IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 January 2024, accepted 12 February 2024, date of publication 19 February 2024, date of current version 23 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3367591

APPLIED RESEARCH

Statistical Design of Experiments for Power System Protection Testing: A Case Study for Distance Relay Performance Testing

MIRKO GINOCCHI[1], (Member, IEEE), THANAKORN PENTHONG[1], FERDINANDA PONCI[1], (Senior Member, IEEE), AND ANTONELLO MONTI[1,2], (Senior Member, IEEE)

[1]Institute for Automation of Complex Power Systems, RWTH Aachen University, 52064 Aachen, Germany
[2]Fraunhofer FIT, Center for Digital Energy, 52068 Aachen, Germany

Corresponding author: Mirko Ginocchi (mirko.ginocchi@eonerc.rwth-aachen.de)

This work was supported by the Interoperability Network for the Energy Transition (IntNET), which is a project funded by the European Union's Horizon Europe Research and Innovation Program under Grant 101070086.

ABSTRACT In modern power systems, testing protection systems and equipment before the field implementation is vital to ensure the correct operation and guarantee reliability, quality and security of electricity supply. This paper focuses on the performance testing of the distance protection, whose state-of-the-art test methodologies (including the recommendations of the IEC 60255-121:2014 standard) can quickly lead to time/money resource limitations. As budget-wise considerations should not justify an "arbitrary" or "convenient" selection of which and how many tests to perform, this paper shows how the current testing methodologies benefits from the statistical design of experiments (stat-DOE). It is proven how the stat-DOE supports the performance testing in the efficient selection of the optimal tests to conduct, in the systematic investigation of the effect of different factors, and in the robust definition of pass/fail criteria for specifying acceptance tests. A step-to-step practical guideline for adopting the stat-DOE is offered to conduct a realistic performance testing, accounting for operator-specific requirements (e.g., maximum affordable number of tests) and physical constraints among factors. The results allow to propose lines of refinement and recommendations for the stakeholders. Finally, not only this paper serves as a guide to replicate the performance testing of other protection functions or in alternative scenarios, but it also ultimately paves the way for the routine adoption of the stat-DOE in the definition and refinement of test methodologies for power system protection testing at large towards the achievement of a standardized basis for it.

INDEX TERMS Design of experiments, distance relay, IEC 60255-121:2014, performance testing, power system protection.

I. INTRODUCTION

The increasing complexity of modern power grids calls for more sophisticated protection functions to ensure correct operation of the power system as well as the reliability, quality and security of supply [1]. Thus, testing the protection functions embedded in protection equipment and systems turns to be a vital activity, to ultimately prevent the potentially disruptive consequences of protection failures, such as major power outages and widespread cascading events. Power system protection testing is a multi-faceted and broad-scope activity. Different types of tests (both device-specific and application-oriented) are identified by the *IEEE Guide for Power System Protection Testing* [2], such as certification, application, commissioning and maintenance tests. Among these, this paper focuses on performance tests, i.e., a type of certification tests typically conducted when a new relay has to be used in the power system or a firmware has to

The associate editor coordinating the review of this manuscript and approving it for publication was Padmanabh Thakur.

be upgraded. These tests assess the ability of the relay to perform based on its specifications and verify its behaviour under realizable power system conditions, including fault conditions. To test the performance of protection functions, testing platforms based on real-time simulation are usually developed in Hardware-in-the-loop (HiL) to model (complex) power systems and enable their connection with the physical equipment under test to investigate the response under conditions similar to real operation (e.g., [3], [4], [5], [6]). Besides testing platforms, testing methodologies are also required to assess the protection schemes, and standardization efforts have been made to help users evaluate protection functions on a formal basis with respect to relay selection, setting, commissioning, application, and operation [7].

Among all the different protection functions, this paper focuses on the distance protection as the function subject of the testing procedure, due to its broad adoption in transmission systems for its simplicity, selectivity, and dependability [8]. The distance protection function works by measuring the line impedance and comparing it to a threshold value. The impedance seen by the relay is calculated based on voltage and current measurements at the relay location. As such impedance is a proportion of the distance from the measurement location to the fault point, the fault location recorded by the relay can be computed. In general terms, the performance of the distance protection function can be seen as a process f:

$$\mathbf{Y} = f(\mathbf{U}) \tag{1}$$

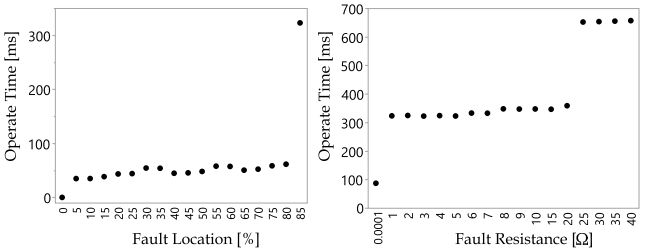
which takes as input a set of factors $\mathbf{U} = \{U_i\}_{i\in\{1,\dots,K\}}$ (e.g., fault location, fault resistance) and outputs a set of measurable response variable(s) $\mathbf{Y}$ (e.g., the operate time of the relay). To assess the distance protection performance and investigate which factors have the biggest impact on it, one or more *experiments* (defined as a series of N_T tests) are conducted by deliberately changing the factors $\mathbf{U}$ at different values (e.g., applying faults at different locations with different fault resistance) and recording the response(s) of interest.

A. STATE-OF-THE-ART OF EXPERIMENTAL STRATEGIES FOR DISTANCE PROTECTION PERFORMANCE TESTING

The experimental strategies met in the literature for the distance protection performance testing can be categorized into one-at-a-time (OAT) and factorial, as discussed next.

1) ONE-AT-A-TIME EXPERIMENTS

The most straightforward experimental strategy is the OAT method, which consists in choosing an initial point $\mathbf{U}^0$ (where all the factors are set at their baseline or nominal values), and successively varying each factor *one at a time* over its range of interest with the other factors held constant at the baseline level, possibly iterating such approach by selecting different nominal points $\mathbf{U}^0$ [9], [10]. Examples of performance tests based on OAT experiments can be found in [5], [11], and [12]. For instance, Figure 1 depicts an OAT experiment to test the distance relay performance (in terms of operate time) depending on changes in two factors, namely fault location and fault resistance, in the case of a single-phase-to-ground fault (AN) with inception angle of 0°. Assuming that Zone 1 covers 85% of the protected line, the effect of different fault locations (from 0% till 85%) is tested (Figure 1a), with fault resistance fixed at 1Ω; then, the fault resistance is varied over multiple values from 0.001Ω to 40Ω, with fault location fixed at 85% (Figure 1b). Such OAT method is undoubtedly appealing and intuitive: changing one factor at a time logically implies that whatever observed effect on the process response can be exclusively attributed to the specific factor being varied in each test. For example, it might be inferred that, at the nominal conditions studied in the two sets of tests (i.e., 1Ω and 85%), the fault resistance has a major effect in increasing the operate time (Figure 1b), with the fault location being important only when the fault occurs at the boundary of Zone 1 (Figure 1a). Yet, OAT methods showcase at least two major methodological flaws [9], [10]. First, they fail in detecting potential interactions between factors, i.e., when the effect of one factor on the observed response depends on the level of another factor (e.g., being magnified or dampened). This may lead to the incorrect classification of the factors' importance for the distance protection performance and ultimately to poor results (which might even change if different nominal points $\mathbf{U}^0$ are chosen). Second, they are ineffective in dealing with the "curse of dimensionality", i.e., the number of tests to conduct might easily exponentially grow as the number of factors to investigate increases.



(a) Fault Resistance fixed at 1Ω (b) Fault Location fixed at 85%

FIGURE 1. Example of an OAT experiment.

2) FACTORIAL EXPERIMENTS

Opposite to the OAT method is the factorial strategy, which is based on varying multiple factors *together* instead of *one at a time*. Although less intuitive than the OAT method, the factorial strategy is a more robust approach when dealing with several factors as interactive effect among factors might emerge and be effectively captured. In its general form, a factorial experiment is based on the so-called "full factorial" design: the effect of K factors $\{U_1, \dots, U_K\}$ on the process response is assessed by discretizing the factor variability range into a given number of values $\{L_1, \dots, L_K\}$, the so-called "levels", and conducting tests for each factor level combination (or "run"), sometimes replicating the same

experimental run R times, for a total of $N_T = L_1 \times \ldots \times L_K \times R$ tests. Examples of performance tests based on full factorial designs can be found in [4], [12], [13], [14], and [15], as well as in the IEC 60255-121:2014 standard [16]. In particular, the latter was issued by the IEC Technical Committee 95 in 2014 with the intent to address the lack of uniformity among testing methodologies, prevent misunderstandings within the protection relay community, and produce a standard procedure to evaluate and compare protection function performance claims from different manufacturers. The sequence of tests suggested by the IEC 60255-121:2014 standard [16] is reproduced in the flowchart of Figure 2: four factors are considered, namely source impedance ratio (SIR), fault location, type, and inception angle, and each of them is studied at specific levels (4, 7, 4 and 4, respectively); all the other factors (e.g., fault resistance, other relay settings) are held constant at nominal values. As each fault is injected 4 times, $N_T = 1792$.

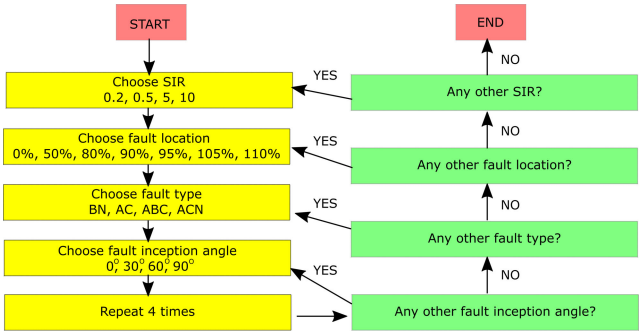


FIGURE 2. Flowchart of the sequence of tests suggested by [16] (adapted).

B. MOTIVATIONS OF THE WORK

In Section I-B1, the main challenges and practice lacks deriving from the above-discussed literature review are presented. Based on these, Section I-B2 illustrates how this paper intends to overcome them and to go beyond the state-of-the-art.

1) MAIN CHALLENGES IDENTIFIED IN THE STATE-OF-THE-ART

From the state-of-the-art review of Section I-A, three main challenges can be identified.

(i) The distance protection performance testing conducted in the literature (employing either OAT or full factorial designs) is limited to measuring one or more responses of interest, without an in-depth analysis of the results. Consider e.g., Table 1, which shows an excerpt of the full factorial experiment performed in [14], where the relay operation is analysed for different fault locations and types ("✗" and "✓" signalize failure and success in detecting the fault, respectively). From Table 1, the only possible inference is that, if an AN fault occurs, the relay fails to operate in both Zone 1 and 2, regardless of the fault location. Although such factorial experiments—which explore all the possible combinations of the factors' levels and are hence more informative and efficient than OAT strategies—can be effective in spotting some combinations of factor levels leading to the relay misoperation, no attempt to perform further analysis is carried out. For example, again with reference to Table 1, no assessment of the statistical significance of the effects of fault location and type is performed, no empirical model of the relay operation as function of the latter is built, and no "operability region" in the design space (i.e., the space containing all the possible factors' combinations) is even roughly defined.

TABLE 1. Example of a factorial experiment (adapted from [14]). Zone 1 and 2 are set to cover 80% and 150% of the protected line, respectively.

Run	Fault location	Fault type	Operation
1	100%	AN	✗
2	100%	BC	✓
3	100%	ABC	✓
4	75%	AN	✗
5	75%	BC	✓
6	75%	ABC	✓
7	50%	AN	✗
8	50%	BC	✓
9	50%	ABC	✓

(ii) When following the recommendations of the IEC 60255-121:2014 standard [16], it turns out to be inevitable to conduct hundreds or thousands of tests. This is even more so when a wider set of scenarios is to be investigated to check compliance with utility-specific requirements [15], and/or more factors than those in [16] are worth to test, such as the fault resistance (as done e.g., in [4], [15], [17]), which [16] suggests to hold constant at 0Ω. These aspects introduce challenges in terms of cost and duration of the testing activity—especially in resource-saving contexts, e.g., due to the required experimental time—, which are sometimes used to justify an "arbitrary" or "convenient" selection of the tests to perform, e.g., *subjectively* limiting the number of factors to study, avoiding the collection of replicated tests, etc.

(iii) The size of the experiment prescribed by factorial designs might not be compatible with the maximum number of tests that the operator can afford. For instance, [13] conducted nearly 10^5 tests with an average of 5 tests per minute, which, considering 8 working hours per day, translates into about 40 days of tests. Also, the scenarios under test might own properties that lead the design space to be characterized by some physical constraints that ultimately translate into correlations among factors and/or disallowed combinations thereof (e.g., phase-to-phase faults having non-null resistance). The full factorial design prescribed by the IEC 60255-121:2014 standard [16] cannot accommodate user-specific requirements as well as physical constraints among factors.

TABLE 2. Overview of the workflow for the stat-DOE [9].

