Robustness study and reliability growth based on exploratory design of experiments and statistical analysis: a case study using a train door test bench

Laurent Cauffriez, Pierre Loslever, Nathalie Caouder, Fabien Turgis, Reynald Copin

Robustness study and reliability growth based on exploratory design of experiments and statistical analysis: a case study using a train door test bench

Laurent Cauffriez, P. Loslever, N. Caouder, F. Turgis, R. Copin

Abstract This paper presents a methodology to improve the reliability/robustness of a complex mechatronic system for passenger access, which is greatly stressed over its life cycle. The proposed methodology aims to weed out design problems during the development phase and to test the robustness of the passenger access system under operating conditions as close to the real service conditions as possible. The empirical study process is based on an experimental design series in order to cover a complex experimental domain in a reasonable number of tests for each experimental design. The exploratory aspect that stands in the experimental stage also stands in the data analysis stage thanks to a descriptive multivariate approach that reduces the initial information as less as possible. An application concerning the passenger access system for a train is presented. The experimental process is mainly based on the *D*-optimal design and is used with a 1:1 scale test bench. An example with five input factors is presented; the output data consist in a multidimensional signal containing open/close cycles that may be perturbed. The statistical analysis process starts by considering a scale windowing while beeping both time and multivariate aspects. The frequency values are then investigated using multiple correspondence analysis. Advantages and disadvantages of the proposed design of experiments and data analysis are discussed.

Keywords Design of experiments · Reliability growth · Robustness validation · *D*-optimal design · Scale windowing · Multiple correspondence analysis · Train door system

L. Cauffriez (✉)
TEMPO/PSI-EA 8530, Université Lille Nord de France,
Université de Valenciennes et du Hainaut-Cambrésis,
Le Mont Houy,
59313, Valenciennes, Cedex 9, France
e-mail: laurent.cauffriez@univ-valenciennes.fr

P. Loslever
LAMIH-FRE 3304, Université Lille Nord de France,
Université de Valenciennes et du Hainaut-Cambrésis,
Le Mont Houy,
59313, Valenciennes, Cedex 9, France

N. Caouder
LAMAV-EA 4015, Université Lille Nord de France,
Université de Valenciennes et du Hainaut-Cambrésis,
Le Mont Houy,
59313, Valenciennes, Cedex 9, France

F. Turgis · R. Copin
BOMBARDIER Transport,
1 place des ateliers,
59154 Crespin, France

Abbreviations

ALT	Accelerated life test
AM	Arithmetic mean
DAP	Data analysis path
DoE	Design of experiments
FLC	Factor level combinations
HALT	Highly accelerated life test
MCA	Multiple correspondence analysis
MS	Multidimensional signal
PCA	Principal correspondence analysis
RMS	Root mean square

1 Introduction

Design for reliability is a set of methodologies and best practices performed during design step intended to increase the certitude that a product will meet its reliability goals. For this, some advanced statistic tools (highly accelerated life testing (HALT), accelerated life testing (ALT), design of experiment (DoE)) are usually used in the design process during the development and testing phase of a product.

HALT is used for finding predominant failure mechanisms in a product. It improves the reliability of a product by gradually increasing stresses until the product fails [1–3]. It is therefore suitable for finding design weakness in a very short period of time (usually hours or days). It is generally applied on a whole system and do not work well when there is a wear-out mechanism involved. It leads to improve the destruct and operational margin in regard with the product operational specifications.

ALT is suitable for finding well-understood failure mechanisms. It is used for determining the reliability of a product in a short period of time (usually weeks or months) by accelerating the use environment. It is usually performed on simple materials or components rather than a full system and is often used to characterize wear-out mechanism [4–6]. Instead of stepping up to failure, a level for which the product will survive at is picked (within relevant failure area) and a test is run at this level until failure. The characterization of the wear-out mechanisms assures that it occurs outside customer expectations and outside the warranty period.

DoE is an engineering technique used for performing experiments over a short time. It is used to cover an experimental domain with a number of optimized tests, which depends on the desired accuracy of the results [7–9].

In some fields like aerospace, analysis and testing processes of prototype using accelerated life test approach are quite common [10]. In the case of railway transportation systems, testing processes of prototype very early in the design phase is quite new. This is due to the recent French market contracts between railway operators and rolling stock suppliers which clearly stipulate the service reliability required and the financial penalties if the desired reliability is not attained. These financial penalties are greater even though the design time has become shorter (i.e., 3 years for a new rolling stock system). During the study phase, each sub-system of the rolling stock must be designed, prototyped, tested, and validated.

The passenger access sub-system with its moving steps and door, which is used as illustrative example, follows the same process. This system is developed by a subcontractor, under specifications established by a train constructor (Bombardier Transport in the present case). It is then tested and integrated into the rolling stock fleet before delivery to the customer. The passenger access system is one of the most critical sub-systems of a rolling stock from the standpoint of reliability because it is responsible for 30 to 40 % of the failures during commercial use. As part of its continual improvement process, Bombardier has accepted the challenge of developing a new methodology in the field of railway engineering in order to increase the system's reliability much more rapidly (reliability growth), thus reducing drastically the number of failures in commercial use [11].

The methodology for testing the passenger access system presented in this article seeks to provide a partial response to the problem of reliability growth of this latter. It intends to minimize the risk that the passenger access will not meet its reliability goals in commercial use. This testing process is executed very early in the design phase. It starts with the delivery of the first prototype by the subcontractor and is completed at the beginning of the commercial operations phase. It is based on DoE and data analysis to reach reliability growth of the passenger access. Although HALT and ALT methods are source of inspiration for this research work, we cannot claim having used these latter strictly speaking. Indeed it was impossible to verify all the assumptions for applying these methods strictly because of industrial costs (only one passenger access for testing) and industrial timing constraints.

2 DoE for failure testing

The target of HALT is not to determine the life of the product but to make the product reliable as possible. In fact, it is used to determine the technological limit of the product [12]. It consists in moving upper and lower destruct limits away from product operational specifications limits leading thus to increase the operating margin of this latter as illustrated Fig. 1 and published by Charki et al. [2].

On the other hand, the target of ALT is to determine the life of the product which is very important for mechanical items which wear over time. The design of ALT tests requires determining length of test, number of samples, goal of test, confidence desired, accuracy desired, cost, acceleration factor based either on existing models (Arrhenius, Coffin–Manson, Norris–Lanzberg), or determined by experimentation which implies lots of samples and time [13–16]. When the acceleration factor is difficult to determine (it is the case for products that are ordinarily not in continuous use), a way to determine the acceleration factor is to accelerate the product use rate. To illustrate this point of view, Pascual et al. takes the example of a bearing for a

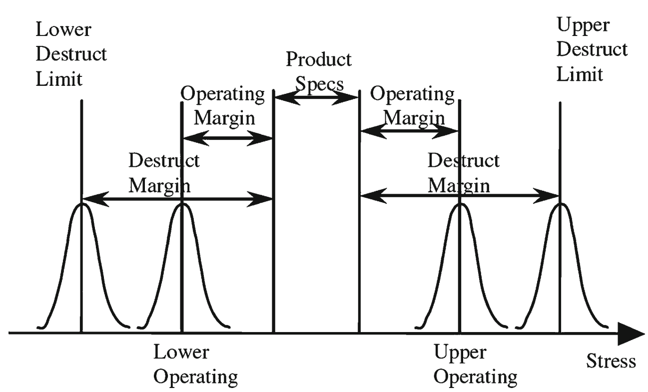


Fig. 1 Determination of the operating limits of the product issued from Charki et al. [2]

washing machine agitator designed for a 12-year median life with a use assumption of eight loads per week [12]. If the machine is tested at 112 loads/week, i.e., 16/day, the median life is thus reduced to 10 months. Nelson points out that failure of bearing can be accelerated by running them at three or more times the normal speed [17].

The DoE method is also crucial to performing experiments over a short time. The definition of an experimental design is to cover an experimental domain with a number of optimized tests, which depends on the desired accuracy of the results. The problem is choosing the type of DoE that is appropriate for the application [18]. Facing with a new complex system (e.g., a passenger access sub-system, with its moving steps and door), using a DoE exploratory approach may be envisaged. Then given the large number of factors and maybe the large number of levels per factor, the industrial experimentation for screening may be performed using the optimal design concept [19].

