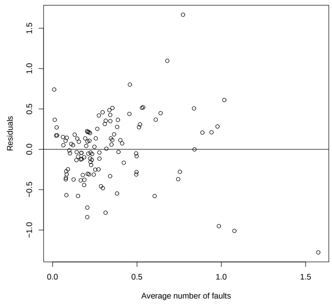

Figure 9: Residuals for the multiple linear regression in Fig. 7.

models.

Table 1 reports the parameter estimates for the model $q_{1-\alpha} = \beta_0 + \beta_1 m$ for the MV network, when α is set by taking $t = 20, 30, 50, 100$ years. Analogous estimates for the LV network appear in Table 2.

Table 1: Parameters of the linear regression model for the elicitation of the exceptionality threshold (MV network).

P&C (MV)	$t = 20$	$t = 30$	$t = 50$	$t = 100$
α	0.9999658	0.9999772	0.9999863	0.9999932
β_0	2.31899	2.44987	2.61476	2.83850
β_1	9.40380	9.77464	10.24184	10.87580

Table 2: Parameters of the linear regression model for the elicitation of the exceptionality threshold (LV network).

P&C (LV)	$t = 20$	$t = 30$	$t = 50$	$t = 100$
α	0.9999658	0.9999772	0.9999863	0.9999932
β_0	3.5753	3.7557	3.9830	4.2915
β_1	7.0716	7.3504	7.7018	8.1785

5 Regulatory decisions (2008)

The Italian regulation of continuity of supply for the period 2007-2011 contains important innovations. Among others, the reward and penalty mechanism that focused only on the SAIDI indicator for long unplanned interruptions was extended to the SAIFI (System Average Interruption Frequency Index) and MAIFI (Momentary Interruption Frequency Index) indicators, calculated respectively on long and short, unplanned interruptions.[9] In addition, a GS was introduced on maximum restoration times, with a specific provision that differentiates penalty payments in case of normal and exceptional circumstances [14].

Building on the statistical analysis described in Section 4, a new criterion for the identification of 'exceptionality' was adopted for both regulatory instruments. In particular, with the Regulatory Order 333/07, AEEG formulates a rule that associates exceptionality to interruptions occurring during an Exceptional Period of time.

For a specific distribution utility an Exceptional Period (EP) is defined as follows [14]. Given a province j, a year t, a voltage level (V), and a 6-hour time interval I, set H_1 at 3 hours before the beginning of I and H_2 at 3 hours after the end of I. The period of time between time H_1 and time H_2 is defined an EP if

$$X_V^j(I) > \beta_0 + \beta_1 \cdot m \tag{1}$$

where:

$X_V^j(I)$ is the observed number of not-notified, long interruptions originated on the MV (LV) network for the province j, in the 6-hour time interval I;

m is the three-year average (over the years $t-2$, $t-3$, $t-4$) of the observed values of X_V^j over all 6-hour time intervals;[10]

The coefficients in (1) are:

- $\beta_0 = 2.3$; $\beta_1 = 9.4$ for the MV level;
- $\beta_0 = 3.5$; $\beta_1 = 7.1$ for the LV level;

In other words, the methodology first identifies EIs, as those 6-hour intervals during which the number of registered faults is higher than the exceptionality threshold. As an example, In Fig. 10 and Fig. 11 we find the boxplots of the distribution of the number of EIs in the three years 2004-2006 for the 113 P&Cs operating in the MV network and in the LV one respectively.

Then, the methodology labels as an EP a larger time span. In the simplest case the EP covers a period of 12 hours. If more than one subsequent EI is found, the EP will include a higher number of hours.

The use of EP in the reward and penalty mechanism is as follows [14]. Given a province j, and a year t, a long interruption is an Exceptional Long Interruption (ELI) if:

[9] Interruptions included in the calculation of the MAIFI indicator are short ones: interruptions longer than one second and shorter than three minutes.

[10] The average includes also those 6-hour intervals with zero interruptions.

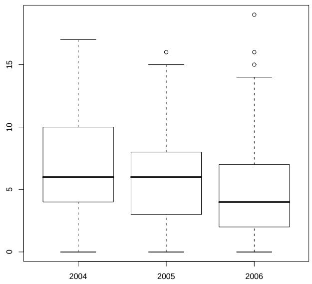

Figure 10: Flanked boxplots for number of EI in the years 2004-2006 (MV network).

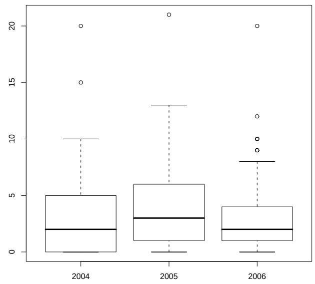

Figure 11: Flanked boxplots for number of EI in the years 2004-2006 (LV network).