STEP 1. Specify problem and objectives
STEP 2. Select the response variable
STEP 3. Choose factors, levels and ranges
STEP 4. Select the design
STEP 5. Conduct the experiment
STEP 6. Statistically analyse the data
STEP 7. Extract conclusions and recommendations

2) GOING BEYOND THE STATE-OF-THE-ART

The above challenges call for strategies which might (i) enable robust and deeper analysis of the experimental results, (ii) aid in the optimal choice of the tests while retaining the experimental effort under control and holding a greater level of *objectivity* in the testing procedure, and (iii) flexibly account for operator-specific requirements and physical constraints among factors. These features can be effectively tackled by the statistical Design of Experiments (stat-DOE), whose steps are reported in Table 2 and detailed in Section III. The stat-DOE combines the strength of the classical DOE [18], i.e., the process of laying out a detailed plan in advance of carrying out the real experiments, with the power of the statistical approach to properly collect data (especially in resource-saving contexts) and statistically analyze them, ultimately enabling scientifically sound and robust inferences (see e.g., [9], [10]). The idea of adopting a stat-DOE based approach for testing procedures is not novel; e.g., in the smart grid interoperability testing methodology [19] developed by the European Union Joint Research Centre, the stat-DOE is an integral part of the test procedure [20], [21]. Yet, to the best of authors' knowledge, never has the stat-DOE theory been applied in the power system protection testing. Thus, the ultimate objective of this paper is to show which benefits the stat-DOE can bring to the testing methodologies for power system protection, with specific focus on the distance protection function.

In this paper, the challenges of Section I-B1 are tackled by resorting to the stat-DOE as discussed hereafter.

(i) Accurately designing an experiment via the stat-DOE enables to robustly analyze the experimental results and equip the testing procedure with a higher degree of objectivity. In fact, the stat-DOE allows to statistically characterize the variation of the response, to learn which factors have the greatest influence on it, to identify cause-effect relationships by fitting an empirical model and using it to predict the process behavior at unexplored factor levels' combinations, to attach a given level of statistical confidence to any statement or conclusion, and ultimately to support decision-making (e.g., compliance check with utility-specific requirements or standard guidelines before the field implementation of protection devices/systems). The wealth of possibilities that would stem at no cost from the statistical methods envisaged by the stat-DOE (i.e., Step 6 of Table 2) is demonstrated in Section V in an exemplary three-factor scenario, where the full factorial design is used to reflect the most widespread state-of-the-art experimental strategy.

(ii) The number of tests with a full factorial design increases exponentially as the number of factors and/or the number of their levels grow. To face such curse of dimensionality, the stat-DOE theory offers alternative designs that allow objectively reducing the number of tests while maintaining a satisfactory level of information. One of the most adopted designs to meet such trade-off among economy and efficiency is the so-called "fractional factorial", which is illustrated in the context of the stat-DOE in Section VI in an exemplary four-factor scenario.

(iii) Full and fractional factorial designs belong to the family of *classical* designs, which have represented the common practice since the early origins of the stat-DOE theory [18], [22], [23]. Classical designs work well when the scenario under study can be securely mapped to any of them, and no particular physical constraints or operator-specific requirements are present. As this is not always the case (or at least only partially), it turns out to be more convenient to create an optimal "custom" design for the specific experimental quest, instead of forcing the problem to fit itself to the available off-the-shelf classical designs: this is the rationale behind the *modern* designs, which have been establishing as the best-practice within the stat-DOE community, and many general-purpose statistical software packages have become available to generate them (e.g., [24], [25], [26]). When compared to classical designs, modern designs (introduced in Section VII) offer wider flexibility, by allowing not only the optimal choice of the tests to conduct in the presence of operator-specific requirements (e.g., maximum number of affordable tests), but also to effectively work with a constrained design space. Due to their broad-scope applicability, modern designs are used in Section VIII, where a guideline of the stat-DOE application is illustrated to emulate a realistic performance testing of a commercial distance relay in an HiL set-up.

C. CONTRIBUTIONS OF THE WORK

This paper brings the following novel contributions:

1) the stat-DOE theory is introduced for the first time in the field of power system protection testing, by showing how the former can be mapped to the latter; the full and fractional factorial designs are revamped to test the distance protection performance in two exemplary scenarios; the benefits of the stat-DOE to the state-of-the-art testing of distance protection performance (including the IEC 60255-121:2014 standard) are demonstrated;
2) the design efficiency of the two main experimental strategies encountered in the literature (i.e., OAT and full factorial) is compared in resource-saving contexts, i.e., in terms of amount of information that can be extracted

when constraints are posed on the number of tests to afford (e.g., due to limiting time/money budget);

3) a stat-DOE application guideline is showcased to conduct an HiL testing of the distance performance by adopting the current best-practice of modern designs; the process and the results are used to compile lines of refinements and recommendations for all the stakeholders involved in the standardization chain (i.e., relay manufacturers, utilities, standardization bodies).

D. STRUCTURE OF THE PAPER

Section II provides theoretical background of the distance protection function. Section III offers a brief overview of the stat-DOE and maps its seven steps to the distance protection performance testing. Section IV describes the test set up adopted for all the experiments performed throughout the paper. Section V describes the full factorial design and discusses the wealth of possibilities stemming from the statistical methods of the stat-DOE in an exemplary three-factor scenario. Section VI describes the fractional factorial design by actualizing it in an exemplary four-factor scenario. Section VII presents the concept of the modern designs and provides a technical comparison of the OAT and factorial designs from the design efficiency viewpoint. Section VIII illustrates the guideline to conduct a realistic performance testing via the stat-DOE with the current best-practice of modern designs, and extracts lines of refinements and recommendations for all the stakeholders. Section IX concludes the paper.

II. BACKGROUND ON DISTANCE PROTECTION

The distance protection (which is the most widely used protection function in transmission systems) estimates the physical distance between the relay's sensors (voltage and current transformers) and the location of the fault, by measuring the line impedance and comparing it to a threshold value. If the impedance seen by the relay is smaller than the threshold value, the relay triggers a signal to the appropriate circuit breaker to isolate the faulted line section.

Distance relays are designed to trip specific breakers to protect limited zones of the power system, each of them characterized by different time grading. Consider the exemplary transmission system of Figure 3 with radial configuration, whereby a distance relay is located at substation A, and the impedance (*R*-*X*) diagram of Figure 4. In general, four protection zones can be set within the relay.

- The relay setting for Zone 1 (in forward direction) usually covers 80% to 85% of the impedance of section A-B.
- The relay setting for Zone 2 (in forward direction) covers 100% of the impedance of section A-B plus 50% of the impedance of the shortest line (section B-C). In the case the remote substation misses the outgoing line (e.g., substation B), but only the transformer bay (TX) is present, Zone 2 covers 100% of the impedance of section A-B plus 20% of the TX impedance.

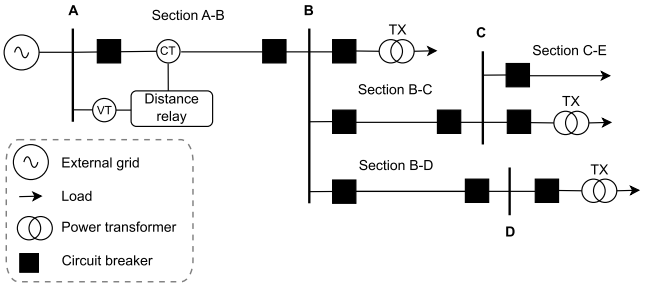


FIGURE 3. Sample transmission system in radial configuration.

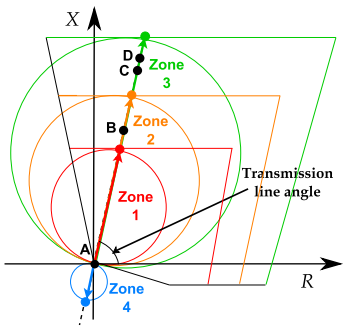


FIGURE 4. Phase and ground characteristic of the distance protection.

- The relay setting for Zone 3 (in forward direction) covers 100% of the impedance of section A-B plus 120% of the impedance of the longest line (section B-D). If there is an overlapping problem of Zone 3 between the relays at substations A and B, the Zone 3 setting of the relay at substation A covers 100% of the impedance of section A-B plus 100% of the impedance of section B-C plus 25% of the impedance of section C-E. In the case the remote substation misses the outgoing line and only the TX is present, Zone 3 covers 100% of the impedance of section A-B plus 60% of the TX impedance.
- The relay setting for Zone 4 (in reverse direction) is usually equal to 15% of the Zone 1 setting. Zone 4 is used as backup protection of the bus bar at the substation A. Yet, for radial passive systems, Zone 4 does not operate since no current reversal is present, i.e., the fault contribution in the reverse direction is null.

Mho and quadrilateral characteristics are used by distance relays to detect phase and ground faults, respectively, as shown in Figure 4. To calculate the fault location, the distance protection function receives voltage V and current I signals from the instrument transformers at the relay location, from which the impedance of the transmission line Z^r is computed [27]. For example, in the case of a single-phase-to-ground fault AN, the impedance seen by the relay Z_A^r is computed as:

$$Z_A^r = \frac{V_A^r}{I_A^r + k_0 I_A^r} \tag{2}$$

where $k_0 = (Z_0 - Z_1)/3Z_1$ is the compensation factor (with Z_0 and Z_1 indicating the zero- and positive- sequence

impedance of the line, respectively). On the other hand, in the case of e.g., a phase-to-phase AB fault, the impedance seen by the relay Z^r_{AB} is computed as:

$$Z^r_{AB} = \frac{V^r_A - V^r_B}{I^r_A - I^r_B} \tag{3}$$

Similar formulas hold for the other fault types [27].

III. STATISTICAL DESIGN OF EXPERIMENTS

In this section, the steps of the stat-DOE workflow (see Table 2) are briefly described and mapped to the specific context of the distance protection performance testing. The JMP® toolkit [26] supports the stat-DOE throughout this work.

A. STEP 1 – SPECIFY PROBLEM AND OBJECTIVES

The problem at hand and the question(s) to address via the specific experiment(s) are defined. In general, experiments can be conducted e.g., for factor screening (to identify the influential factors out of many factors to investigate); optimization (to find the optimal settings of the influential factors which yield ''desirable'' values of the response); confirmation (to verify that the process behaves according to some preliminary knowledge or hypothesis); robustness (to study under which conditions the response takes on values in ''unwanted'' regions). In practice, the (distance) protection performance testing can be undertaken not only by the relay manufacturers to evaluate the relay design and check the compliance with standards, but also by the utilities to check the accuracy of the protection relay and define acceptance tests, including acceptance criteria and requirements related to the specific application [2], [7], [28].

B. STEP 2 – CHOOSE THE RESPONSE VARIABLE

One or more measurable responses of interest are identified which provide useful insight about the process under study and can satisfactorily address the experiment's objective. For example, valid metrics for evaluating the distance protection performance can be the operate time, or the selectivity of the relay at a specific protection zone [4]. Statistical indicators are sometimes extracted from the measured responses, e.g., the minimum and maximum values of the operate time, as well as the mode, median and mean [16].

C. STEP 3 – CHOOSE FACTORS, LEVELS AND RANGES

The factors which might affect the response and are the main target of the experiment are carefully selected, whereas the others which may exert some effect on the response but either are not of interest or it is convenient to ''control'' are set at a specific level. Selecting which factors belong to either classes (referred to as *design factors* and *held-constant factors*, respectively) can be done by resorting to contributions from stakeholders, subject expertise, standard guidelines, etc. For example, the IEC 60255-121:2014 standard [16] considers SIR, fault location, type and inception angle as the design factors to evaluate the dynamic performance of distance relays (see Figure 2), whereas the held-constant factors are the fault resistance (to be set at 0Ω, or to the minimum allowed value if numerical limitation arises) and all other settings needed for distance protection to perform correctly (to be set to the most common values suggested by the manufacturer).

In addition, the ranges over which the design factors will be varied as well as the levels at which these will be tested are defined. For example, the fault location can be tested from 0% to the boundary of Zone 1 (e.g., 85% of the protected line) if the performance in such protection zone is of interest, or AN, BN, and CN fault types can be injected if only single-phase-to-ground faults are to be studied. Similarly to the selection of design and held-constant factors, the factor ranges to choose as well as the levels for each factor to study are typically the outcome of some form of process knowledge (e.g., in terms of practical experience, theoretical understanding, etc.) or technical guidelines [16], [29].

D. STEP 4 – SELECT THE DESIGN

Out of the plethora of the available designs, the fittest for the objective and specifications of the experiment is selected. Based on the chosen design, a ''design matrix'' is generated (such as the second and third columns of Table 1): each row defines the nth factor level combination to test in each run ($n = 1, \ldots, N$) and each column specifies the level at which each factor is set for the nth run. The design selection should account for the size of the experiment in terms of number of tests that can be afforded (potentially considering replicated runs), and preliminary assumptions on a tentative empirical model linking the design factors to the response variable (e.g., presence of interactions among factors, or quadratic effects for some factors). Two main classes of designs can be identified: classical and modern designs.

- Classical designs are mostly used to introduce stat-DOE concepts, and include e.g., full and fractional designs, response surface designs, mixture designs, Taguchi array designs, split plot designs [9]. Upfront and quite good knowledge of the system/process is required when choosing classical designs, as they prescribe the number/type of factors, the factor levels and model effects as well as the number of tests to conduct.
- Modern designs are the current best-practice for carrying out experiments. They are mostly computer-generated designs that can be optimally customized to meet specific features of the problem at hand [9], hence providing the operator with more flexibility. For instance, the number/type of factors, factor levels and model effects can be freely specified (e.g., based on subject matter expertise), the number of tests to conduct can be tuned to match given operator-specific requirements, and physical constraints among factors can be handled.

E. STEP 5 – CONDUCT THE EXPERIMENT

The experiment designed according to Step 1 to 4 is conducted by testing all the factors' combinations reported in the design matrix. A script might come in handy to automatize the sequential reading of the tests to perform, the configuration of the testing platform with the corresponding power system conditions, and the results' extraction [15].

F. STEP 6 – STATISTICALLY ANALYSE THE DATA

To guarantee that results and conclusions are reliable and objective, statistical methods are employed to analyze the experimental data. These include e.g., statistical test of hypothesis, confidence interval estimation, ANalysis Of VAriance (ANOVA) and graphical methods [9]. Multivariate regression analysis is also used to derive an empirical model that best emulates the unknown functional relationship of (1) between the observed response(s) and the (influential) design factors. Also, follow-up runs usually planned according to design augmentation strategies as well as confirmation tests can be performed to validate the experiment conclusions.