The optimality theory in DoE is historically based on a linear model (e.g., regression and variance analyses) and begins with work of Kiefer, Wynn, and Gauchi, as cited in Fedorov [20] and Gauchi [21]. A D-optimal design experiment with N experiments, where D letter is used to denote determinant, minimizes the determinant of the variance matrix. Factor heterogeneity, including the quantitative and qualitative variables, does not lead to use orthogonal designs or centered composite designs. However, these two categories of DoE require a very large number of tests (see Table 1).

For the DoE choice, the most appropriate solution is to choose optimal designs, especially the D-optimal design, which is the DoE currently used in industry. The other kinds of DoE, which have specific applications (e.g., for mixture designs in chemistry), have been not considered for this paper. The heterogeneity of factors in this paper requires the use of algorithmic methods to define the best possible experimental design for the experimental domain. Thus, for our analysis, we used the algorithm associated with software SAS-JMP: Bayesian D-optimal coordinate exchange [22]. The futures of the passenger access system are the following: The system is not in continuous use and is event driven. The door with its moving steps is opened if and only if there is a passenger that pushes the button. The closing only happens if there was previously an opening cycle. The system is greatly stressed over its life cycle: pushing of the passengers on the closed door, obstruction to closing, vibration, shocks, vandalism, dust, humidity, and temperature changing… All of these factors may have an impact on the reliability. A sensitivity analysis based on design of experiments for failure testing seems to be a first approach for operational margin improvement process. The proposed methodology based on DoE aims to test the robustness of the passenger access system under operating conditions as close to the real service conditions as possible. In this first DoE approach, the target of the experiments is to accelerate the frequency of the use of this event-driven system and to act on some operating conditions for constructing robust design of the passenger access. The acceleration factor can thus be expressed as a ratio of the number of cycles executed during tests divided by the average door cycles to be theoretically executed per year.

3 Exploratory experimental process of a passenger access system

Notations and terms used in this text:

1. The *passenger access* system is composed of three subsystems: a door, a gap bridge, and a movable step.
2. A *sub-system* contains a set of elementary components.
3. An *experimental factor* is a specific entry variable in the DoE and corresponds to the actions to be simulated in a system stress test (except when specifically mentioned, the word *factor* will be used instead of *experimental factor*).
4. A *cycle* is composed of the opening of passenger access system, followed immediately by the closing of said system.
5. A *test* is composed of a specific number of cycles.

Table 1 Comparison of the number of trials according to Goupy and Creighton [18]

Design type	Calculation formula k is the number of factors investigated q describes the size of the fraction of the full factorial used	Example for the number of trials with k=5 q=1
Full factorial	2^k	32
Centered composite	$2^{k-q}+n_{\text{starpoint}}+n_{\text{center point}}$	16+16+10=42
Fractional factorial (requires 2^q trials less than a full factorial)	2^{k-q}	16
D-optimal	$\geq (k+1)$	6

6. An *experimental campaign* is defined by the type of investigation applied to the system and by the experimental goal.
7. An *experimental design*, which is the result of a *design of experiment*, includes the number of tests performed during the experimental campaign and the modality of the factors applied during each test.
8. An *experimental process* is a sequence of experimental designs for a given system.

The passenger access system is composed of three sub-systems: a door, a gap bridge, and a movable step (Fig. 2). These three sub-systems are considered independent, for two reasons, one mechanical, and one electronic:

– Mechanical—the three sub-systems are all integrated into a rigid frame and do not interact mechanically.
– Electronic—the three sub-systems are controlled by the same electronic control unit, which has not been considered in this study.

The procedure used to study this system has four main phases (several loops may be present):

1. Definition of the experimental domain (phase $\phi 1$): A factor taxonomy is first suggested, each experimental factor class being defined exhaustively. Measurement variables are also determined which is useful in the perspective of the experimental environment design/building, including the test bench.
2. Definition of the chronology of DoE ($\phi 2$): For each experimental step, the possible factor and variable sets are defined and then the DoE (for each DoE, the test chronology may be optimized).
3. Design and building of the test bench ($\phi 3$): The main elements to be defined and implemented are the mechanical structure, actuators, programmable logic controllers, and measurement devices.
4. Data collection ($\phi 4$): This last stage yields data sets including signals with quantitative (e.g., position, speed, current) or qualitative scales (e.g., on/off indicator).

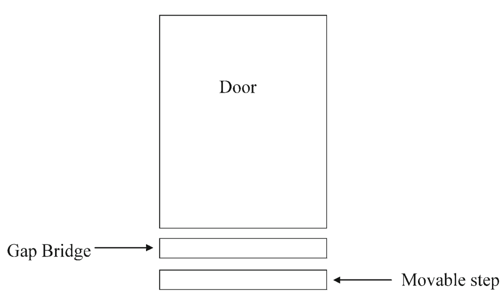


Fig. 2 The passenger access system

3.1 Experimental domain definition

The first step is to specify the experimental domain, which is defined by all areas of variation of the factors making up the DoE [23].

The first task is to list stresses that may act on the passenger access system. To do this, we organized several brainstorming sessions with expert engineers to see the problems that the passenger access system experienced, which led to an exhaustive list of the identified stresses. This list was then translated into factors and parameters for the testing process.

This list of stresses can be divided into four main categories:

– The first category, external factors, defines factors acting on the passenger access system during its operating phase, for example, super-elevation of the train, the weight of the passengers on the train, the weight of passengers climbing on/getting off the movable step, passengers who push on the closed door, pressure waves, shocks, and obstacles (e.g., vandalism).
– The second category, interfacing factors, concerns all the factors due to the interface between the passenger access system and the train. The passenger access system is fixed to a welded metal structure, which has a limited number of manufacturing tolerances that may act upon the features of the passenger access system with respect to the specification sheets.
– The third category, settings factors, refers to all the internal settings of the passenger access system. We assumed that, although a system can be adjusted during the assembly or maintenance phases, the adjustment is likely to be lost during the system's operational phase (e.g., positioning of limit switches, adjustment of belt tension, adjustment of mechanical part). It is therefore necessary to study the influence that can have these settings on the passenger access system's performance.
– The fourth category, climatic factors, deals with environmental factors. Since the trains operate in moderate climate, these factors have a minor influence and are not all considered, apart from the effect of rain, dust, and various washing products.

The experimental domain that was defined in first step highlights three major difficulties. First, the experimental domain consists of a total of more than ten factors. During the testing period, it is desirable to analyze the direct influence of factors, but also the influence of interactions between factors. Ultimately, this leads to a very large experimental domain that is difficult to cover in a reasonable number of tests. Second, the system consists of three sub-systems that are mechanically independent due to design choices. This independence makes the experimental domain complex because the factors must be applied specifically to the individual sub-systems and not to all the sub-systems simultaneously. For example, when the movable steps are deployed, the weight of passengers on the steps will not have a simultaneous interaction with the force applied by the passengers pushing on the door. The last difficulty lies in the nature of the experimental domain. The set of factors contains quantitative factors (e.g., the super elevation of railway is a quantitative factor ranging from $-\alpha$ to $+\alpha$ angular degree) and qualitative factors (e.g., the presence of an obstacle when closing the door may take only two levels: the presence or absence of obstacle).

Taking all these constraints into account makes the experimental domain very complex and requires a specific approach, which is explained in the section below.

3.2 Inter- and intra-experimental design chronologies

The question of the length of the experimental period is crucial. Given the large experimental domain described above, it is not realistic to try to take all the factors simultaneously into account due to industrial timing constraints. In fact, out of the 3 years of development, the design and construction of a basic prototype on average take a year and a half. The time for completing the reliability and robustness testing of the prototype is thus reduced to a year and a half, which includes the time for integrating the prototype on the test bench, the time for running the tests, and the time required to analyze the failures and to take corrective action for both the prototype and the test bench (it is worth to note that the experimental environment is complex because it contains actuators that perform perturbing actions on the door system).