- it begins during an EP and
- its duration (in minutes) is above the 3rd quartile of the three-year distribution (over the years $t-2$, $t-3$, $t-4$) of the durations of long interruptions, for the province j.

Moreover, given a province j (for local distribution companies, the portion of the province served by the utility), and a year t, a short interruption is an Exceptional Short Interruption (ESI) if it begins during an EP.

The customer-minutes-lost of ELIs are not considered in the calculation of the actual-SAIDI; in addition, ELIs, and ESIs do not enter in the computation of, respectively, the actual-SAIFI and actual-MAIFI.

The use of EP within the framework of the GS on maximum restoration times requires a detailed explanation [14].

The GS on maximum restoration times obliges distribution utilities to pay a reimbursement to customers when maximum restoration times are exceeded. Standards are differentiated as per voltage level and territorial density (Table 3). Similarly, reimbursements are differentiated as per voltage level and installed capacity of the customer.

One of the main difficulties with GS on very long interruptions is the identification of exemptions from payments or, equally, the definition of those cases when the interruption was not under the responsibility of the distribution utility. Such difficulty emerges for the regulator in the relationship with both the utility and the customers. For the utilities the problem is one of limiting their the financial exposure. For the consumers, the problem is in communicating to a large public that there are circumstances when reimbursements will not be paid.

The Italian regulation addresses these issues in an innovative way. On the one hand, reimbursements are automatically paid to customers under all possible circumstances, whenever maximum restoration times are exceeded (also in case

Table 3: GS on maximum restoration times, Italy

Interruption type	Territorial density	Maximum time, LV [hours]	Maximum time, MV [hours]
Not-notified	High density	8	4
	Medium density	12	6
	Low density	16	8
Notified	All densities	8	8

of faults on the transmission grid).[11] On the other hand, the utility responsibility is accounted for in the following way. Reimbursements are normally paid by the distribution utility, or eventually shared between the distribution and the transmission company (in proportions given by the respective fault durations). However, the financial exposure of the utilities is capped[12] and three forms of exemptions are considered:

- For interruptions occurring during an EP;
- For interruptions caused by Exceptional Climatic Events: climatic conditions that exceed the technical limits of the network;
- For the period of time called Suspension of Clock: the clock counting the interruption duration may be suspended when the utility considers it not safe for repair teams to carry out the work necessary to restore supply.

All reimbursements that are not covered by the utilities are paid by an Emergency Fund. This fund is created with contributions from customers and from regulated utilities. Customers contribute with a small adjustment in the distribution tariff.[13] Distribution companies contribute with payments that are a function of the number of LV customers that received interruptions longer than 8 hours in the previous year.

Provisions such as the cap on the financial exposure of the utilities and the suspension of clock, or the idea that consumers may contribute to the expenses for reimbursements (when payments exceed the cap), are similar to rules that are found also in the UK regulation. Nonetheless, the creation of the Emergency

[11] The only exception is when customers have been evacuated by order of the public authority, in case of a natural calamity.

[12] At 2% and 7% of revenues respectively for distribution and transmission companies.

[13] For instance, a LV domestic customers contributes with 0.35 €/year.

Fund enables AEEG to secure reimbursements to customers for a larger set of interruption types, including those originating on the transmission grid. More notably, this setting enables AEEG to define maximum restoration times that are significantly lower that those set by the UK regulation.

6 Conclusions and further work

The use of statistical methodologies for the identification of exceptional events may present desirable properties such as fairness to all regulated utilities and simplicity in application. Nevertheless, particular care is necessary in the choice of both the continuity indicator as well as the spatial and temporal unit of the analysis. Moreover, specific strengths and weaknesses of the methodology might emerge more clearly only after a few years of experience. Therefore, it is important to keep monitoring it and to identify and propose potential improvements.

As far as the Italian experience is concerned, a new statistical methodology for the identification of exceptional events was adopted in 2008. Several issues with the use of the Two-Step method, employed in the previous four years, suggested the use of a different indicator of performance (the number of faults, instead of the SAIDI and SAIFI), of a larger spatial unit (the province, instead of the district) and of a shorter and more flexible time frame (the 6-hour time interval, extendable to include additional hours).

An elaborate and accurate statistical analysis was carried in order to support these choices and to define the exceptionality thresholds. The methodology finds an application in two regulatory instruments that control continuity of supply. In particular, the structure of penalty payments under the new GS on maximum restoration time is remarkable in the sense that customers receive reimbursements for very long interruptions also in cases of exceptional conditions.