G. STEP 7 – CONCLUSIONS AND RECOMMENDATIONS

Once the data have been analyzed, practical conclusions about the results are drawn and actions are recommended to the involved stakeholders. For example, the statistically designed experiment can suggest which factors have the biggest influence in driving the operate time above utility-specific thresholds, or which factors have small (or negligible) impact and can be studied by considering just fewer levels (or even removed from the set of design factors) in subsequent experiments, hence reducing the number of tests to conduct. Also, the whole standardization chain can receive useful inputs for further refinement, in terms of e.g., whether it is worth integrating additional factors in the performance tests (such as the fault resistance, that the IEC 60255-121:2014 standard prescribes to be excluded from the set of design factors) or indication on how pass/fail criteria can be rigorously defined based on the specific application (about which no directions are given in the IEC 60255-121:2014 standard).

IV. SCENARIOS UNDER STUDY AND TEST SET-UP

A. SCENARIOS UNDER STUDY

In this work, the distance protection performance is tested in three different scenarios. Two exemplary scenarios are elaborated to facilitate the understanding of the full and fractional factorial designs, which are described in Sections V and VI, respectively. In Section VIII, a more close-to-reality scenario is developed to conduct a realistic performance testing and provide a guideline of the stat-DOE application by adopting the current best-practice of modern designs.

B. TEST SET-UP

The test set-up adopted to conduct all the experiments in these three scenarios is described in Figure 5, which depicts the HiL

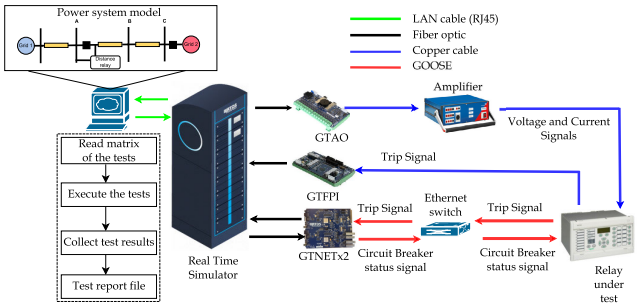


FIGURE 5. Testing platform developed for this work.

testing platform developed, consisting of real-time simulator, CMS-356 amplifier, relay under test, router, as well as GTAO, GTFPI, and GTNETx2 cards. An interface with both GOOSE and copper-wire (output contact) is implemented via GTNETx2 and GTAO/GTFPI cards, respectively.

The adopted power system model, shown in Figure 6 and simulated in the RSCAD® software, is a PI model transmission line consisting of a bundle circuit per phase (connected to the 115 kV grids on both sides as recommended by the IEC 60255-121:2014 standard [16]), with:

- $R_0 = 0.2135\Omega$/km, $X_{0L} = 1.3294\Omega$/km, $X_{0C} = 0.567\Omega$/km for the zero sequence;
- $R_1 = 0.0429\Omega$/km, $X_{1L} = 0.2677\Omega$/km, $X_{1C} = 0.232\Omega$/km for the positive sequence.

The sensors modeled in the power system of Figure 6 are ideal voltage and current transformers (with 115kV/115V and 1800A/1A ratios, respectively); no saturation is considered.

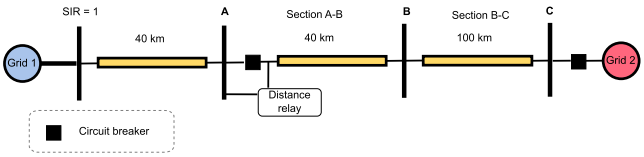


FIGURE 6. Single line diagram of the power system model.

The device under test is the Schneider P543 relay, with four protection zones implemented therein. The phase and ground characteristics of the distance protection function are mho and quadrilateral, respectively. Taking into account the utility requirements, the resistive reach for all the zones of the quadrilateral element is set to 40 primary Ω. The service settings of the relay, calculated according to the power system model of Figure 6, are reported in Table 3.

TABLE 3. Service settings of the distance relay under test.

Zone	Setting (% of line impedance)	Operate time [ms]
1	85% of A-B	0
2	100% of A-B plus 50% of B-C	300
3	100% of A-B plus 120% of B-C	600
4	15% of the Zone 1 setting	600

A script file is used, for the nth test, to (i) read the corresponding row of the design matrix, (ii) configure the

necessary power system conditions (e.g., line impedance based on the fault location, SIR), (iii) emulate the fault inception, and (iv) record the test result. One test is physically conducted as follows: when a fault is applied in the power system, the relay receives voltage and current signals (pre-fault and during fault conditions) via the amplifier. Based on formulas like (2) and (3), the relay decides whether the occurrence of the fault is located inside the protection zones. If so, the relay sends the trip signal using both GOOSE and output contact to the real time simulator, which correspondingly measures the operate time. The closed/open status of the circuit breaker at the relay location is sent to the relay via GOOSE by the real time simulator. The simulation is run for 12 s to ensure that the power system returns to the normal conditions before proceeding to the next test. All the tests are conducted automatically, and the results for each test, including current and voltage waveform, are recorded as a COMTRADE file [30] in the case further analyses are needed.

C. ASSUMPTIONS OF THIS WORK

The assumptions used in this work are described hereafter.

- The measured response of the system is the operate time t, defined as the duration of the time interval between the fault inception and the receiving of the trip signal via GOOSE (without the circuit breaker operation).
- The typical operate time of the circuit breaker (including mechanical process and arc extinguishing) is 100 ms.
- For the sake of simplicity, only the performance of the relay at Zone 1 is evaluated.
- Unless otherwise specified, the utility's viewpoint is considered when discussing pass/fail criteria and operability regions for the relay. Although the main protection zone (Zone 1) shall operate "instantaneously", the utility can accept to have the main protection operate just *faster* than the backup protection (Zone 2 and 3), to prevent wide-area blackout caused by the latter. As shown in Table 3, the operate time of Zone 2 is set at 300 ms. This is done to account for the circuit breaker operation, and to introduce some safety margin which is quantified in 200 ms to include the relay error, the overshoot time of the relay, and the error of the instrument transformers. Hence, the Zone 1 operation of the relay is assumed to be "correct" (i.e., according to the utility requirements) as long as $t < 200$ ms, with 200 ms being considered as the maximum acceptable operate time.

V. FULL FACTORIAL DESIGN

In this section, an experiment based on the full factorial design is elaborated to test the distance protection performance. Such design is described in Section V-B while being actualized in the three-factor scenario of Section V-A. Later, Section V-C shows the wealth of possibilities that the statistical analysis of the experimental data (as foreseen by Step 6 of the stat-DOE) brings if the full factorial design would be adopted in the broader context of the stat-DOE. Different statistical analyses are described in detail and then applied to prove the practical insights that could be gained to ultimately extract statistically sound inferences. Finally, Section V-D discusses the main downsides of the full factorial design and sets the scene for introducing the fractional factorial design.

A. DESCRIPTION OF THE SCENARIO

In this scenario, the effect of fault resistance, location and inception angle is tested on the relay performance. The three factors are tested at two levels: $\{1, 3\}\Omega$ for the fault resistance; $\{84, 85\}\%$ for the fault location (assuming that only the boundary of Zone 1 is of interest); $\{0, 15\}°$ for the fault inception angle. The generic model (1) thus becomes $Y = f(U_1, U_2, U_3)$, where U_1, U_2, U_3 are fault resistance, location and inception angle, respectively, and Y is the operate time.

B. DESCRIPTION OF THE DESIGN

In general terms, a full factorial design with K factors at two levels is called 2^K design. By convention, the "low" and "high" levels of each factor (U_i^+ and U_i^-) are indicated with the orthogonal coding, i.e., with "-1" and "$+1$".

For the scenario of Section V-A, a 2^3 factorial design generates the design matrix reported in Table 4 (second, third and fourth columns), with $N = 8$ rows and $K = 3$ columns: each row indicates the combination of levels at which the ith factor is set at the nth run. By repeating the 8 experimental runs 4 times (as suggested by the IEC 60255-121:2014 standard [16]), 32 tests are conducted. The operate times are collected in the columns Y_1 through Y_4. The columns $Y_{n\bullet}$ and $\bar{Y}_{n\bullet}$ report the sum and the average of the four replicates of the nth run. The grand total $\sum_{n=1}^{4} Y_{n\bullet}$ and the grand mean $\frac{1}{4}\sum_{n=1}^{4} \bar{Y}_{n\bullet}$ are indicated as $Y_{\bullet\bullet}$ and $\bar{Y}_{\bullet\bullet}$, respectively.

The 8 factor levels' combinations to test with a 2^3 design are displayed geometrically as the cube depicted in Figure 7. To identify what factor levels' combination yields the (sum of the) responses of each run, the convention is used to denote the "$+1$" level of any factor U_i by the respective lowercase letter, and the "-1" level of any factor U_i by the absence

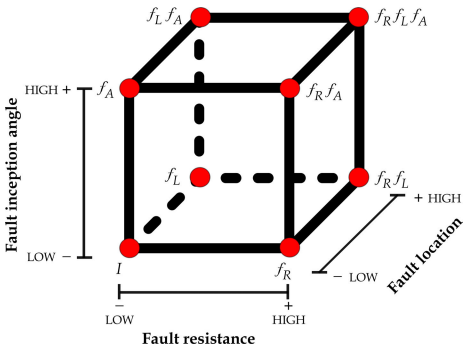


FIGURE 7. Geometric view of the 2^3 factorial design.

TABLE 4. Design matrix of the 2^3 design and measured response (in terms of operate time) for the three-factor scenario of Section V-A.

Run	Fault Resistance	Fault Location	Fault Inception Angle	Operate time [ms] of the replicates Y_1	Y_2	Y_3	Y_4	Totals [ms] $Y_{n\cdot}$	Averages [ms] $\bar{Y}_{n\cdot}$
1	−1	−1	−1	94.5	98.05	92.75	93.1	$I = 378.4$	
2	+1	−1	−1	323.1	323.1	323.9	323.6	$f_R = 1293.7$	
3	−1	+1	−1	323.3	324.5	323.8	324.2	$f_L = 1295.8$	
4	+1	+1	−1	324.5	323.3	322.1	323.0	$f_R f_L = 1292.9$	
5	−1	−1	+1	87.8	94.4	89.8	83.3	$f_A = 355.3$	
6	+1	−1	+1	323.6	323.7	323.6	322.0	$f_R f_A = 1292.9$	
7	−1	+1	+1	325.7	323.8	323.5	324.1	$f_L f_A = 1297.1$	
8	+1	+1	+1	322.3	323.4	323.2	323.8	$f_R f_L f_A = 1292.7$	
								$Y_{\cdot\cdot} = 8498.8$	$\bar{Y}_{\cdot\cdot} = 265.6$

of the corresponding letter. For example, f_R corresponds to the response obtained when the fault resistance is tested at its "+1" level and the fault location and fault inception angle are at their "−1" levels. When all the factors are at their "−1" levels, I is used to denote the corresponding response.

C. CONDUCT STEP 6 OF THE STAT-DOE

As said in Section I-B1, recording the operate time (and possibly extracting summary statistics such as mean or median values) represents the culmination of the state-of-the-art approaches adopting full factorial designs to test the distance protection performance. Yet, if these factorial designs were used in the broader context of the stat-DOE, much deeper insight could be gained at no cost by resorting to Step 6 of the stat-DOE, which envisages the statistical analysis of the experimental data, as shown hereafter.

1) ESTIMATING THE FACTORIAL EFFECTS

With a 2^3 (2^K) design, 7 ($2^K - 1$) factorial effects can be estimated: the three main effects of fault resistance, location and inception angle (F_R, F_L and F_A), and the four interactions $F_RF_L, F_RF_A, F_LF_A, F_RF_LF_A$. Consider, say, F_R, i.e., the main effect of the fault resistance, defined as the change in the operate time due to changing the fault resistance from its low to high level, averaged over the levels of the other factors. From Figure 7, it is evident that, for each replicate, 4 estimates of F_R are available: $f_r - I, f_Rf_L - f_L, f_Rf_A - f_A$ and $f_Rf_Lf_A - f_Lf_A$. By averaging them out, the main effect F_R is computed as:

$$F_R = \frac{f_r - I + f_Rf_L - f_L + f_Rf_A - f_A + f_Rf_Lf_A - f_Lf_A}{4R} \quad (4)$$

The numerator of (4) is called a "contrast", i.e., a linear combination of the responses at each factor levels' combination (the column $Y_{n\cdot}$ of Table 4) with either "+1" or "−1" signs. All the seven factorial effects are estimated by using the table of contrast coefficients reported in Table 5, which is obtained by extending the design matrix with columns that represent the interactions among factors (column 5 through 8), plus a column with the identity vector **I**, representing the average over the whole experiment. The signs of the interactions result from multiplying the signs of their respective factors (e.g., $F_RF_L = F_R \times F_L$).

TABLE 5. Table of contrast coefficients for the 2^3 design.