Therefore, it was decided to define an experimental methodology (Fig. 3) that includes four successive test campaigns, each with a specific goal. The three first campaigns focused on the early failures of the passenger access system and correspond to a reliability growth program. The last campaign focused on validating the robustness and the global reliability of the passenger access system. The proposed methodology can thus be broken down into four activities: DoE-1 test campaign (A1, focus on the door only), DoE-2 test campaign (A2, focus on both door and gaps) for reliability growth, DoE-3 test campaign for robustness validation (A3, focus on both door and gaps), and DoE-4 test campaign for global reliability validation (A4, focus on both door and gaps).

The evolution of the growth of reliability is shown in the succession of prototype versions at the output of each Ai activity, i.e., improved, mature, series, and validated series prototype. The output of the experimental method is a robust series prototype, as well as the identification of the most influential factors for the passenger access system during its life cycle. The results presented in this paper concerns the data analysis of the first DoE test campaign, i.e., DoE-1.

3.2.1 Design of experiments

Among the various types of existing DoE, it was decided to use *D*-optimal principle because of the complexity of the experimental domain, which had a lot of factor types and constraints [21]. As stated above, the DoE was performed using the *D*-optimal criterion. However, we used several simplifying assumptions:

- Assumption 1: The experimental domain was divided into three distinct experimental domains because the three sub-systems (i.e., the door, the gap bridge and the movable step) of the passenger access system are mechanically independent.
- Assumption 2: Strictly, using DoE involves using a new passenger access system for each test, which is not feasible/realistic due to budgetary constraints. In this study, the number of cycles performed during each test is far less than the number of cycles that the door and the two steps may perform over their whole life cycle.
- Assumption 3: The reliability of the tests is sufficient to avoid having to use DoE with a replicate.
- Assumption 4: The interfacing factors representing the vehicle structure are not feasible within the context of

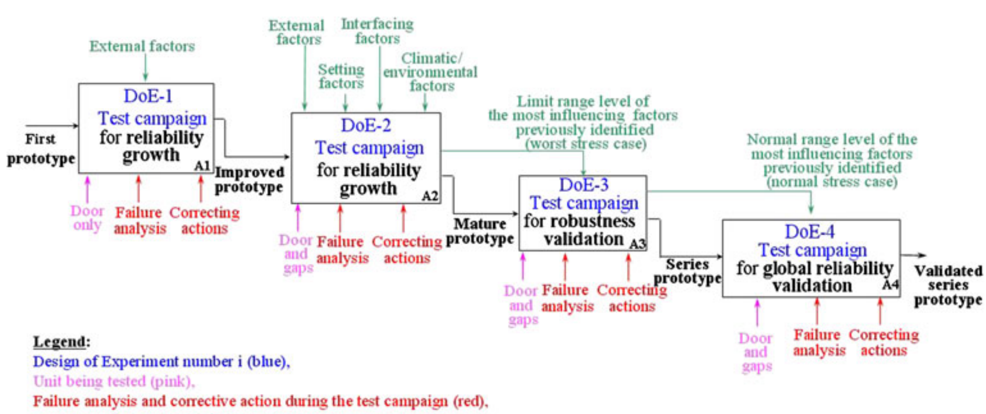


Fig. 3 DoE for the robustness and the reliability growth of the passenger access system

this study. It would be too restrictive and too expensive to provide a set of interchangeable structures to represent all the possible configurations. We therefore chose a “worst case” in terms of manufacturing tolerances for the vehicle structure.

3.2.2 Reducing the experimental domain

The best way to decrease the number of tests is to act directly on the factors. The dimension of an experimental matrix depends only on the number of factors being considered. Since the passenger access system is considered as three independent sub-systems, the number of factors is divided by 3. However, these three sub-systems can have common factors whose particularity is that they can be applied in a same way to the whole passenger access system and not to one sub-system. In order to decrease the number of tests, we then decided to reduce the analysis to the common and independent factors in each sub-system. If all the factors of each sub-system are in equal number and in equal nature, one DoE per campaign is needed to cover the behavior of the whole passenger access system.

3.2.3 Optimizing the test sequences

The sequence of the tests is randomly generated by the DoE. For this case, it must be optimized to save time in the experiments. Each transfer from test i to test $i+1$ requires setting a new configuration of factors. Some factors require more time to set than others (e.g., the transition from a super elevation of $-\alpha$ to $+\alpha$ angular degree requires 2 min to adjust while passing a belt tension from a minimum value to a maximum value takes 30 min to adjust). Choosing an appropriate sequence of tests helps to drastically reduce the adjustment time between tests and thus improves the overall time of test campaigns. Therefore, test sequences were randomly generated by DoE and were optimized using an algorithm that schedules the tests according to the time of adjustment and maintenance operations on the passenger access system. Let us consider the DoE-1 example (for confidential reasons, some factor levels may not be indicated).

From a preliminary study involving reliability and production engineers, the experimental domain is composed of U=5 factors:

- X_1, the super elevation (cant) of the railway
- X_2, the shocks on the door
- X_3, the obstacles in the door
- X_4, the vertical loading of passengers
- X_5, the pushing of the passengers on the door

The factor set containing quantitative and qualitative scale factors, thus each one with two levels at least, the total number of experimental adjustments (trials) is large (more than 2^5 for a full-factorial design). Thanks to D-optimal option of software SAS-JMP, the DoE-1 sets the number of trials equal to I=29 trials (24 trials recommended by SAS-JMP plus five trials situated in the center of the factor region to improve the experimental design precision).

A trial is a factor-level combination, labeled (α,β,χ,δ,ε), which are the five elements corresponding to the levels of the U=5 factors. The number of cycles of door opening and closing is limited to D=160 cycles per trial in order to respect the imposed industrial timing plan. Table 2 presents the trial sequence initially proposed by SAS-JMP [20], and Table 3 reports the results of the optimization process applied to this initial sequence.

Shocks and obstacles stresses applied at each cycle can theoretically lead to the degradation of the door being tested because technological limits are reached in this “systematically” case. To reproduce on the testing bench the commercial use conditions—it is a fact that shocks and obstacles occur on a randomly way during commercial use—it was decided to adopt a pseudo-random approach for the generation of shocks and obstacles stresses. Thus, for a given quintuplet of factor levels, the experimental protocol is conducted so that all the cycles are not disturbed: only one cycle out of every 10 for the obstacles (cycles 10, 20, 30,…) and one cycle out of every 20 for the shocks is disturbed (cycles 20, 40, 60,…) resulting in a double “impact and obstacle” disturbance only for cycles 20, 40, 60, 80, 100, 120, 140, and 160.

The door obstacles are performed when the relative door position equals 1 m (i.e., mid-point between fully closed (the distance between the two door leaves is 0 m) and fully open (the distance between the two door leaves is 2 m)). The door impacts are performed when the door is either fully open or fully closed, as follows: cycle 20, right leaf when the door is closed; cycle 40, left leaf when the door is open; cycle 60, left leaf when the door is closed; cycle 80, right leaf when the door is open; and so on up to cycle 160.

3.2.4 Success criteria

The goal of the DoE-1 and DoE-2 campaigns is to detect significant failures according to the reliability growth program. Therefore, any time the test campaign is stopped due to the failure of the unit being tested can be seen as a success, except for a deficiency of the test bench, of course. When D cycle test sequences are fully completed, it can be considered as a success. If a failure occurs before the end of the D cycles, it is then necessary to analyze the failure and to repair it in order to complete the remaining test sequences.