The attention of the Italian regulator for consumer protection under extreme conditions is now finding a new direction in monitoring the definition of Emergency Plans (a new requirement for all distribution utilities) and of the design criteria of overhead lines (a work undertaken by CEI, the Italian Electrotechnical Committee).

References

[1] CEER, "Third benchmarking report on quality of electricity supply," Council of European Energy Regulators, Brussels, 2005.

[2] E. Fumagalli, L. Lo Schiavo, F. Delestre, "Service quality regulation in electricity distribution and retail," Springer-Verlag, Berlin Heidelberg, 2007.

[3] Ofgem, "Revised standards of performance arrangements for electricity distributors," Consultation Document, January 2005, 03/05. [Online]. Available: www.ofgem.gov.uk.

[4] E. Fumagalli, L. Lo Schiavo, S. Salvati, P. Secchi, "Statistical identification of Major Event Days: an application to continuity of supply regulation in Italy," *IEEE Transactions on Power Delivery*, vol. 21, no. 2, pp. 761-767, April 2006.

[5] R. D. Christie, "Statistical classification of major event days in distribution system reliability," IEEE Trans. Power Del., vol. 18, no. 4, pp. 1336–1341, Oct. 2003.

[6] IEEE Standard 1366-2003. *IEEE Guide for Electric Power Distribution Reliability Indices*, IEEE, New York, NY, May 2004.

[7] Ofgem, "Standards of performance: interim arrangements," Document 151/03, October 2003. [Online]. Available: www.ofgem.gov.uk.

[8] S. Salvati, E. Fumagalli, L. Lo Schiavo, P. Secchi, "Analysis of the methodology adopted by the Italian regulatory authority for identifying Major Event Days," 9th International Conference on *Probabilistic Methods Applied to Power Systems*, Stockholm, Sweden, June 11-15, 2006.

[9] C. Warren, B. Saint, "IEEE reliability indices standards - Major event day calculations and how it effects small utilities," IEEE Conference Paper No. 04 D2.

[10] Ofgem, "The Electricity (Standards of Performance) Regulations 2005," Statutory Instruments 2005 No. 1019. [Online]. Available: www.opsi.gov.uk.

[11] Ofgem, "Electricity distribution price control review. Final proposals," Document 265/04, November 2004. [Online]. Available: www.ofgem.gov.uk.

[12] J. A. Hartigan, M. A. Wong, "A K-means clustering algorithm," *Applied Statistics*, vol.28, pp. 100–108, 1979.

[13] A. Struyf, M. Hubert, P. J. Rousseeuw, "Clustering in an Object-Oriented Environment," *Journal of Statistical Software*, 1, 1996.

[14] AEEG. (2007, December) Regulatory Order no. 333/07. [Online]. Available (in Italian): www.autorita.energia.it

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

6/2008 E. Fumagalli, L. Lo Schiavo, A.M. Paganoni, P. Secchi:
Statistical analyses of exceptional events: the Italian experience

5/2008 S. Badia, A. Quaini, A. Quarteroni:
Modular vs. non-modular preconditioners for fluid-structure systems with large added-mass effect

4/2008 R. Milani, A. Quarteroni, G. Rozza:
Reduced basis method for linear elasticity problems with many parameters

3/2008 P. Zunino, C. D'Angelo, L. Petrini, C. Vergara, C. Capelli, F. Migliavacca:
Numerical simulation of drug eluting coronary stents: mechanics, fluid dynamics and drug release

2/2008 S. Berrone, M. Verani:
An Adaptive Gradient-DWR Finite Element Algorithm for an Optimal Control Constrained Problem

1/2008 R. Rosso, M. Verani:
Stabilizing rôle of a curvature correction to line tension

26/2007 M. Colecchia, N. Nicolai, P. Secchi, G. Bandieramonte, A.M. Paganoni, L.M. Sangalli, L. Piva, R. Salvioni:
Penile Superficial Squamous cell Carcinoma (SCC) Submitted to CO_2 Laser Excision only: Oncologic Outcome of T1 Disease in 25 Years-Long Experience

25/2007 S. Micheletti, S. Perotto:
Space-Time Adaption for Advection-Diffusion-Reaction Problems on Anisotropic Meshes

24/2007 L.M. Sangalli, P. Secchi, S. Vantini, A. Veneziani:
A Case Study in Functional Data Analisys: Geometrical Features of the Internal Carotid Artery

23/2007 L.M. Sangalli, P. Secchi, S. Vantini, A. Veneziani:
Efficient estimation of 3-dimensional centerlines of inner carotid arteries and their curvature functions by free knot regression splines