Factor level combination	Factorial effect F_R	F_L	F_A	F_RF_L	F_RF_A	F_LF_A	$F_RF_LF_A$	**I**
I	−1	−1	−1	+1	+1	+1	−1	+1
f_R	+1	−1	−1	−1	−1	+1	+1	+1
f_L	−1	+1	−1	−1	+1	−1	+1	+1
f_Rf_L	+1	+1	−1	+1	−1	−1	−1	+1
f_A	−1	−1	+1	+1	−1	−1	+1	+1
f_Rf_A	+1	−1	+1	−1	+1	−1	−1	+1
f_Lf_A	−1	+1	+1	−1	−1	+1	−1	+1
$f_Rf_Lf_A$	+1	+1	+1	+1	+1	+1	+1	+1

The general equation for computing the contrasts for any factorial effect for a 2^K design is:

$$\text{Contrast}_{U_1 \cdots U_K} = (u_1 \pm 1)(u_2 \pm 1) \cdots (u_K \pm 1) \quad (5)$$

where the sign in each set of parentheses is negative if the factor U_i is included in the effect, and positive otherwise. For example, the contrast for the two-factor interaction among fault location and fault inception angle F_LF_A is found by expanding the right hand side of:

$$\text{Contrast}_{F_LF_A} = (f_R + 1)(f_L - 1)(f_A - 1) \quad (6)$$

and by replacing 1 with I in the result. As the contrasts for the seven factorial effects are orthogonal, the 2^3 design (and all the 2^K designs) is orthogonal. After the effects' contrast are found, the corresponding effects can be obtained as:

$$U_1 \cdots U_K = \frac{2}{R2^K}(\text{Contrast}_{U_1 \cdots U_K}) \quad (7)$$

which are reported in the second column of Table 6. Examining the magnitude and direction of the factorial effects provides guidance on which factors are likely to be important in determining the observed response.

The overall variability in the data is given by the total sum of squares SS_{Tot}:

$$\text{SS}_{Tot} = \sum_{n=1}^{N}\sum_{r=1}^{R} Y_{nr}^2 - \frac{Y_{\cdot\cdot}^2}{NR} \quad (8)$$

where Y_{nr} is the response value of the nth run in the rth replicate. To quantify the amount of overall data variation due to each effect $U_1 \cdots U_K$, the sum of squares (SS) of the effects

is computed as:

$$\mathrm{SS}_{U_1 \cdots U_K} = \frac{1}{R2^K}(\mathrm{Contrast}_{U_1 \cdots U_K})^2 \tag{9}$$

From (8) and (9), the fundamental ANOVA equation can be adopted to decompose the overall data variability (SS_{Tot}) into its components [9]. For the 2^3 design of Table 4 it writes:

$$\begin{aligned}\mathrm{SS}_{Tot} = {} & \mathrm{SS}_{F_R} + \mathrm{SS}_{F_L} + \mathrm{SS}_{F_A} + \mathrm{SS}_{F_R F_L} + \mathrm{SS}_{F_R F_A} \\ & + \mathrm{SS}_{F_L F_A} + \mathrm{SS}_{F_R F_L F_A} + \mathrm{SS}_{Error}\end{aligned} \tag{10}$$

where the SS of the experimental/measurement error (SS_{Error}) is found by subtraction of the SS of each effect in (10) from the SS_{Tot}. Each SS is associated with a given number of degrees of freedom (df), i.e., the number of independent elements inside the SS. The ratio between an SS and its own df yields a mean sum of squares (MS). Thus, the statistical significance of each factorial effect can be assessed by resorting to a rigorous hypothesis testing [9], which, stated formally in the case of, say, F_R, reads:

$$\begin{aligned}H_0 &: F_R = 0 \\ H_1 &: F_R \neq 0\end{aligned} \tag{11}$$

where H_0 and H_1 are the null and alternative hypotheses, respectively. As the MS is formally a variance, the F_0 ratio of the MS of the effect to the MS of the error (MS_{Error}) is a test statistic for the hypothesis H_0 that the effect is not significant. For example, for the main effect F_R, the F_0 ratio:

$$\mathrm{F}_0 = \frac{\mathrm{MS}_{F_R}}{\mathrm{MS}_{Error}} = \frac{\mathrm{SS}_{F_R}/\mathrm{df}_{F_R}}{\mathrm{SS}_{Error}/\mathrm{df}_{Error}} \tag{12}$$

follows an F distribution $\mathrm{F}_{\mathrm{df}_{F_R},\mathrm{df}_{Error}}$, where $\mathrm{df}_{F_R} = L_1 - 1$ are numerator degrees of freedom and $\mathrm{df}_{Error} = L_1L_2L_3(R-1)$ are the denominator degrees of freedom. The null hypothesis H_0 would be rejected if the numerator MS_{F_R} is "significantly" greater than the denominator MS_{Error}, or, more formally, if F_0 is higher than the reference value $\mathrm{F}_{1,24}$. To facilitate the decision making on the effect statistical significance, the p-value is computed, which indicates the smallest level α at which the null hypothesis H_0 is rejected, or, in other terms, at which the corresponding effect would be significant. Values of $p < 0.05$ or $p < 0.01$ are often considered evidence that the corresponding effect is significant. For the scenario of Section V-A, Table 6 reports the ANOVA of the factorial effects. By looking at both the magnitudes of the factorial effects and their p-values, it turns out that the main effects of fault resistance and fault location (F_R and F_L) along with their second-order interaction (F_RF_L) dominate the process, whereas the other effects are negligible. Also, the positive values of F_R and F_L signalize that the effect of fault resistance and location is directly proportional to the operate time: faults with high resistance values and closer to the boundary make the operate time increase. On the other hand, an almost null value of F_A signalizes that changing the fault inception angle from 0° to 15° does not have any impact on the operate time.

It is noteworthy though that the main effects do not have much meaning when they are involved in significant interactions, as it is the case for F_R and F_L: in fact, the effect of F_R on the operate time depends on the level at which F_L is studied, and vice versa. Figure 8, which plots their second-order interactive effect, is the key to clarifying the situation: the effect of the fault resistance varies depending on the level at which the fault location is studied. In particular, on the one hand, the fault resistance has almost null effect if the fault happens at 85% (see the flat blue line): whatever the fault resistance, the operate time is always higher than 200 ms. On the other hand, the effect of the fault resistance is large if the fault happens at 84% (see the great slope of the red line): for values of fault resistance close to 1Ω, the relay operates in time, whereas the relay operates with times higher than the threshold of 200 ms if the fault resistance increases up to 3Ω. It is noteworthy that such interactive effect among fault resistance and fault location could not be detected if the OAT approach (see Section I-A1) would be adopted [10].

TABLE 6. ANOVA of the factorial effects. The asterisk signalizes the significant ones at the $\alpha = 1\%$ level of significance.

Effect	Effect estimate	df	SS	F_0	p-value
F_R	115.34	1	106450.75	28085.92	< 0.0001*
F_L	116.14	1	107909.16	28470.70	< 0.0001*
F_A	−1.42	1	16.32	4.3049	0.0489
F_RF_L	−116.26	1	108141.56	28532.02	< 0.0001*
F_RF_A	1.30	1	13.59	3.5843	0.0704
F_LF_A	1.56	1	19.61	5.1737	0.0322
$F_RF_LF_A$	−1.48	1	17.78	4.6899	0.0405

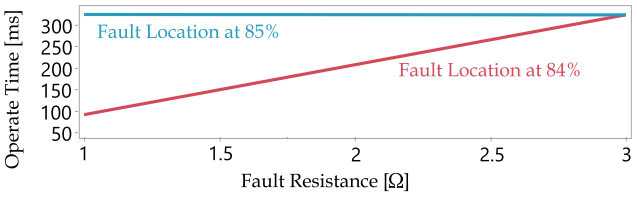


FIGURE 8. Plot of the second-order interaction between fault resistance and fault location for the three-factor scenario of Section V-A.

2) FITTING AN EMPIRICAL MODEL

The ANOVA of the factorial effects treats the factors as if they are qualitative. If the designed experiment involves at least one quantitative factor, an empirical model of the process can be built from the available data, and used as an interpolation equation to predict the response at unexplored combinations (i.e., at factors' levels other than those actually used in the experiment). The general approach to fit empirical models is the regression analysis [9].

The full regression model that can be fit to the experimental data from a 2^3 design (if the number of replicated runs is $R > 1$) is the first order model with interactions of the form:

$$\begin{aligned}\hat{Y} = {} & \hat{\beta}_0 + \hat{\beta}_1 U_1 + \hat{\beta}_2 U_2 + \hat{\beta}_3 U_3 + \hat{\beta}_{12} U_1 U_2 \\ & + \hat{\beta}_{13} U_1 U_3 + \hat{\beta}_{23} U_2 U_3 + \hat{\beta}_{123} U_1 U_2 U_3 + \epsilon\end{aligned} \tag{13}$$

In (13), the intercept $\hat{\beta}_0$ (representing the constant response not depending on the factors' levels), and the other coefficients $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$, $\hat{\beta}_{12}$, $\hat{\beta}_{13}$, $\hat{\beta}_{23}$ $\hat{\beta}_{123}$ are the unknown model coefficients to be estimated from the experimental data (e.g., via least-squares methods), whereas ϵ is a random error term accounting for the experimental/measurement error[1] in the system/process under study (here, the measurement error coming from the instrument transformer, amplifier, measurement unit of the distance relay).

3) STATISTICAL ANALYSIS OF THE EMPIRICAL MODEL

Once a regression model is fit to the experimental data, the ANOVA is again used to formally test the significance of the empirical model. For example, Table 7 shows the ANOVA for the full model (13). The overall variability of the response values (given by SS_{Total}, with df_{Total} equal to $N_T - 1$) is partitioned in two sources of variation: the data variability explained by the fitted model SS_{Model}, which reads

$$\mathrm{SS}_{Model} = \mathrm{SS}_{F_R} + \mathrm{SS}_{F_L} + \mathrm{SS}_{F_A} + \mathrm{SS}_{F_R F_L} + \mathrm{SS}_{F_R F_A} + \mathrm{SS}_{F_L F_A} + \mathrm{SS}_{F_R F_L F_A} \quad (14)$$

and the data variability SS_{Error} that remains unexplained. By performing the ANOVA of the full model (13), the very small p-value signalizes that at least one of its terms is significant.

TABLE 7. ANOVA of the full empirical model (13).

Source of variation	SS	df	MS	F_0	p-value
Model	322568.76	7	46081.3	12158.06	< 0.0001
Error	90.96	24	3.8		
Total	322659.72	31			

Table 8 reports detailed statistical information on the coefficients of the fitted model, whose 95% confidence interval lower and upper bounds are reported in the last two columns. It is noteworthy that, as a consequence of the ± 1 orthogonal coding, the least-square estimates of the $\hat{\beta}$ coefficients of the model (13) are one-half the corresponding effect estimates in Table 6. This property is useful to easily interpret the model coefficients and determine the relative size of the factor effects. In other words, the magnitudes of the model coefficients are directly comparable: being dimensionless, they measure the effect of changing each design factor over a one-unit interval. For example, the effects of fault resistance and fault location on the operate time are almost the same ($\beta_1 = 57.58$ and $\beta_2 = 58.07$, respectively), and both are as large as their interaction effect ($\beta_{12} = -58.13$). The standard error (SE) of the model coefficients, reported in column 3, is computed as

$$\mathrm{SE}(\hat{\beta}) = \sqrt{\frac{\mathrm{MS}_{Error}}{R2^K}} \quad (15)$$

[1]Note that if no replicated run is available, no internal estimate of the error ϵ would be produced.

and is the same for all the coefficient estimates due to the orthogonality of the 2^K design. Interestingly, there is no other 8-run design on the three-dimensional design space bounded by ± 1 (i.e., the cube of Figure 7) that makes the SE of the model coefficients smaller than $\sqrt{\mathrm{MS}_{Error}/R2^K}$.

TABLE 8. Statistical details of the coefficients of the full model (13).

Term	Estimate	SE	p-value	Low 95% CI	High 95% CI
β_0	265.59	0.34	< 0.0001	264.88	266.29
β_1	57.58	0.34	< 0.0001	56.97	58.39
β_2	58.07	0.34	< 0.0001	57.36	58.78
β_3	-0.71	0.34	0.0489	-1.42	-0.004
$\beta_1\beta_2$	-58.13	0.34	< 0.0001	-58.84	-57.423
$\beta_1\beta_3$	0.65	0.34	0.0704	-0.059	1.362
$\beta_2\beta_3$	0.78	0.34	0.0322	0.073	1.493
$\beta_1\beta_2\beta_3$	-0.75	0.34	0.0405	-1.456	-0.035

A summary of the goodness-of-fit of the full model obtained from the experimental data can be extracted from a set of metrics as shown in the second column of Table 9.

The coefficient of model determination R^2, which is defined as:

$$\mathrm{R}^2 = \frac{\mathrm{SS}_{Model}}{\mathrm{SS}_{Tot}}, \quad \mathrm{R}^2 \in [0, 1] \quad (16)$$

quantifies the proportion of the overall data variation being explained by the model (with $\mathrm{R}^2 = 1$ signalizing a perfect fit). Since R^2 tends to increase simply as more factors (although insignificant) are added to the model, the adjusted $\mathrm{R}^2_{\mathrm{Adj}}$ is sometimes preferred, and computed as:

$$\mathrm{R}^2_{\mathrm{Adj}} = 1 - \frac{\mathrm{SS}_{Error}/\mathrm{df}_{Error}}{\mathrm{SS}_{Tot}/\mathrm{df}_{Tot}} \quad (17)$$

As the $\mathrm{R}^2_{\mathrm{Adj}}$ is "adjusted" for the size of the model (i.e., the number of factors), it can actually decrease when non-significant terms are added to the model, and can be used to evaluate the impact of increasing or decreasing the number of model terms. To measure how well the model will predict new data, the prediction error sum of squares (PRESS):

$$\mathrm{PRESS} = \sum_{n=1}^{N} (\hat{Y}_{\sim n} - Y_n)^2 \quad (18)$$

is computed with "leave-one-out" cross validation. In particular, the nth data point is predicted ($\hat{Y}_{\sim n}$) with a model that includes all the observations except the nth one (Y_n) and computing the residual by comparing it with Y_n. After repeating this for all the observations, the sum of the squared prediction errors is obtained: a small PRESS indicates that the model is likely to be a good predictor. From the PRESS, the R^2_{Pred} can be computed as

$$\mathrm{R}^2_{\mathrm{Pred}} = 1 - \frac{\mathrm{PRESS}}{\mathrm{SS}_{Tot}} \quad (19)$$

which indicates the proportion of the variability in new data that the full model is expected to explain.