The goal of DoE-3 is to examine the failure due to the worst stress case, i.e., the most influencing factors previously identified, and the goal of DoE-4 is to validate the robustness and reliability of the unit. Thus, any fully completed D cycle test sequence is declared a success. In other

Table 2 SAS-JMP software experimental matrix: the initial trial sequence. Li represents the value of the factor for the current DoE (the real values of Li cannot be published for reasons of confidentiality)

SAS-JMP sequence of trials	*X*1 Super elevation of the railway	*X*2 Shocks on the door	*X*3 Obstacles in the door	*X*4 Vertical loading of passengers	*X*5 Pushing of the passengers on the door
1	L3	L1	L1	L2	L3
2	L3	L1	L2	L2	L1
3	L3	L1	L2	L1	L3
4	L3	L3	L2	L1	L1
5	L1	L3	L2	L2	L1
6	L1	L1	L1	L2	L1
7	L1	L1	L2	L1	L1
8	L1	L3	L1	L1	L1
9	L1	L1	L1	L1	L3
10	L2	L1	L2	L1	L1
11	L2	L3	L2	L2	L1
12	L1	L2	L2	L3	L2
13	L3	L1	L1	L1	L1
14	L1	L3	L1	L2	L3
15	L2	L2	L2	L3	L2
16	L1	L1	L2	L2	L3
17	L3	L3	L1	L2	L1
18	L2	L2	L1	L3	L2
19	L1	L3	L2	L1	L3
20	L1	L2	L1	L3	L2
21	L2	L1	L1	L2	L1
22	L2	L1	L1	L1	L3
23	L2	L3	L1	L1	L1
24	L3	L2	L1	L3	L2
25	L2	L3	L2	L1	L3
26	L2	L1	L2	L2	L3
27	L3	L3	L1	L1	L3
28	L3	L3	L2	L2	L3
29	L2	L3	L1	L2	L3

words, there should not be any more significant system failures during the DoE-4 test campaign since the sub-system being tested has been improved and revised during the three previous campaigns.

The experimental methodology requires several experimental designs. Yet, only the first experimental design has been completed and the recorded data analyzed. For this reason, only the data analysis of DoE-1 is presented in this paper. Our goal was to suggest a very first data analysis procedure in which the initial time database is summarized as little as possible.

3.3 Test bench design

The test bench must be able to reproduce simultaneously all the mechanical, internal, and environmental stresses encountered by the passenger access system while it is operating [24]. The test bench must also be able to reproduce the same mechanical interfaces between the passenger access system and the train. These aspects led us to choose a scale 1:1 test bench in which the system is integrated into a metallic structure representing the train, as shown in Fig. 4. This structure is also subjected to the action of several jacks to simulate the various stresses. The failures detected during the tests highlight the "weaknesses" of the system, and adjustments following the failures make the system more robust and thus more viable at least in the first part of its life cycle.

3.3.1 Description of the test bench

The test bench can be broken down into five sub-parts (Fig. 4):

– The passenger access system, which is composed of three independent mechanical sub-systems and the associated control equipment

Table 3 Experimental matrix that resulted from applying the optimization process to the initial trial sequence (grey zones model the new trial concatenation to reduce the adjustment time between tests and thus to improve the overall time of test campaigns)

Optimized sequence of trials	X1 Super elevation of the railway	X2 Shocks on the door	X3 Obstacles in the door	X4 Vertical loading of passengers	X5 Pushing of the passengers on the door
4	*L3*	*L1*	*L2*	*L3*	*L1*
5	*L1*	*L2*	*L2*	*L3*	*L1*
11	*L2*	*L2*	*L2*	*L3*	*L1*
28	*L3*	*L2*	*L2*	*L3*	*L3*
19	*L1*	*L1*	*L2*	*L3*	*L3*
25	*L2*	*L1*	*L2*	*L3*	*L3*
27	*L3*	*L1*	*L1*	*L3*	*L3*
8	*L1*	*L1*	*L1*	*L3*	*L1*
23	*L2*	*L1*	*L1*	*L3*	*L1*
17	*L3*	*L2*	*L1*	*L3*	*L1*
14	*L1*	*L2*	*L1*	*L3*	*L3*
29	*L2*	*L2*	*L1*	*L3*	*L3*
15	*L2*	*L3*	*L2*	*L2*	*L2*
12	*L1*	*L3*	*L2*	*L2*	*L2*
20	*L1*	*L3*	*L1*	*L2*	*L2*
18	*L2*	*L3*	*L1*	*L2*	*L2*
24	*L3*	*L3*	*L1*	*L2*	*L2*
2	*L3*	*L2*	*L2*	*L1*	*L1*
7	*L1*	*L1*	*L2*	*L1*	*L1*
10	*L2*	*L1*	*L2*	*L1*	*L1*
3	*L3*	*L1*	*L2*	*L1*	*L3*
16	*L1*	*L2*	*L2*	*L1*	*L3*
26	*L2*	*L2*	*L2*	*L1*	*L3*
1	*L3*	*L2*	*L1*	*L1*	*L3*
6	*L1*	*L2*	*L1*	*L1*	*L1*
21	*L2*	*L2*	*L1*	*L1*	*L1*
13	*L3*	*L1*	*L1*	*L1*	*L1*
9	*L1*	*L1*	*L1*	*L1*	*L3*
22	*L2*	*L1*	*L1*	*L1*	*L3*

– The structure representing the train in which the passenger access system operates in a worst-case configuration
– The frame of the test bench and its actuators, which includes the different electro-mechanical and electro-pneumatic devices used to evaluate the unit being tested
– The controller, which controls all the actuators
– The control and acquisition unit, which controls the unit being tested, synchronizes the different test bench actuators, and centralizes data produced by the sensors placed on the passenger access system and on the test bench itself

3.3.2 Monitoring the unit being tested

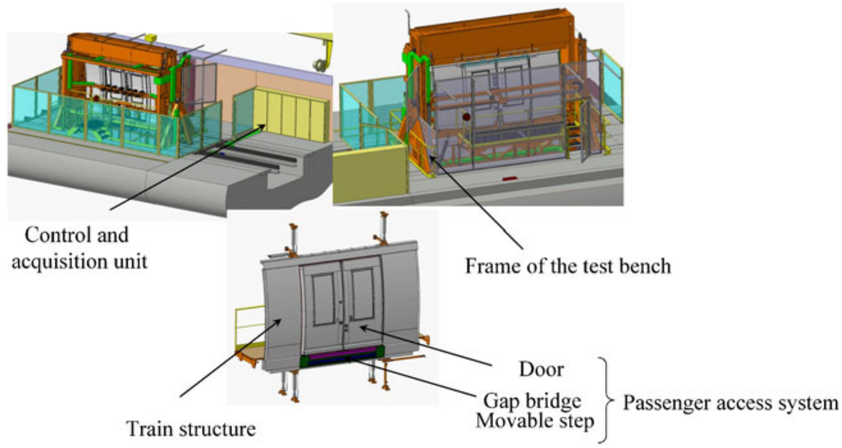
The stresses applied during experiments can theoretically lead to the degradation of the unit being tested, thus making the unit inoperable because the technological limits have been reached. This inoperability can be tolerated in an elementary component but not for the complete system. Therefore, the test bench has been equipped with strain gauges that check if the system being tested is still working in the elastic range or if the elastic limits have been exceeded (i.e., finite element calculations). If the elastic limits have been exceeded, the test bench is immediately stopped, and this is considered as a system design failure.

3.3.3 Measured and computed variables

For the door, the main variables are the following:

– Y_1, relative position signal (i.e., the distance between the two leaves, which is between 0 and 2 m) given by the train door's control unit with a sampling rate of 10 Hz
– Y_2, relative door speed signal, here calculated from a smoothing procedure and derivation computation, used with the position signal, which should run between 0 and about 0.8 m/s

Fig. 4 The developed test bench for passenger access system reliability growth



– Y_3, motor current signal, given by the train door's control unit with a sampling rate of 10 Hz, which should run between 0 and 15 A
– Y_4, door closed signal, produced by the train door's limit switch with a sampling rate of 100 kHz (response time of the limit switch is between 30 and 100 ms), which makes it possible to give the opening and closing instants

For each trial, these variables yield a multidimensional signal (MS) with heterogeneous components, with some being quantitative and some being qualitative. Given the presence of several experimental designs, each one yielding many MS, maybe with doubtful data pieces, results cannot be reached using a single step.

4 Exploratory statistical analysis of a passenger access system

In the following text, to facilitate comprehension, we have used the following notational system:

– An outlined letter designates a set (e.g., \set $\mathbb{E}$);
– An uppercase letter designates the set size (e.g., $E0$= card($\mathbb{E}$)), where card is the cardinality function
– A boldface letter designates either a vector or a matrix (e.g., $\boldsymbol{x}=(x_1, x_2, x_3)'$ or $\boldsymbol{X}=\boldsymbol{x}\times\boldsymbol{x}'$).
– Element (of a set, a vector, …) and the number of elements will be noted using the same letter, with lower and upper characters respectively, such as with i=1, 2, …I.

Given these basic notations, $\mathbb{U}=\{X_u, u=1,\ldots,U\}$ and $\mathbb{V}=\{Y_v, v=1,\ldots,V\}$, are, respectively, the initial factor set (i.e., independent variables; even though *time* is a factor when time data is present, *time* will be not considered as belonging to U) and the initial variable set (i.e., dependent variables). In our case, the dependent variables yield multidimensional raw signals with V components.

Yet, only DoE-1 data are available, the settings, programming of control/acquisition unit, and data collecting have lasted for about 6 months. The procedure used to analyze the collected data has five main phases [25]:

1. Data characterization (φ1): The initial time data sets are changed into new ones, the latter being more or less large, i.e., summarizing indicators (arithmetic mean, standard deviation, root mean square, extreme value, spectral analysis coefficient) are computed for the whole signal or specific time windows. These indicators become the analysis variables.
2. Data coding (φ2): The values are transformed so that they are homogeneous in a multivariate analysis perspective (e.g., principal component analysis or hierarchical clustering method [25]) or so that a scale model matches the requiring input scale model of a specific method (e.g., quantitative values changed into order values for some nonparametric tests).
3. Data organization (φ3): The data sets are organized into a series of two-entry tables. The rows correspond to the experimental situations and the columns to the variables (i.e., time variables or indicators) and the factors (i.e., experimental and time factors).
4. Data table study (φ4): During this stage, the data tables are analyzed in order to find relationships between the variables and/or between the factors and the variables.
5. Results presentation (φ5): The results are summarized using one to three of the usual model forms, namely *verbal* (e.g., a text that states the most influent factors), *mathematical* (e.g., a regression model), and *graphical* (e.g., variable vs. factor plot, box plot). These models must be used so that φ4 outputs are understandable for non expert of methods employed in φ4.

In the section below, the sequence (φ1, φ2, φ3, φ4, φ5) will be named *data analysis path* (DAP), where each phase requires one or more specific methods. A priori, many DAP are available; for instance, one can imagine a taxonomic dimension with two classes: descriptive vs. inferential statistics DAP.

The results presented in this paper concern the door subsystem (activity A1 of Fig. 3). In the following, the generic value of the measurement variable will be labeled $\mathbf{y}_{\alpha\beta\chi\ldots,abc\ldots}$, where the subscripts before the comma designate the empirical situations and the subscripts after the comma are used to designate the variables. For instance, the generic value of a time variable is labeled $y_{ij,v}$ (remember that i corresponds to a specific combination of the five factors; $i(i=1,\ldots, I=29)$ is used instead of the quintuplet $(\alpha,\beta,\chi,\delta,\varepsilon)$ in order to simplify the notation, and j corresponds to the time sample). From the perspective of an exploratory data analysis, a first DAP will be used to mine the data (DAP I), and then a second DAP will be used in order to show the influence of the factors (DAP II).

4.1 First data analysis path (DAP I): data mining

The DAP I purpose is to mine the data. Thus, we will try to find if a malfunction is possible for the three trials that stopped before $D=160$ cycles: For trials 9, 11, and 15, the number of cycles is 138, 158, and 156, respectively; thus, the total number of cycles is not $I\times D=4{,}640$ but $DT=4{,}640-(22+2+4)=4{,}612$. In the next section, the number of door opening/closing cycles per trial i will be named D_i.

4.1.1 Data characterization

Although the database is rather large and complex, the MS must not be summarized too much. Each of the $I=29$ trials is an execution of a MS with $V=3$ components (position, speed, and motor current for the door); each component is obtained for a chronology of D_i door opening/closing cycles, with the time distance between two samples j and $j+1$ being 0.1 s. Thus, for a very first analysis, let us characterize the behavior of each time variable v ($v=1, \ldots, V=3$) during each door cycle d ($d=1, \ldots, D_i$) with its magnitude histogram.

Given the high value of the total number of cycles ($DT=4{,}612$) and the desire to find main trends for this DAP I, only $S=5$ space windows are considered at first (labeled with an indicator running from 1 to 5, where labels 1 and 5 correspond to windows containing the lowest and highest values, respectively). Thus, this characterization step yields $f_{id,vs}$ as generic value (i.e., the frequency within a space window s for the variable v obtained for the cycle d of the trial i).

To illustrate this data characterization stage, let us focus on the position and current variables ($v=1$ and $v=3$) for $i=1$ and $d=1$. Figure 5a, b shows the corresponding data characterizing stage inputs. Magnitude histogram analysis playing a main role in the space windowing-based characterization, Fig. 5c, d displays the two corresponding histograms. To line out the possible presence of outliers, the two magnitude histograms obtained from all time sample for the 29 trials are worth to be visualized. The two overall magnitude histograms, Fig. 5e, f, highlight that there are overlarge values for the motor current, but the frequency of such a behavior is rather low. Consequently, in an exploratory multivariate analysis context, using a space windowing can be seen as a good starting point, more particularly better than using usual indicators such as the arithmetic mean or the RMS value. With the two time data sets displayed Fig. 5a, b, the two frequency output sets are $\mathbb{F}\mathbb{I}_{11,1}=\{30, 16, 15, 16, 23\}$ for the position signal and $\mathbb{F}\mathbb{I}_{11,3}=\{63, 22, 9, 5, 1\}$ for the motor current signal.

4.1.2 Data coding

Each generic data set $\mathbb{F}\mathbb{I}_{id,v}$ ($i=1, \ldots, 25$; $d=1, \ldots, D_i$; $v=1, \ldots, V=3$) contains only frequency data. Thus, this data coding phase is not required for a multivariate analysis (instead of having three different units of measurements —meters, meters per second, and ampere—only one unit is still present—percentage).

4.1.3 Data organization

The frequency values are organized in a two-entry table, a row r corresponding to an empirical situation, i.e., a pair (i, d), and a column c corresponding to a system state, i.e., a pair (v,s). This data organization stage yields a main table $F1$ with $R=DT=4{,}612$ rows and $C=V\times S=3\times 5=15$ columns. Due to the presence of the space windowing, the sum across columns for any row is

$$\sum_{v=1}^{V=3}\sum_{s=1}^{S=5} fid, vs = fid, \bullet\bullet = V \times 100 \tag{1}$$

The frequency values are in percentages and a point means a sum over the corresponding subscript.

4.1.4 Data table analysis

Table $F1$ may be analyzed using multiple correspondence analysis (MCA) [26]. Unlike principal component analysis (PCA), in MCA, the distance between two rows and the distance between two columns are based on the same metric (i.e., the chi-squared (χ^2) metric). With our data, the distance between two row points within the space R^{15} is:

$$\begin{aligned} d^2(r,r') &= d^2((i,d),(i',d')) \\ &= \sum_{v=1}^{V=3}\sum_{s=1}^{S=5}\frac{1}{f\bullet\bullet, vs}\left(\frac{f_{id,vs}}{f_{id,\bullet\bullet}}-\frac{f_{i'd',vs}}{f_{i'd',\bullet\bullet}}\right)^2 \end{aligned} \tag{2}$$