TABLE 9. Summary of fit of the full (13) and reduced (20) model.

	Full model	Reduced model
R^2	0.999718	0.99951
R^2_{Adj}	0.999636	0.999457
PRESS	161.7144	206.6951
R^2_{Pred}	0.9995	0.9995

4) REFINE THE EMPIRICAL MODEL

The full model (13) initially fit to the experimental data of Table 4 contains all the main effects as well as all the higher-order interactions. In general, it might be desirable to fit the lowest order polynomial that adequately describes the system/process under study.

According to the ANOVA of Table 8, the model (13) can be refined by screening the non-significant effects out of it (considering a significance level of $\alpha = 1\%$). Hence, the reduced model to predict the operate time would be:

$$\hat{Y} = 265.59 + 57.68\, U_1 + 58.07\, U_2 - 58.13\, U_1 U_2 \quad (20)$$

which, in addition to the intercept, contains only the coefficients $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_{12}$. This is an appeal of the "sparsity-of-effects" principle, an heuristics based on which most systems/processes turn to be dominated by some of the main effects and low-order interactions, with most high-order interactions being negligible [31]. The ANOVA of the reduced model (20) is given in Table 10: the MS_{Error} is now composed not only by a "pure" error component due to the presence of replicated runs, but also by a "lack-of-fit" component made of the SS of the factors dropped from the full model (13). The summary of the goodness-of-fit of the reduced model (20) is given in the third column of Table 9. Although four coefficients have been removed from the original model, no appreciable degradation of the goodness-of-fit metrics is visible. For example, by observing the values of the PRESS and R^2_{Pred}, the reduced model (20) owns almost the same predictive capability of the original full model (13).

TABLE 10. ANOVA of the reduced model (20).

Source of variation	SS	df	MS	F_0	p-value
Model	322501.47	3	107500	19020.51	< 0.0001
Error	158.25	28	5.651819		
Total	322659.72	31			

5) GRAPHICAL ANALYSIS OF THE DATA

After building a suitable empirical model, it is useful to resort to a graphical analysis to ease the interpretation of the results. For example, plots of the main effects and/or interactions can be built to inform about the magnitude and direction of change of the response due to the factors (e.g., see Figure 8 for the interactive effect among fault resistance and fault location). Additionally, if at least two factors are quantitative, three-dimensional response surfaces or two-dimensional contour plots can be produced to predict $\hat{Y}$ at various factors' combinations. For example, Figure 9 depicts the contour plot of the operate time in the design space spanned by the variation range of fault resistance $(1\Omega, 3\Omega)$ and fault location $(84\%, 85\%)$. In general, if the contour lines are curved, the interaction is important. Under the assumption that the relay is considered to fail if $t > 200$ ms, the contour plot allows to easily identify its operability region, which can be visualized in the bottom-left triangular area with the green contour lines. Such contour plot can be also read along each factor dimension. For instance, if the fault resistance is $> 2\Omega$, the operate time would be > 200 ms irrespective of the location at which the fault happens. Conversely, if the fault location is close to 85 % of the protected line, the relay operates with $t > 200$ ms even for values of fault resistance close to 1Ω. It is noteworthy that the operability region of Figure 9 can be mathematically described by the empirical model (20), which, under another perspective, can be seen as the *prediction rule* that defines the combinations of values of fault resistance and fault location leading to misoperation of the relay ($t > 200$ ms).

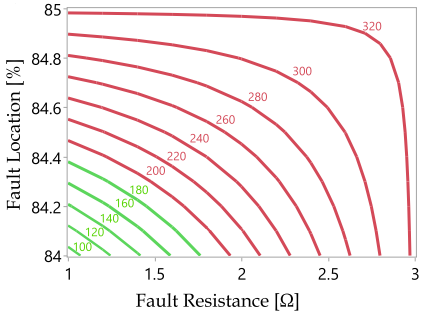


FIGURE 9. Contour plot of the operate time t in the design space spanned by the variation ranges of fault resistance and fault location. Based on the utility's requirements, failure of operation of the relay (red contour lines) is assumed for t >200 ms.

6) CHECK MODEL ADEQUACY AND AUGMENT THE DESIGN

The full regression model that can be fit from a 2^K design is the first order model with interactions (13). Such regression models work well even if the linearity assumption holds only very approximately, as the interactive terms can capture some sort of curvature in the response function. When the curvature in the response function cannot be adequately captured by such regression models, pure quadratic effects of the form $\beta_{ii} U_{ii}^2$ must be introduced, yielding a second-order response surface model. However, fitting a second-order model is not always necessary: in fact, checking the adequacy of the first-order model (13) can be first performed by augmenting the original N-run design with so-called "center points", i.e., N_C replicated runs at the center of the design space. In particular, by comparing the average response at the N points with that at the N_C points, a formal hypothesis test can check whether the center points lie on (or near) the plane passing through the N factorial points, or, in other words, whether the quadratic terms $\beta_{ii} U_{ii}^2$ (leading to a response function with

quadratic curvature) are statistically significant. If such test cannot exclude the presence of a quadratic curvature in the response function, a second-order model becomes necessary. In this case, the original N-run design is augmented with axial runs, leading to a so-called "central composite" design. For the scenario of Section V-A, by adding four more runs at the center point $\{2\Omega, 84.5\%, 7.5°\}$, the curvature test cannot exclude the presence of a quadratic curvature. Hence, a central composite design is created by augmenting the original design of Figure 7 with axial runs at the six points on the center of the six cube sides plus the center point. The second-order model fit to this composite design suggests that the fault resistance does have a non-negligible quadratic effect.

D. REMARKS ON THE FULL FACTORIAL DESIGN

In Section V-C, some of the opportunities offered by the statistical analysis of the experimental data (Step 6 of the stat-DOE) have been shown when the full factorial design is used, and evidence of the deeper insight achievable (with minimal effort) into the distance protection performance has been provided. However, as the number of factors increases, the number of tests required for a full factorial design grows exponentially ($N_T = R2^K$). For example, with 6 factors, an unreplicated 2^6 design (with $R = 1$) would require 64 runs: in this case, 63 df would be available in total, but only 6 of them are used to estimate the main effects, and only 15 for the two-factor interactions, with the other 42 df being used to estimate three-factor and higher-order interactions. In this context, the already mentioned sparsity-of-effects principle can justify the assumption of regarding high-order interactions as essentially "inert", and neglecting them. Hence, main effects and low-order interactions can be obtained by running only a *fraction* of the complete factorial experiment: this leads to the fractional factorial designs, which are described hereafter.

VI. FRACTIONAL FACTORIAL DESIGN

In this section, an experiment with the fractional factorial design is elaborated to test the distance protection performance. Such design is described in Section VI-B while being actualized in the four-factor scenario of Section VI-A. Later, Section VI-C provides final remarks on the fractional factorial design and introduces cornerstone principles of the stat-DOE.

A. DESCRIPTION OF THE SCENARIO

In this scenario, the three-factor scenario of Section V-A is extended with a fourth factor U_4, i.e., the fault type. This factor is studied at two levels, namely the single-phase-to-ground fault AN and the phase-to-phase fault ABC.

B. DESCRIPTION OF THE DESIGN

With 4 factors, a complete 2^K design would require 16 runs; imagine though that, because of time/money constraints, only 8 runs can be afforded, i.e., one half fraction of the 2^4 design.

TABLE 11. Design matrix of the 2^{4-1} fractional factorial design.

Run	Fault Resistance	Fault Location	Fault Inception Angle	Fault Type
1	−1	−1	−1	−1
2	+1	−1	−1	+1
3	−1	+1	−1	+1
4	+1	+1	−1	−1
5	−1	−1	+1	+1
6	+1	−1	+1	−1
7	−1	+1	+1	−1
8	+1	+1	+1	+1

TABLE 12. Estimates of the aliased effects for the 2^{4-1} design of the four-factor scenario of Section VI-A.

Alias structure	Estimate	p-value
$[F_R] \rightarrow F_R + F_L F_A F_T$	122.73	< 0.0001
$[F_L] \rightarrow F_L + F_R F_A F_T$	122.556	< 0.0001
$[F_A] \rightarrow F_A + F_R F_L F_T$	−5.994	< 0.0001
$[F_T] \rightarrow F_T + F_R F_L F_A$	5.194	< 0.0001
$[F_R F_L] \rightarrow F_R F_L + F_A F_T$	−122.681	< 0.0001
$[F_R F_A] \rightarrow F_R F_A + F_L F_T$	6.668	< 0.0001
$[F_R F_T] \rightarrow F_R F_T + F_L F_A$	−6.218	< 0.0001

Such fractional design is referred to as 2^{4-1} design, given that $2^4 \frac{1}{2} = 2^4 2^{-1}$ runs are employed. In general, a 2^{K-1} design can be formed by writing down a basic design consisting of the runs for a full 2^{K-1} design, and then appending the Kth factor by identifying its plus and minus levels with the plus and minus signs of the highest order interaction $U_1 U_2 \ldots U_K$. For example, the design matrix of a 2^{4-1} fractional factorial design is reported in Table 11: the second, third and fourth columns form the design matrix of a 2^3 full factorial design (see Table 4), and the fifth column contains the algebraic signs associated to the highest order interaction (i.e., the three-factor interaction $F_R F_L F_A$, see the last column of Table 5). Yet, as only 8 runs are available in the 2^{4-1} design, not all of the 15 effects can be estimated in a standalone manner. This leads to the so-called "confounding" of the effects; for example, as the three-factor interaction $F_R F_L F_A$ is used to accommodate the fourth factor "fault type", its main effect F_T cannot be distinguished from the $F_R F_L F_A$. In formal terms:

$$[F_T] \rightarrow F_T + F_R F_L F_A$$

where the $[\cdot]$ indicates that the main effect F_T is the result of the linear combination of F_T itself and $F_R F_L F_A$. Yet, under the sparsity-of-effects principle, since the three-factor interaction would be small enough to be ignored (as confirmed by the ANOVA of Table 8), it can be logically concluded that $[F_T]$ provides an estimate of the main effect F_T. Two or more effects possessing such property are called "aliases". The alias relationships for such a 2^{4-1} design are reported in the first column of Table 12. A design with such alias structure has a *resolution* IV, and it is indicated as 2^{4-1}_{IV}; in resolution IV designs, the main effects are aliased neither with other main effects nor with other two-factor interactions, but the two-factor interactions are aliased with each other. Resolution IV designs must contain at least $N = 2K$ runs,

so e.g., if 9 factors are to be studied, a resolution IV design requires at least 18 runs. In addition to resolution IV designs, of particular importance are resolution III and V designs.

In resolution III designs, e.g., 2^{4-1}_{III}, no main effect is aliased with other main effects, but main effects are aliased with two-factor interactions, and some two-factor interactions may be aliased with each other. It is possible to construct resolution III designs to investigate up to $K = N - 1$ factors in only N runs, where N is a power of 2. Particularly useful resolution III designs are e.g., to study 3 factors with 4 runs, 7 factors with 8 runs and up to 15 factors with 16 runs, as well as the so-called Plackett-Burman designs (with N being also multiple of 4) which allow to study e.g., 11, 19 and 23 factors with 12, 20 and 24 runs, respectively [32].

In resolution V designs, e.g., 2^{5-1}_{V}, no main effect or two-factor interaction is aliased with any other main effect or two-factor interaction, but two-factor interactions are aliased with three-factor interactions.

Intuitively, fractional designs with the highest possible resolution are desirable, as the higher the resolution, the less restrictive the assumptions required regarding which interactions are negligible. Hence, resolution V designs are very powerful, in that they allow unique estimation of all main effects and two-factor interactions, provided that all three-factor and higher-order interactions be negligible. However, as K grows, resolution V designs might require a higher number of runs N: for this reason, resolution IV designs are usually the optimal trade-off choice to start with, in that they avoid the confounding of main effects and two-factor interactions typical of resolution III designs, while preventing the larger sample size requirements of resolution V designs.

For the scenario of Section VI-A, Table 12 reports the effect estimates with the associated p-values. Looking at the magnitude of the alias chains, only $F_R + F_L F_A F_T$, $F_L + F_R F_A F_T$ and $F_R F_L + F_A F_T$ have a great estimate. Under the justified assumption that the three factor interactions $F_L F_A F_T$ and $F_L + F_R F_A F_T$ are negligible, it is logical to conclude that the fault resistance F_R and fault location F_L have dominant effects. On the other hand, the alias chains $F_A + F_R F_L F_T$ and $F_T + F_R F_L F_A$ have relative small estimates, leading to the conclusion that the main effects of fault inception angle F_A and fault type F_T are very small. Hence, if F_T has negligible effect, also the two-factor interaction $F_A F_T$ is negligible, as it is quite unlikely for a factor to have a negligible main effect but a large interactive effect; this leaves with the explanation that the large estimate of the alias chain $F_R F_L + F_A F_T$ is due only to the large effect of the second-order interaction $F_R F_L$. Such interpretations agree with the conclusions stemming from the analysis of the full 2^4 design (whose results are not shown here for brevity). Consequently, if only the effects of fault resistance and fault location are included, the reduced model obtained from such 2^{4-1} design is:

$$\hat{Y} = 263.09 + 61.37\, U_1 + 61.28\, U_2 - 61.34\, U_1 U_2 \quad (21)$$

The value of $R^2 = 0.996459$ confirms the goodness of the empirical model, which is very close to (20) obtained for the scenario of Section V-A: including the fault type as additional design factor does not impact the dominant importance of fault resistance and fault location on the operate time.