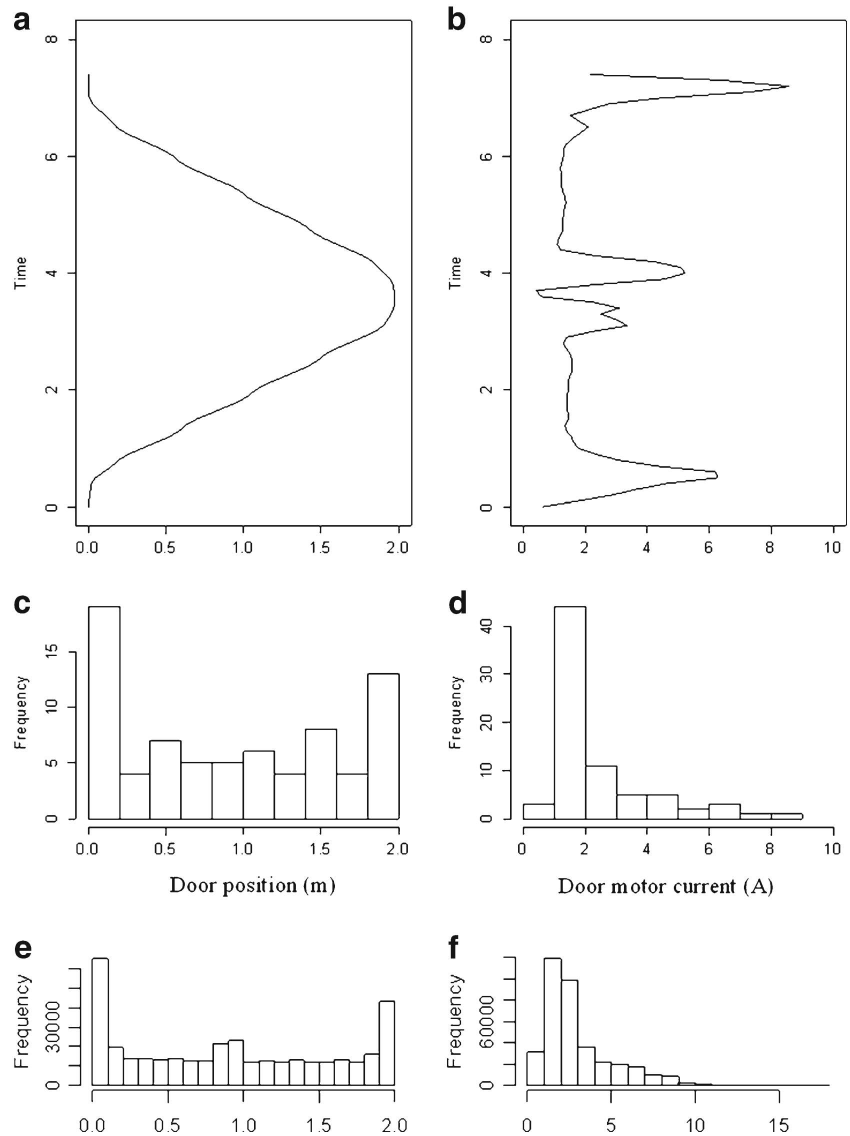


Fig. 5 Example of signals (**a**, **b**) and magnitude histograms for the door position and door motor current (**c**, **d**) for the cycle $d=1$ with the trial $i=1$ (i.e., from about 70 time samples at 10 Hz). The bottom figures (**e**, **f**) show the magnitude histograms computed for all the time samples of all the cycles of all the trials (i.e., from about 400,000 points)

The distance between two column points within the space $\mathbb{R}^{4612}$ is:

$$d^2(c,c') = d^2((v,s),(v',s'))$$
$$= \sum_{i=1}^{I=15} \sum_{d=1}^{Di} \frac{1}{fid,\bullet\bullet} \left(\frac{f_{id,vs}}{f_{\bullet\bullet,vs}} - \frac{f_{id,v's'}}{f_{\bullet\bullet,v's'}} \right)^2 \quad (3)$$

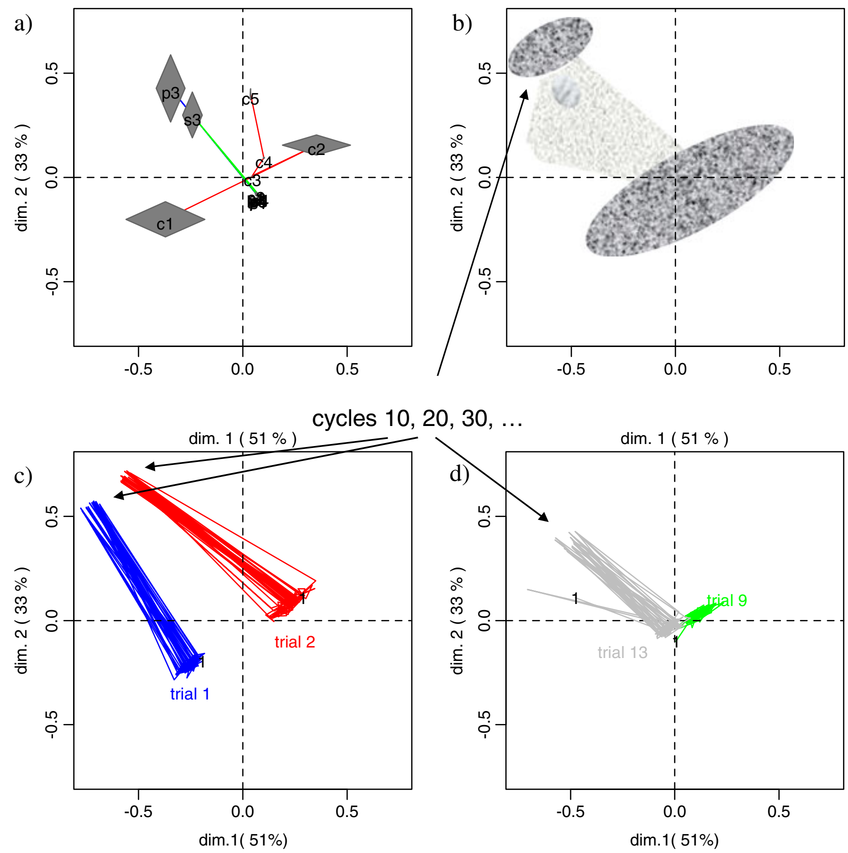
Figure 6 presents some of the graphical outputs that can be drawn for MCA of $F1$.

Figure 6a, which explains about 84 % of the total inertia, shows that the main distinction between the cycles come from the space windows 1 and 2 of the door motor current (windows labeled c1 and c2) and space windows 3 for the door position and speed (labeled p3 and s3). The main roles played by these four space windows are first due to the obstacle effect (Fig. 6b).

For instance, let us have a look at trial 1 cycles: There are cycles 10, 20, 30, … in the positive side of axis 2 and other cycles on the negative side (Fig. 6c). Given the relative positions of these two kinds of cycles and the top position of p3 and s3 (Fig. 6a), cycles 10, 20, 30, … should present higher frequency values for p3 and s3. This is consistent with the values given in Table 4 (the same conclusion can be drawn for trial 2) and the way impacts are performed, i.e., when the door is in an intermediate position (p3 space window).

Given these intra-trial differences outlined by axis 2, axis 1 rather shows inter-trial differences instead. For instance, given the overall position of trials 1 and 2 (Fig. 6c) and space windows c1 and c2 (Fig. 6a), trial 1 should have higher-frequency values for c1 and lower values for c2. Table 4 confirms this. The values are consistent with the experimental design adjustment: The main difference between trials 1 and 2 is that the vertical loading of passengers is much higher in trial 2 than in trial 1, involving a larger motor current to move the door.

Fig. 6 MCA output of table $F1_{4,612\times15}$ with one and two main axes. **a** Projection of the $V\times S=3\times5=15$ column points (for each variable v, the $S=5$ space windows are linked in increasing order. **b** Schematic presentation of the 4,612 row point positions, yielding two main clusters and some points between these two clusters. **c** Projection of the 160+160 door opening/closing cycles corresponding to trials 1 (in *blue*) and 2 (in *red*). **d** Projection of the 138+160 cycles corresponding to trials 9 (in *green*) and 13 (in *gray*). For subplots **c** and **d**, "1" indicates the first cycle of the corresponding trial



MCA also allows to highlight outliers. This is the case of the first cycle of trial 13 during which the set of factor levels do not permit the closing of the door within 3.5 s. In fact, the door control unit (DCU) operations are based on a learning feedback from the previous cycles. For this reason, only the first cycle is mainly impacted (Fig. 6d).

Finally, let us consider trial 9, which stopped at cycle 138 (instead of 160). For this trial, the experimental design and protocol settings did not planned impact that is why there is no outlier for green points of Fig. 6d (trial 9). If Figure 6d shows that intra-trial differences are rather low for trial 9 (thus consistent with input reference), the relative positions of the corresponding cycle points do not show any trend allowing to explain why this trial 9 has stopped at cycle 138 instead of cycle 160. In this case, a posteriori data checking showed that the malfunction was due to a brutal loss of communication between DCU and the test bench. So, even though 160 cycles were performed, only 136 cycles were recorded, thus the problem is not due to mechanics but to informatics.

4.1.5 Results presentation

In a data mining perspective, we consider that MCA output, with Fig. 6 as an output subpart (this figure only considers axes 1 and 2 but there are next main axes), can be seen as a good result presentation way. Even through this figure is complex, intra- and inter-individual trial differences can be

Table 4 Example of frequency values to show some differences highlighted by MCA for $F1$ (approximations may yield total slightly different from 100 %)

Situation		Position						Speed						Current					
Trial	Cycle	*p*1	*p*2	*p*3	*p*4	*p*5	Total	*s*1	*s*2	*s*3	*s*4	*s*5	Total	*c*1	*c*2	*c*3	*c*4	*c*5	Total
1	1	30	16	15	16	23	100	21	20	20	18	21	100	63	22	9	5	1	100
1	10	15	7	60	7	12	100	9	11	63	7	10	100	64	16	8	6	6	100
2	1	30	15	17	15	23	100	22	21	16	19	22	100	17	60	11	8	4	100
2	10	14	8	61	6	11	100	10	12	61	7	10	100	45	29	8	10	8	100

shown at a glance, as well as outliers. A careful MCA output study yields *when* these outliers occur (which cycle of which trial?) and *where* (which space window of which MS component?). This long but essential data mining DAP being performed, let us now consider a new DAP where the main goal is to reach the experimental design initial aim, i.e., to show factor effects.

4.2 Second data analysis path (DAP II): influence of the factors

This second DAP will maintain both the multivariate and descriptive contexts.

4.2.1 Data characterization

With the perspective to underscore nonlinear relational phenomena and to reduce as less as possible information loss, the space windowing concept will be maintained. Each trial i that does not contain an impact or an obstacle is characterized when averaging, for each space window$s$ (s=1, …, S=5) of each MS component v (v=1, …, V=3), the frequency values over the D_i cycles it contains, using the following equation:

$$\bar{f}_{i\bullet,vs} = \frac{1}{Di}\sum_{d=1}^{Di} f_{id,vs} \qquad (4)$$

For the other trials with impacts and/or obstacles, the corresponding cycles are removed before averaging since they are highly disturbed. Thus, the data characterization yields the generic set $\boldsymbol{F2}_{i,v}$ (i=1, …, 29; v=1, …, V=3), each set containing S=5 frequency values.

4.2.2 Data coding

As with the previous DAP, this phase is not required since all the values are homogeneous (frequencies).

4.2.3 Data organization

The frequency data are organized by generating a table $F2$ with R=I=29 rows and C=V×S=3×5=15 columns. To show how the factor influences can be highlighted, let us consider two factors the super elevation with three levels (−7°, 0°, and 7° levels) and the vertical loading of passengers with 3 levels (zero, medium, and high; for confidentiality reasons, the value are not indicated). The corresponding data set can be organized by generating a table $F2'$ with six rows, one for each level of each factor and 15 columns, in which each frequency row is obtained when averaging over the trials that contain a given factor level.

To summarize, the data organization output is a set of two-entry tables, the row corresponding to a trial (thus a combination of the U=5 factors) or a trial subset (thanks to an averaging operation over the factors) and a column to a (v, s) pair (thus a given space window of a given MS component).

4.2.4 Data table analysis

To show the differences between the I=29 trials and where these differences come from (i.e., from which (v,s) pairs), table $F2$ can be analyzed using MCA. In order to show the super elevation and load influences, $F2'$ rows can be considered as six supplementary row points. These six points do not participate in the positioning of the main axes [26].

Figure 7 gives the MCA output for axes 1 and 2. These axes exhibit about 98 % of the total inertia. It is worth noting that these two axes are mainly controlled by the motor current space windows. For main axis 1, the motor current windows that play a main role in distinguishing the I=29 trials are c1 and c2, this in the following way: Trials with motor current values mainly situated within window c1 are situated on the left side, such as trials 1 or 27. On the opposite side are the trials with motor current values mainly situated within window c2, such as trials 19 or 22. This distinction, shown by axis 1, is consistent with the way the I=29 trials are designed. For instance, a main distinction between trial 1 (on the very left side) and trial 22 (on the right side) is that the former is performed with a 0 level for factor X_5 (the pushing of passengers on the door), which is not the case for trial 22.

The passengers' pushing is the stress that increases the friction on the door system the most. To maintain the contractual performance value (i.e., to maintain the opening/closing time), the door control unit must increase the current value to face up with the friction increase. Axis 2 allows us to distinguish trials with higher values for the current (presence of c5 window), such as trial 17.

Once the main role played by the C=V×S=15 columns points and I=15 row points have been established, it is now possible to show the factor influence using the concept of supplementary row point (with one point per factor level). According to axis 1, the super elevation influence is displayed with rather large distances for points labeled "+7°" (on the left side) and "0°" (near the crossing of the two axes) and points "0°" and "−7°" (on the right side).

To summarize, given the multifactor and multivariate aspects of the database (here with U=5 factors and V=3 variables), MCA allows to show:

– The most influenced variables and space windows (here c1 and c2)

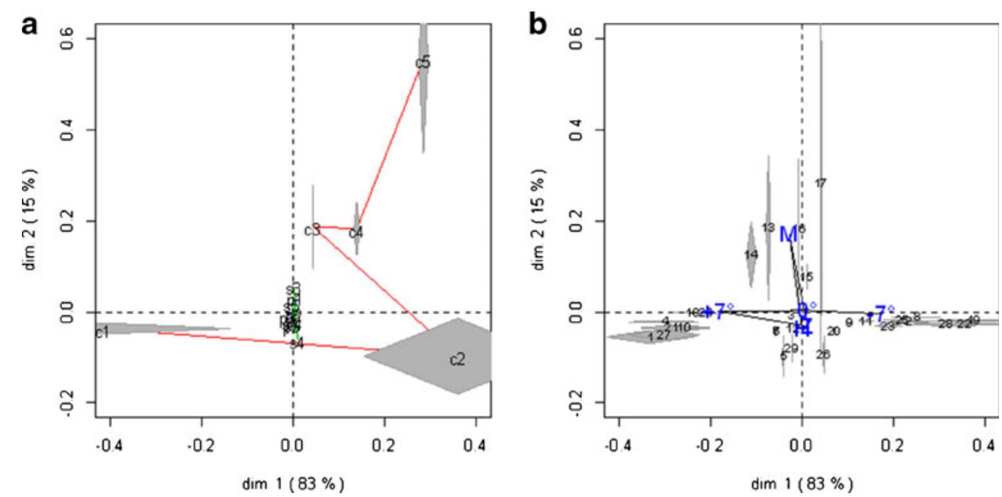


Fig. 7 MCA output of $F2_{29\times 15}$ with one and two main axes. **a** Projections of the $V\times S=3\times 5=15$ column points (for each variable *v*, the $S=5$ space windows are linked in the increasing order; the two lozenge sizes indicate the point contribution in positioning axes 1 and 2); **b** projection of $I=29$ row points with an active status and six row points with passive status: three supplementary points for the super elevation factor (levels are labeled −7°, 0°, and +7°, in *large blue symbols*) and three supplementary points for passenger load factor (levels are labeled *Z* for zero, *M* for medium, and *H* for high, in *large blue symbols*)

– The most influent factors (here super elevation, for confidentiality reasons all the factors are not considered)
– How such factors influence the variables (here moving from −7° to +7° yields a current decrease from c2 to c1)

4.2.5 Results presentation

Characterizing phase output data being data frequency sets, we suggest to consider figures showing such kind of values, i.e., to show histograms. To compare a set of $S=5$ frequencies with a usual indicator, each histogram subplot will also display indicator value. For instance, let us consider the arithmetic mean and the root mean square (RMS). Figures 8 and 9 show the differences and/or similarities highlighted by MCA more quantitatively.