C. REMARKS ON THE FRACTIONAL FACTORIAL DESIGN

A major use of fractional factorials is for screening experiments, i.e., when experiments with many factors are considered and the goal is to identify those having large effects while retaining the experimental effort under control. The factors identified as important can then be investigated more thoroughly by exploiting the projectivity property of fractional factorial designs and/or by performing further experiments based on the sequential experimentation principle.

The projectivity property allows each fractional factorial design of resolution P to contain complete factorial designs (possibly replicated) in any subset of $P - 1$ factors. For example, for the four-factor scenario of Section VI-A, since both the fault inception angle and fault type have revealed to be inert, they can be dropped from consideration; hence, the original 2^{4-1}_{IV} can be projected into a full 2^2 design in the remaining active factors (fault resistance and fault location), and can be analyzed as in Section V-C.

The sequential experimentation principle suggests that it is usually best to start with a smaller design, which can be always augmented later if necessary. For instance, it is always possible to combine the runs of two or more fractional factorial designs and sequentially build a larger design to estimate specific factor effects and/or interactions of interest. Also, follow-up experiments can be conducted by e.g., performing one or more confirmation runs to verify the conclusions inferred from the original experiment, augmenting the design with strategies such as full/partial fold-over to de-alias specific effects of interest, adding some runs to modeling additional terms such as quadratic effects, and varying the factor ranges to move the analysis to another experimental region that is more likely to contain response values of interest. For more details, the reader is referred to e.g., Chapter 8 of [9].

VII. THE MODERN PHILOSOPHY OF CUSTOM DESIGNS

Section V and VI describe the full and fractional factorial designs and apply them to study two exemplary scenarios to test the distance protection performance. Yet, these (and other) classical designs might not fit the scenario under test when the features of the latter do not perfectly match the requirements typical of classical designs. For example, the available resources might pose a burden which makes the operator's requirements (e.g., in terms of maximum number of affordable tests) incompatible with the size of the experiment prescribed by classical designs; also, constraints might exist on the design space (in the form of disallowed combinations of factors or linear/nonlinear constraints among them), which cannot be handled by any of the classical

designs. When none of the classical designs fit the bill, it is common practice to create a "custom" design that is optimal for the problem at hand and can accommodate operator specific requirements and physical constraints among factors: this is the philosophy behind the modern designs, which are discussed in Section VII-A. Modern designs have been establishing as best-practice in the stat-DOE community: in fact, if the problem under study turns out to be a "standard" one, the generated custom design will correspond to one of the classical designs (e.g., Section VII-B), otherwise it will be the *optimal* one to fit the configuration of the scenario under test.

A. CUSTOM DESIGNS AND OPTIMALITY

In general, a custom design allows e.g., to elaborate an experiment with different types of factors (e.g., continuous, discrete numeric, qualitative, constant, uncontrolled) with any number of levels, to specify a non-regular (i.e., constrained) design space, to prioritize the estimability of the effects of factors of primary interest over those for which estimability is only desirable but not necessary, and to select the number of runs that matches the budget for the given experiment.

To construct the appropriate design that meets the specific requirements of the scenario under test, custom designs employ an optimality criterion, i.e., a criterion for optimally selecting the design points whereby tests should be conducted. After the early works on the design optimality theory [33] and the first algorithms to build optimal designs [34], most of the state-of-the-art methods for optimal designs currently use a "coordinate-exchange" algorithm [35], in which individual design coordinates are systematically searched to find the optimal settings that maximize a given optimality criterion. The *D*-, *G*- and *I*-optimality criteria are the most widely used.

The *D*-optimal design minimizes the SE of the model parameters $\hat{\beta}$ (15), or, in other words, the region defined by their joint confidence intervals. As a smaller confidence region means more precise estimates of the $\hat{\beta}$'s, a *D*-optimal criterion is suitable when the experiment goal is mainly the estimation of the factors' effects and testing of their significance, and the identification of the active factors in screening experiments. The *G*-optimal design minimizes the maximum variance of the prediction over the design region, and the *I*-optimal design minimizes its average value (i.e., the ratio of the prediction variance integrated over the design space to the area of it). As a smaller prediction variance means more precise estimates of the response prediction, *G*- and *I*-optimal designs are suitable when the accurate prediction of the response takes precedence over the precise estimation of the factors' effects, e.g., to predict the response variable at untried factors' combinations, to determine optimum operating conditions and regions in the design space where the response falls within an acceptable range. For more details on these and other optimality criteria, the reader is referred to e.g., Chapter 6 of [9].

B. COMPARING OAT AND FACTORIAL EXPERIMENTS

As said in Section I-A1, some literature works follow approaches for distance protection performance testing based on OAT experiments, which showcase at least two methodological pitfalls: lack of ability to capture interactions among factors and limited capacity to cope with the curse of dimensionality. After having introduced the concept of design optimality in Section VII-A, the performance of OAT designs can now be evaluated also under further viewpoints when compared to classical/custom designs.

Without loss of generality, consider the model $Y = f(U_1, U_2)$ with two quantitative factors U_1 and U_2, whose variation range is between -1 and $+1$ (leading to a squared design space of area equal to 4), and imagine that only $N_T = 8$ tests can be afforded. Imagine to design two types of experiments, a custom design and an OAT design. Their cross-comparison is done in terms of ability to fit (with the same N_T) the first-order model without interactions:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 U_1 + \hat{\beta}_1 U_2 \tag{22}$$

due to the impossibility of the OAT design to capture two-factor interactions and quadratic effects.

Figure 10a represents the optimal custom design for such problem, which turns out to be a replicated 2^2 full factorial design. Thus, the latter is taken, for the sake of the comparison, as the reference design for the family of classical/custom designs. On the other hand, Figure 10b represents a typical OAT design: assuming the nominal point $\{0, 0\}$, the factor U_1 is first varied four times over its variation range with $U_2 = 0$, and then U_2 is similarly varied by keeping $U_1 = 0$.

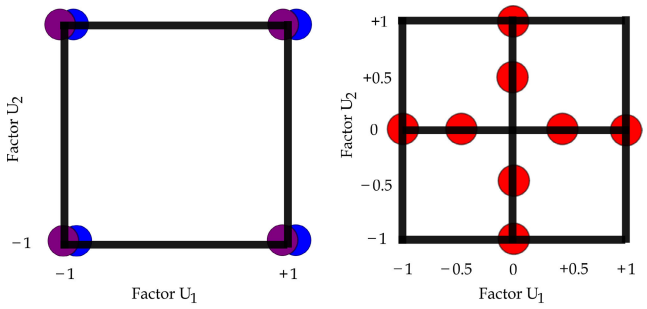


FIGURE 10. (a) Custom design (replicated 2^2 full factorial). (b) OAT design.

Various diagnostic metrics (some of which are shown hereafter) can be used to assess the performance of both designs to fit the model (22) even before running the experiments.

Figure 11 shows the profile of the Unscaled Prediction Variance (UPV) over the factors' ranges:

$$\text{UPV} = \frac{Var[\hat{Y}(U_1, U_2)]}{\sigma^2} \tag{23}$$

where $Var[\hat{Y}(U_1, U_2)]$ is the variance of the predicted response at every design point (U_1, U_2), and σ^2 is the error

variance (corresponding to MS_{Error}). The prediction variance $Var[\hat{Y}(U_1, U_2)]$ depends on the error variance, which is unknown before running the experiment. Yet, the UPV, i.e., the ratio of the $Var[\hat{Y}(U_1, U_2)]$ to the error variance, is not a function of σ^2, and hence depends only on the design type and on the factor settings: consequently, the UPV can be calculated even before acquiring the data, and provides indication of where the predicted response will have more or less variability. Obviously, low values of UPV are desirable. As shown in Figure 11, the UPV of the custom 2^2 design is constantly way lower than the OAT design. The two designs can be compared also based on the max(UPV), i.e., the maximum value of the UPV: for the custom 2^2 design, max(UPV)—which is the least desirable from the design viewpoint—is more than two times smaller than the OAT design.

Figure 12 depicts, for both designs, the Fraction of Design Space (FDS) plot, which shows the UPV on the vertical axis and the proportion of the design space (ranging from 0% to 100%) on the horizontal axis. The FDS plot gives indication of how the UPV is distributed throughout the whole design space: for a point (x, y) falling on the curve, the value x is the proportion of the design space with UPV $\leq y$. An ideal FDS plot would be flat with a small value of UPV. From Figure 12, it can be derived that the custom 2^2 design has uniformly smaller UPV than the OAT design. Also, the two black dashed crosshairs (centered at 50% of the design space) indicate that the UPV of the custom 2^2 design will be at most 0.217 over a region that covers 50% of the design region, whereas at such FDS the UPV of the OAT already reaches almost 0.4.

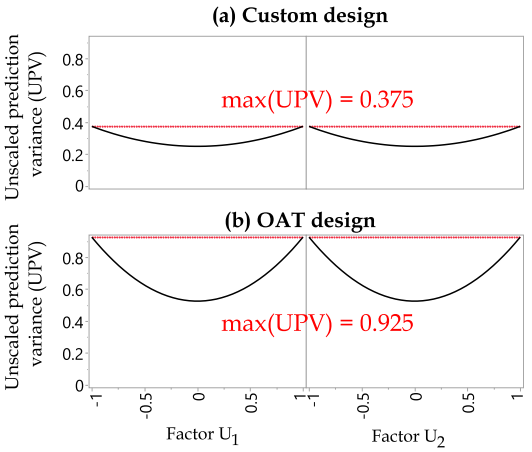


FIGURE 11. UPV profile of the (a) custom 2^2 design and (b) OAT design. The horizontal red dotted line indicates the maximum value.

Table 13 shows the values of the D-, G- and I-optimality criteria for the custom 2^2 and OAT designs. The custom 2^2 design turns to be D-, G- and I-optimal for fitting the first-order model (22), with relative efficiency values almost always higher than 2 times when compared to the OAT design. Interestingly, the custom 2^2 design outperforms the OAT design although the latter covers the variation range of the factors U_1 and U_2 with more levels than the former (see

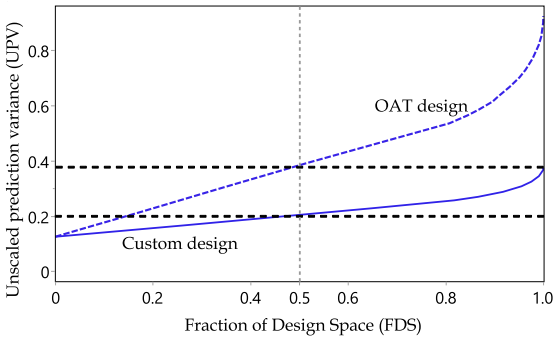


FIGURE 12. FDS plot for the custom 2^2 design and OAT design.

TABLE 13. Efficiency of the custom 2^2 design and OAT design to fit the first order model (22) in terms of optimality criteria.

	Custom 2^2 design	OAT design	Relative efficiency
D-optimality	100%	46.05%	2.172
G-optimality	100%	40.54%	2.467
I-optimality	100%	53.19%	1.880

the allocation of the eight design points in the design spaces of Figure 10). Also, the presence of replicated runs allows the custom 2^2 design to provide an internal estimation of the experimental error, unlike the OAT design.

Finally, not only the custom 2^2 design is D-, G- and I-optimal to fit as well the first-order model with interactions:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 U_1 + \hat{\beta}_2 U_2 + \hat{\beta}_{12} U_1 U_2 \tag{24}$$

but also, by adding just one more run, it boils down into an unreplicated 3^2 factorial design with the two factors at the three levels $\{-1, 0, 1\}$. Such 3^2 factorial design would allow fitting even a second-order regression model of the form:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 U_1 + \hat{\beta}_2 U_2 + \hat{\beta}_{12} U_1 U_2 + \hat{\beta}_{11} U_1^2 + \hat{\beta}_{22} U_2^2 \tag{25}$$

VIII. STAT-DOE APPLIED FOR PERFORMANCE TESTING

In this section, the stat-DOE is adopted to conduct a realistic performance testing of a commercial relay. A guideline of the stat-DOE application is provided by actualizing all the steps of Table 2 in such scenario, and thoroughly explaining the rationale behind each of them. The derived results allow not only to test and characterize the performance of the device under test, but also to sketch out lines of refinements and recommendations for all the stakeholders.

Although the performance of the distance protection only within Zone 1 is tested, testing the performance within other protection zones or of other protection functions (e.g. back-up protection, tele-protection schemes) can be straightforwardly conducted by replicating the different steps presented next.

A. STEP 1 – SPECIFY PROBLEM AND OBJECTIVES

An electricity utility has to conduct an acceptance test of a Schneider P543 relay before its field implementation.

Without loss of generality, the performance of the relay only within Zone 1 is considered. The technical manual from the relay's manufacturer [29] reports that the operate times have an average value of 30 to 35 ms in Zone 1. Nonetheless, the utility is aware that protection relays, when deployed in the field, rarely experience such small operate times under real operating conditions (e.g., faults with resistance higher than 0Ω); hence, the utility can accept that the relay operates *correctly* as long as $t < 200$ ms, to account for the safety margin as explained in Section IV-C. To conduct such performance test, the test set-up of Section IV-B is employed, and an experiment is laid out with the stat-DOE strategy. Given the available time/money resources, it is assumed that only 100 tests can be afforded.

B. STEP 2 – CHOOSE THE RESPONSE VARIABLE

The response variable measured to test the distance protection performance is the relay's operate time t, defined as the duration of the time interval between the fault inception and the receiving of the trip signal, excluding the operation time of the circuit breaker.

C. STEP 3 – CHOOSE FACTORS, LEVELS AND RANGES

According to the IEC 60255-121:2014 standard [16], four are the factors influencing the distance protection performance, namely fault location, fault inception angle, SIR and fault type; these factors should be tested by considering the levels reported in Figure 2. Yet, it is assumed that the protection engineer is interested in considering also the effect of the fault resistance (similarly to other literature works e.g., [4], [15], [17]), so to reflect more realistic operating conditions. Hence, five are the design factors subject of the experiment; these, their levels, and the held-constant factors are reported in Table 14. The rationale behind the selection of the design factors and the associated levels is discussed next.

TABLE 14. Design factors, their levels, and held-constant factors used for the experiment designed for the performance testing in Section VIII.

Design factors	Factor levels	Constant factors
Fault resistance [Ω]	0.001, 3, 6, 9	Service settings
Fault location [%]	5, 25, 50, 75, 85	based on
Fault inception angle [°]	0, 15, 30	Section IV-B
SIR [-]	0.2, 2, 5	
Fault type [-]	BN, AC, ABC, ACN	

- Faults of type BN, AC, ABC, ACN are injected, as prescribed in the IEC 60255-121:2014 standard [16].
- As the IEC 60255-121:2014 standard [16] considers only the positive side of the sine wave of the fault inception angle, and the worst situation occurs at 0° (whereby the fault current magnitude is the highest), faults with inception angles equal to 0°, 15° and 30° are applied.
- As the interest is the performance of the protection function in Zone 1 (set to cover 85% of the protected line), faults at locations corresponding to 5%, 25%, 50%, 75% and 85% are applied, with more focus on locations close to the boundary between Zone 1 and Zone 2.
- Long, medium and short lines are investigated, as indicated in the IEC 60255-121:2014 standard [16]. Given that the SIR is the preferred method to classify the electrical length of a line for the purpose of applying protective relays, SIR values feasible for long, medium and short lines are chosen, i.e., 0.2, 2 and 5, respectively [8].
- The IEC 60255-121:2014 standard [16] does not consider the fault resistance as design factor, and suggests to set it at 0Ω (or to the smallest possible value if numerical limitation arise). However, faults with resistance higher than 0Ω are common in reality. Hence, a variation range up to 9Ω is considered based on [36], and fault resistance values of 0.001Ω, 3Ω, 6Ω, 9Ω are chosen.

D. STEP 4 – SELECT THE DESIGN

From Step 3, a mixed-level, hybrid design space is originated: the five factors have different number of levels, and they are either quantitative (fault resistance, fault location, fault inception angle and SIR) or qualitative (fault type). It is noteworthy that, to accommodate such design space, the IEC 60255-121:2014 standard [16] suggests to adopt a full factorial design (see Figure 2), which would require $4 \times 5 \times 3 \times 3 \times 4 = 720$ runs, for a total of 2880 tests if each fault inception is replicated 4 times. As only 5 tests can be conducted every minute (see Section IV), this would lead to a total experimental time of more than 9 hours. Since, in principle, protection zones other than Zone 1 would be subject to the performance testing, and different scenarios (e.g., evolving and simultaneous faults) or even other ancillary functions (e.g., tele-protection) could be tested [15], conducting tests for several days would be inevitable. Such experimental size might be incompatible with the utility's time/money resources (which here are assumed to limit the number of affordable tests to 100), and it already motivates resorting to a custom design. In addition to this, Step 3 implies that the set of design factors include both fault type and fault resistance; hence, the full factorial design would be composed of runs without practical meaning in the reality, e.g., phase-to-phase faults (AC, ABC) with non-null resistance. Mathematically, this yields a constrained design space, which neither the full factorial design nor other off-the-shelf classical designs can handle. On the other hand, custom designs are able to easily accommodate such feature (see Section VII). Hence, for all these motivations, a custom optimal design is adopted.

In detail, the custom design is generated with JMP [26] by:

(i) choosing the *D*-optimality criterion, assuming that the main focus is estimating the effects of the five factors and identifying their statistical significance;
(ii) imposing that the fault resistance of AC and ABC faults shall be 0.001Ω;
(iii) choosing the effects of primary interest (i.e., whose estimability is considered as "necessary") to be all

TABLE 15. Excerpt of the design matrix generated with the custom design.

Run	Fault Resistance [Ω]	Fault Location [%]	Fault Inception Angle [°]	Fault Type [-]	SIR [-]
1	6	5	0	BN	2
2	6	25	0	ACN	2
3	0.001	25	30	ABC	0.2
4	9	25	0	BN	2
5	9	85	30	BN	2
6	6	25	0	BN	0.2
7	0.001	85	0	ABC	5
8	3	25	15	ACN	0.2
9	0.001	25	30	ABN	0.2
10	0.001	85	0	AC	0.2

the main effects, all the second-order interactions (e.g., $F_R F_A$), and all the quadratic single-factor effects (e.g., F_A^2); on the other hand, interactions of higher order than the second (e.g., $F_R F_L F_A$) as well as quadratic multi-factor effects (e.g., $F_R^2 F_A$) are considered either negligible or not physically meaningful and hence excluded from the definition of the initial empirical model;

(iv) setting the constraint on the number of tests to 100.

The first ten runs of the generated 100-test custom design used for Step 5 to 7 are reported in Table 15, from which it can be noted that phase-to-phase fault types are associated only with null values of fault resistance (0.001 Ω).

Table 16 compares some of the diagnostic metrics of Section VII-B for both the 100-test custom design, and the full factorial one that would originate if the IEC 60255-121:2014 standard [16] is followed. As the full factorial design cannot deal with constraints on the design space (here, the presence of disallowed combinations of factors), the custom design evaluated in Table 16 is generated *without* the constraint (ii). From Table 16, it emerges that, with only 100 tests, the custom design has larger values of the optimality criteria than the full factorial, as well as a satisfactory average UPV. Also, as far as the minimum number of required tests N_{MIN} is concerned, the custom design is way more "economical".

TABLE 16. Comparison of the custom design (neglecting the constraint (ii)) and the full factorial design as per IEC 60255-121:2014 standard [16].

	Custom design without constraint (ii)	Full factorial design
D-optimality	57.09%	50.53%
G-optimality	71.29%	53.68%
Average UPV	0.218	0.026
N_{MIN}	26	1920

E. STEP 5 – CONDUCT THE EXPERIMENT

The 100-test design matrix generated by the custom design including the constraint (ii) is used to conduct the physical experiment with the test set-up and test procedure explained in Section IV-B. The experimental results in terms of operate time are recorded and further elaborated in Step 6.

F. STEP 6 – STATISTICALLY ANALYSE THE DATA

Figure 13 summarizes the experimental results. The histogram shows that the distribution of the operate time values is bimodal, with a neat separation of tests whereby the relay operates "correctly" and "incorrectly" (green and red bars, respectively), according to the operability threshold of 200 ms accepted by the utility. The cumulative distribution plot shows that $t < 200$ ms in almost 80% of the tests.

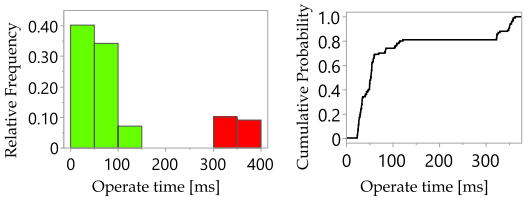


FIGURE 13. Histogram (left) and cumulative distribution plot (right) of the values of operate time for the scenario of Section VIII.

After fitting a quadratic model to the experimental data ($R^2 = 0.89$), the ANOVA shows that the significant effects are fault location, fault resistance and SIR. As found already for the scenarios studied in Section V and VI, the second-order interaction among fault location and fault resistance has a great effect. In addition, faults simulated for short lines lead to greater operate times than for medium and long lines, due to the greater instrument transformer measurement error because of the higher ratio between the impedance in front of and behind the relay. It is worth noting that these results match those obtained by using the full factorial design with 2880 tests suggested by the IEC 60255-121:2014 standard [16]; in other words, the same information can be obtained with an efficiency almost 30 times higher.

Also, a set of follow-up runs are used to validate the experimental conclusions. In particular, 50 new experimental points are randomly sampled from the design space defined by the factor ranges of Table 14, and the empirical model fit with the original 100-test custom design is used to predict the operate time at these untried factors' combinations. As seen in Figure 14, the regression line shows perfect agreement among the predictions and the experimental results.

G. STEP 7 – CONCLUSIONS AND RECOMMENDATIONS

Finally, the experimental results allow extracting conclusions useful not only for the utility, but also for other stakeholders such as relay's manufacturers and standardization bodies.

The empirical model built in Step 6 can be used as a prediction rule to define the pass/fail regions of the relay, or, in other words, the combinations of values of fault location, fault resistance and SIR—i.e., the three factors having significant effects—leading to misoperation of the relay, which can be of practical use from the utility viewpoint. Knowledge of the operability boundary as well as

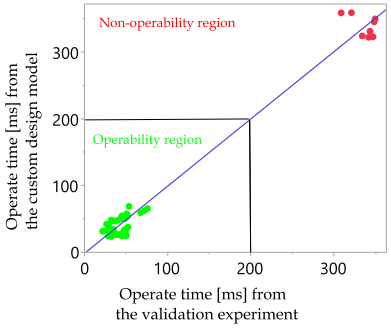


FIGURE 14. Regression between the values of operate time coming from the 50-run validation experiment and those obtained with the empirical model fitted on the experimental data from the custom design.

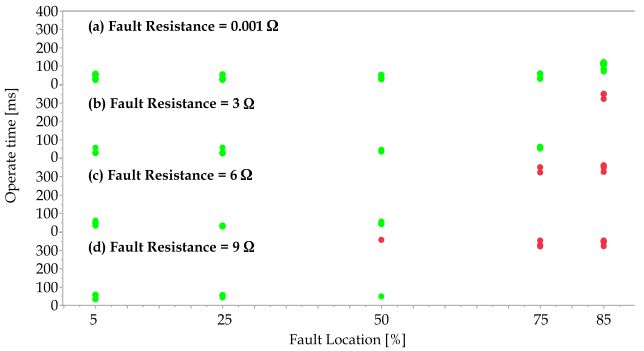


FIGURE 16. Operate time for Zone 1 for different fault resistance values.

identification of the inconsequential factors would support the actualization of Step 3 in further experiments (e.g., to conduct performance tests of other vendors), in terms of choice of the factors to consider and variation ranges to investigate.

Plots involving the three significant factors can be produced. For example, Figure 15 shows a modified version of the so-called "SIR diagrams" suggested by the IEC 60255-121:2014 standard [16]. In the latter, the SIR diagrams separately report the average, minimum and maximum operate times (y-axis) at the different fault locations (x-axis), for various SIRs. Unlike these, the SIR diagrams suggested in Figure 15 report all the recorded operate times at each fault location, and their classification is made based on the operability threshold $t = 200$ ms. In particular, the SIR diagrams of the average operate times are not shown, as this information would not be of much significance for cases where, for a given fault location, some operate times are greater and others are lower than the threshold (see e.g., Figure 15a for the fault location equal to 85%). Also, it is convenient to produce similar diagrams for different values of fault resistance, as done in Figure 16: for null values of fault resistances, the relay always operates within the threshold also at locations close to the boundary (i.e., 85%), whereas, for greater fault resistances, the relay degrades its performance and operates above the threshold even for faults occurring way before the boundary.

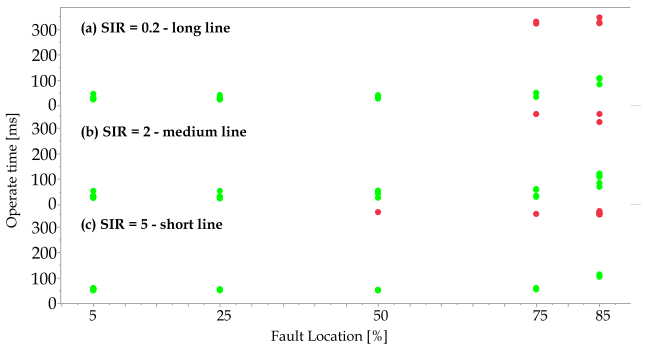


FIGURE 15. Operate times for long (a), medium (b) and short lines (c).

Overall, some lines of refinement and recommendations for all stakeholders can be extracted, as discussed hereafter.

- The reporting of the test results recommended by the IEC 60255-121:2014 standard [16] can be improved. In particular, as long as the SIR diagrams are concerned, *all* the operate times collected at *all* the tests may be reported in one single diagram, as shown in Figure 15, to avoid meaningless average values in case the variability of the operate times is large. Moreover, if other factors are considered, similar diagrams would help to comprehensively characterize the performance of the relay (e.g., Figure 16a–d for different values of fault resistance).
- The IEC 60255-121:2014 standard [16] suggests to conduct the performance testing by holding the fault resistance constant at 0Ω, supposedly to account for the relay manufacturer's viewpoint. The experiments conducted in Section V, VI and VIII show that, as expected, the fault resistance has always a great effect (both alone and in combination with the fault location), and thus cannot be neglected for a thorough assessment of the relay performance in realistic operating conditions. With the stat-DOE, it is proven how the utilities can effectively investigate the effect of the fault resistance together with other factors. In particular, twofold is the benefit given by investigating the effect of fault resistance on the relay performance over a wider range, e.g., from 0 to 9 Ω, as reported in Figure 16. In fact, on the one hand, the tests for which faults with null resistance are applied (e.g., Figure 16a) allow the utilities to check the compliance of the manufacturer's claims with the IEC 60255-121:2014 standard [16] (as reported in the relay technical manual). On the other hand, the tests for which faults with resistance values higher than 0Ω are applied (e.g., Figure 16b-c-d), help the utilities know the values of fault resistance above which misoperation of the relay is recorded in realistic conditions; this helps identifying further directions of improvement, such as which kinds of

additional protection functions the distance protection should be equipped with to deal with high resistance faults (e.g., directional over-current function).

- The IEC 60255-121:2014 standard [16] provides no indication on how to define performance criteria for distance relays, which are left to the utilities [28]. Although it is difficult to univocally define such criteria, the stat-DOE supports the utilities in elaborating robust pass/fail criteria based on their own requirements. For example, if the utility accepts the distance relay to operate "correctly" in Zone 1 as long as $t < 200$ ms (considering the safety margin before the Zone 2 operation), the empirical model fit on the data of the experiment designed with the stat-DOE can be used to extract the relay's operability boundary, which, in turn, allows inferring the combinations of the design factors' values that make the relay operate "incorrectly".

IX. CONCLUSION

Power system protection testing is vital to investigate the performance of protection systems and equipment before their field implementation. Yet, time/money resource constraints can prevent a thorough testing activity. This paper shows how the stat-DOE can be beneficial for the testing activity (especially in a resource-saving context) by focusing on the distance protection performance testing.

The state-of-the-art experimental strategies can greatly profit from the statistical techniques envisaged by the stat-DOE (e.g., statistical test of hypothesis, ANOVA, multivariate regression, etc.). Statistically designing the experiments via the stat-DOE allows to perform a thorough and broader-than-standard evaluation of the relay performance under realistic operating conditions without "arbitrarily" selecting *which* and *how many* tests to conduct. Operator-specific requirements (such as the maximum number of affordable tests) or physical constraints among factors can be effectively handled by modern custom designs that, compared to the state-of-the-art experimental strategies, are more efficient in producing the same amount of information at much less price in terms of number of tests to conduct. For instance, the custom design adopted to test the performance of a commercial distance relay shows an efficiency almost 30 times higher than the design used by the IEC 60255-121:2014 standard [16]. All the actors of the standardization chain can benefit from the recommendations and potential lines of refinement extracted from this work, in terms of how to e.g., study the effect of factors others than those suggested by the IEC 60255-121:2014 standard [16] (e.g., the fault resistance), extract pass/fail criteria for defining acceptance tests based on the utility's requirements, and publish the test results.

This work provides a replicable guideline to test the performance of other protection functions (e.g., transformer differential) and ancillary functions (e.g., tele-protection) or even in different scenarios (such as evolving or simultaneous faults [15]). Other use cases might have their own features in terms of number and types of factors, type of response to measure, operator-specific requirements, etc. Nonetheless, although its instantiation may differ, the stat-DOE can represent an enabler towards a common basis for guaranteeing replicability, robustness and objectivity of the testing activity. More broadly, it can ultimately be a precious tool towards the refinement of interoperability testing procedures [19], [37].

REFERENCES

[1] M. A. Rios, D. S. Kirschen, D. Jayaweera, D. P. Nedic, and R. N. Allan, "Value of security: Modeling time-dependent phenomena and weather conditions," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 543–548, Aug. 2002.

[2] *IEEE Guide for Power System Protection Testing*, Standard IEEE C37.233-2023, Inst. Electr. Electron. Engineers, 2023.

[3] N. I. Santoso and J. Y. Avins, "Real-time software testing for microprocessor-based protective relays," *IEEE Trans. Power Del.*, vol. 9, no. 3, pp. 1359–1367, Jul. 1994.

[4] M. Kezunovic, Y. Q. Xia, Y. Guo, C. W. Fromen, and D. R. Sevcik, "Distance relay application testing using a digital simulator," *IEEE Trans. Power Del.*, vol. 12, no. 1, pp. 72–82, Jan. 1997.

[5] J. R. Camarillo-Pe naranda, M. Aredes, and G. Ramos, "Hardware-in-the-loop testing of a distance protection relay," *IEEE Trans. Ind. Appl.*, vol. 57, no. 3, pp. 2326–2331, Jun. 2021.

[6] T. Penthong, M. Ginocchi, A. Ahmadifar, F. Ponci, and A. Monti, "IEC 61850-based protection scheme for multiple feeder faults and hardware-in-the-loop platform for interoperability testing," *IEEE Access*, vol. 11, pp. 65181–65196, 2023.

[7] T. Bardou, A. Bonetti, V. Leitloff, and M. Yalla. *Protection Relay Functional Standards for All*, Standard IEC 60255, Electr. Tester Mag., 2021, pp. 63–69.

[8] *IEEE Guide for Protective Relay Application Transmiss. Lines*, Standard IEEE Standard C37.113-2015 (Revision IEEE Std C37.113-1999), Inst. Electr. Electron. Engineers, 2015.

[9] D. C. Montgomery, *Design and Analysis of Experiments*, 8th ed. New York, NY, USA: Wiley, 2012.

[10] M. Ginocchi, F. Ponci, and A. Monti, "Sensitivity analysis and power systems: Can we bridge the gap? A review and a guide to getting started," *Energies*, vol. 14, no. 24, p. 8274, Dec. 2021.

[11] K. Jia, T. Bi, W. Li, and Q. Yang, "Ground fault distance protection for paralleled transmission lines," *IEEE Trans. Ind. Appl.*, vol. 51, no. 6, pp. 5228–5236, Nov. 2015.

[12] N. George and O. D. Naidu, "Distance protection issues with renewable power generators and possible solutions," in *Proc. 16th Int. Conf. Develop. Power Syst. Protection (DPSP)*, vol. 2022, Mar. 2022, pp. 373–378.

[13] A. Novikov, J. J. de Chavez, and M. Popov, "Performance assessment of distance protection in systems with high penetration of PVs," in *Proc. IEEE PowerTech*, Milan, Italy, Jun. 2019, pp. 1–6.

[14] T. S. Sidhu, R. K. Varma, P. K. Gangadharan, F. A. Albasri, and G. R. Ortiz, "Performance of distance relays on shunt-FACTS compensated transmission lines," *IEEE Trans. Power Del.*, vol. 20, no. 3, pp. 1837–1845, Jul. 2005.

[15] T. Penthong, M. Ginocchi, F. Ponci, and A. Monti, "Testing methodology for performance evaluation of distance protection relays for transmission systems," in *Proc. IEEE PowerTech*, Belgrade, Serbia, Jun. 2023, pp. 1–6.

[16] *Measuring Relays and Protection Equipment—Part 121: Functional Requirements for Distance Protection*, International Electrotechnical Commission, Geneva, Switzerland, 2014.

[17] A. L. P. De Oliveira and P. M. Da Silveira, "Evaluation of distance protection performance applied on series compensated transmission lines using real time digital simulation," in *Proc. IEEE/PES Transmiss. Distribution Conf. Expo., Latin Amer.*, Aug. 2006, pp. 1–6.

[18] R. A. Fisher, *The Design of Experiments*. Edinburgh, Scotland: Oliver and Boyd, 1935.

[19] I. Papaioannou, S. Tarantola, A. Lucas, E. Kotsakis, A. Marinopoulos, M. Ginocchi, M. Olariaga Guardiola, and M. Masera, *Smart Grid Interoperability Testing Methodology*. Luxembourg, U.K.: Publications Office of the European Union, 2018. [Online]. Available: https://op.europa.eu/en/publication-detail/-/publication/f307f230-f45d-11e8-9982-01aa75ed71a1

[20] N. Andreadou, A. Lucas, S. Tarantola, and I. Poursanidis, "Design of experiments in the methodology for interoperability testing: Evaluating AMI message exchange," *Appl. Sci.*, vol. 9, no. 6, p. 1221, Mar. 2019.
[21] M. Ginocchi, A. Ahmadifar, F. Ponci, and A. Monti, "Application of a smart grid interoperability testing methodology in a real-time hardware-in-the-loop testing environment," *Energies*, vol. 13, no. 7, p. 1648, Apr. 2020.
[22] J. F. Box and R. A. Fisher, *The Life of a Scientist*, 8th ed. New York, NY, USA: Wiley, 1978.
[23] G. E. P. Box and K. B. Wilson, "On the experimental attainment of optimum conditions," *J. Roy. Stat. Soc., Ser. B*, vol. 13, no. 1, pp. 1–45, Jan. 1951.
[24] *Minitab®*, Minitab, State College, PA, USA, 2023.
[25] *Design Expert®*, Design Expert, Minneapolis, MN, USA, 2023.
[26] *JMP® Version 17*, SAS Institutem Cary, NC, USA, 2023.
[27] G. Ziegler, *Numerical Distance Protection: Principles and Applications*. Hoboken, NJ, USA: Wiley, 2008.
[28] A. Bonetti, M. V. V. S. Yalla, and S. Holst, "The IEC 60255–121:2014 standard and its impact on performance specification, testing and evaluation of distance protection relays," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo.*, May 2016, pp. 1–6.
[29] Schneider Electric, *Technical Manual—Easergy MiCOM P54x (P543, P544, P545 & P546)—Current Differential Protection Relay*, Software Version K3, Schneider Electr., Rueil Malmaison, France, 2020.
[30] *IEEE Standard Common Format for Transient Data Exchange (COMTRADE) for Power Systems*, IEEE Standard C37.111-1999, 1999, pp. 1–55.
[31] X. Li, N. Sudarsanam, and D. D. Frey, "Regularities in data from factorial experiments," *Complexity*, vol. 11, no. 5, pp. 32–45, May 2006.
[32] R. L. Plackett and J. P. Burman, "The design of optimum multifactorial experiments," *Biometrika*, vol. 33, no. 4, pp. 305–325, Jun. 1946.
[33] J. Kiefer and J. Wolfowitz, "Optimum designs in regression problems," *Ann. Math. Statist.*, vol. 30, no. 2, pp. 271–294, Jun. 1959.
[34] T. J. Mitchell, "An algorithm for the construction of 'D-optimal' experimental designs," *Technometrics*, vol. 16, no. 2, pp. 203–210, 1974.
[35] R. K. Meyer and C. J. Nachtsheim, "The coordinate-exchange algorithm for constructing exact optimal experimental designs," *Technometrics*, vol. 37, no. 1, pp. 60–69, Feb. 1995.
[36] E. Sorrentino and V. De Andrade, "Optimal-probabilistic method to compute the reach settings of distance relays," *IEEE Trans. Power Del.*, vol. 26, no. 3, pp. 1522–1529, Jul. 2011.
[37] K. Heussen, C. Steinbrink, I. F. Abdulhadi, V. H. Nguyen, M. Z. Degefa, J. Merino, T. V. Jensen, H. Guo, O. Gehrke, D. E. M. Bondy, D. Babazadeh, F. Pröstl Andrén, and T. I. Strasser, "ERIGrid holistic test description for validating cyber-physical energy systems," *Energies*, vol. 12, no. 14, p. 2722, Jul. 2019.

MIRKO GINOCCHI (Member, IEEE) received the M.Sc. degree in sciences and technology for environment and landscape from the University of Milano-Bicocca, Milan, Italy, in 2014. He joined the Sensitivity Analysis of Model Output Group, Competence Centre on Modelling and the European Commission Joint Research Centre, Ispra, Italy, in 2016. He joined the E.ON Energy Research Center, Institute for Automation of Complex Power Systems, RWTH Aachen University, Aachen, Germany, in 2018, where he is currently a Research Assistant and the Ph.D. Candidate. His research interests include uncertainty and sensitivity analysis for power system applications, interoperability testing for smart grids, and statistical design of laboratory experiments.

THANAKORN PENTHONG received the M.Eng. degree in electrical power engineering from Kasetsart University, Bangkok, Thailand, in 2013. He has been with the Power System Control and Protection Division, Provincial Electrical Authority of Thailand, since 2010. He is currently a Research Assistant with the E.ON Energy Research Center, Institute for Automation of Complex Power Systems, RWTH Aachen University, Aachen, Germany. His current research interests include power system protection, monitoring, and control.

FERDINANDA PONCI (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from Politecnico di Milano, Italy, in 2002. She joined the Department of Electrical Engineering, University of South Carolina, Columbia, SC, USA, as an Assistant Professor, in 2003, and was a tenured promoted in 2008. In 2009, she joined the E.ON Energy Research Center, Institute for Automation of Complex Power Systems, RWTH Aachen University, Aachen, Germany, where she is currently a Professor for monitoring and distributed control for power systems. Her research interests include advanced measurement, monitoring, and automation of active distribution systems. She is an Elected Member of the Administration Committee of the IEEE Instrumentation and Measurement Society and the Liaison with the IEEE Women in Engineering.

ANTONELLO MONTI (Senior Member, IEEE) received the M.Sc. (summa cum laude) and Ph.D. degrees in electrical engineering from Politecnico di Milano, Italy, in 1989 and 1994, respectively. He started his career with Ansaldo Industria and then moved to Politecnico di Milano, as an Assistant Professor, in 1995. In 2000, he joined the Department of Electrical Engineering, University of South Carolina, USA, as an Associate Professor and then a Full Professor. Since 2008, he has been the Director of the E.ON Energy Research Center, Institute for Automation of Complex Power System, RWTH Aachen University. Since 2019, he holds a double appointment with Fraunhofer FIT, where he is developing the new Center for Digital Energy, Aachen. He is the author or coauthor of more than 400 peer-reviewed papers published in international journals and in the proceedings of international conferences. He was a recipient of the 2017 IEEE Innovation in Societal Infrastructure Award. He is an Associate Editor of *IEEE Electrification Magazine*. He is the Editorial Board Member of the *SEGAN* (Elsevier) and the Founding Board Member of the *Energy Informatics* (Springer).