Figure 8 shows histograms for three trials: two trials (1 and 19) that appear very separated according to axis 1 and one trial (17) that had a top position according to axis 2 (see Fig. 7). The nine subplots confirm that the main differences are present for the motor current variable and for space windows 1 and 2.

For instance, trial point 1 which has a very left position (Fig. 7) should have motor current values more often within the space window 1 (labeled c1, Fig. 7a). Right column of Fig. 8 confirms this. The contrary stands for trial 19. Still watching the Fig. 8 right column, it is worth noting that if the three histograms are different, the three RMS values are identical.

Figure 9 displays the super elevation and load factor influences for the three MS components. The relative positions of the six level points according to axis 1 (see Fig. 7b) show that the influence of the super elevation factor should be much higher than the influence of the load factor, and these influences should mainly present for the motor current signal. Figure 9a, b confirms this.

In summary, given that the statistical entity is the trial and the trial behavior entity is the space window frequency, DAP II allows first to get inter-trial differences and where these differences come from (i.e. which space window of which MS component). Then, given the interesting possibility for MCA to consider supplementary points, DAP II allows to show the factor influences. These influences appeared quite different from more usual indicators such as the arithmetic mean or the RMS value. Let us now examine the advantages and disadvantages of our exploratory approach for both designing experiments and analyzing data.

5 Discussion

A given experimental design followed by the analysis of the data produced must be seen as system analysis pair (DoE, DAP), which requires many choices. The DoE literature often considers (1) the number of factors, (2) their levels (both the number and the scale), (3) the way the factor level combinations (FLC) are considered, and (4) the factor influence models [27]. With less than four factors, less than three levels per factor, with the same mathematical scale model for all factors (e.g., a quantitative model), the DoE is rather simple, even though many factor level combination sets may exist.

In all the other cases, the DoE problem involves many choices that are rather difficult to justify. One of them is the

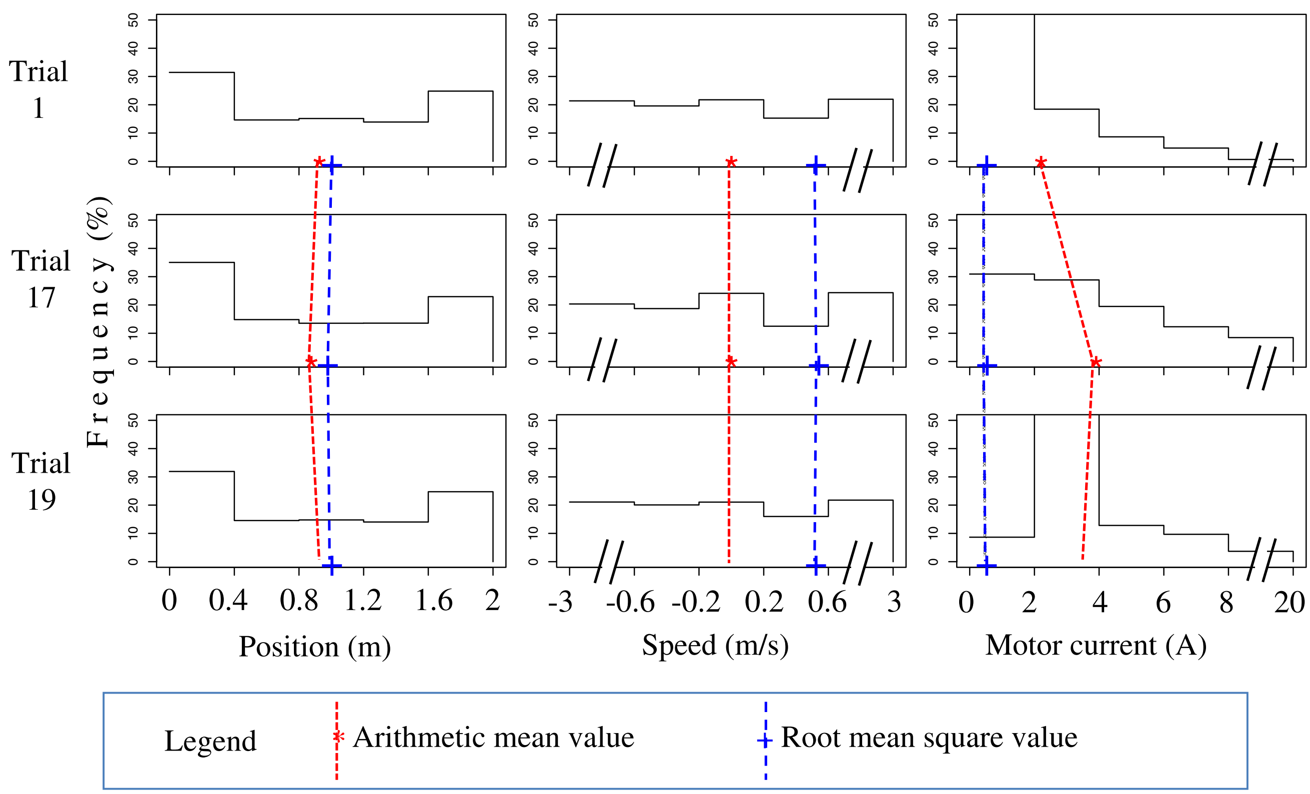


Fig. 8 Example of differences/resemblances pointed out by MCA (Fig. 7). Case of three trials: one situated on the very left of axis 1 (trial 1), one situated on the very right of axis 1 (trial 19), and one situated on the top of axis 2 (trial 17)

FLC chronology. With a *single* train door system (only *one* due to its high cost and the length of time to obtain it), one might think that performing I trials does not fulfill the independence hypothesis, physically and statistically (as with I test tubes with I FLC in a chemistry DoE). In the same way, one might think that performing D cycles with each trial i (i=1, …,I), where all the cycles are non-identical, does not fulfill the repeatability hypothesis, physically and statistically.

In our case, we must acknowledge that intra- and inter-trial chronology differences are somewhat higher than those usually practiced. Since the system being studied is intrinsically based on the "opening/closing" cycle, each factor level must be considered only in terms of the relationship with such a cycle's existence. Let a be the basic cycle and b the disturbed cycle. Several ways can be proposed to design a trial. Briefly, let us consider k=2 since the periodicity may be either introduced or not introduced in the cycle sequence. In the first case (k=1), an elementary sequence could be *aaab* or *aabb*, the four-cycle sequence being played T times with $4 \times T = D$. Thus, the generic sequence can be seen as playing A times the cycle a and B times the cycle b with A= 3 and B=1 in the first example and A=2 and B=2 in the second example. The alternative (k=2) is to introduce cycle b randomly within the D cycles.

In this paper, the first solution (k=1) was preferred in order to have more control over the total duration of a trial and to avoid too many disturbed cycles, with the possibility of damaging the door. Given the total length of time to complete an experimental design (3 weeks) and the duration of a type a cycle (t_a=7.5 s) and the duration of a type b cycle (t_b may run from 8 s (for a cycle with only one shock) to 20 s (for a cycle with an obstacle that is present for 12 s)), the total number of trials was adjusted to 29 trials (I=29). The length for each trial i was adjusted to 160 opening/closing cycles (D=160), and the basic sequence was A=9 times the cycle a and B=1 times the cycle b: T=D/(A+B)= 16. Given our openness to criticism of our choices of I, D, A, and B, the chronology of the 29 trials was finally adjusted in order to minimize the maintenance duration from a trial i to a trial i+1 (see Tables 3 and 4).

Many choices are also possible for the DAP, i.e., for the phase quintuplet ($\varphi 1$, $\varphi 2$, $\varphi 3$, $\varphi 4$, $\varphi 5$), this more particularly due to the presence of time data with multivariate, cyclic, and not stationary aspects. First of all, it is worth noting that an inference approach is used in most cases, although there are L=2 two main statistical approaches: introducing a probabilistic model (l=1) vs. not introduced such a model (l=2). In fact, as soon as the data collection step has been finished, researchers (not only engineers, but also physicians or

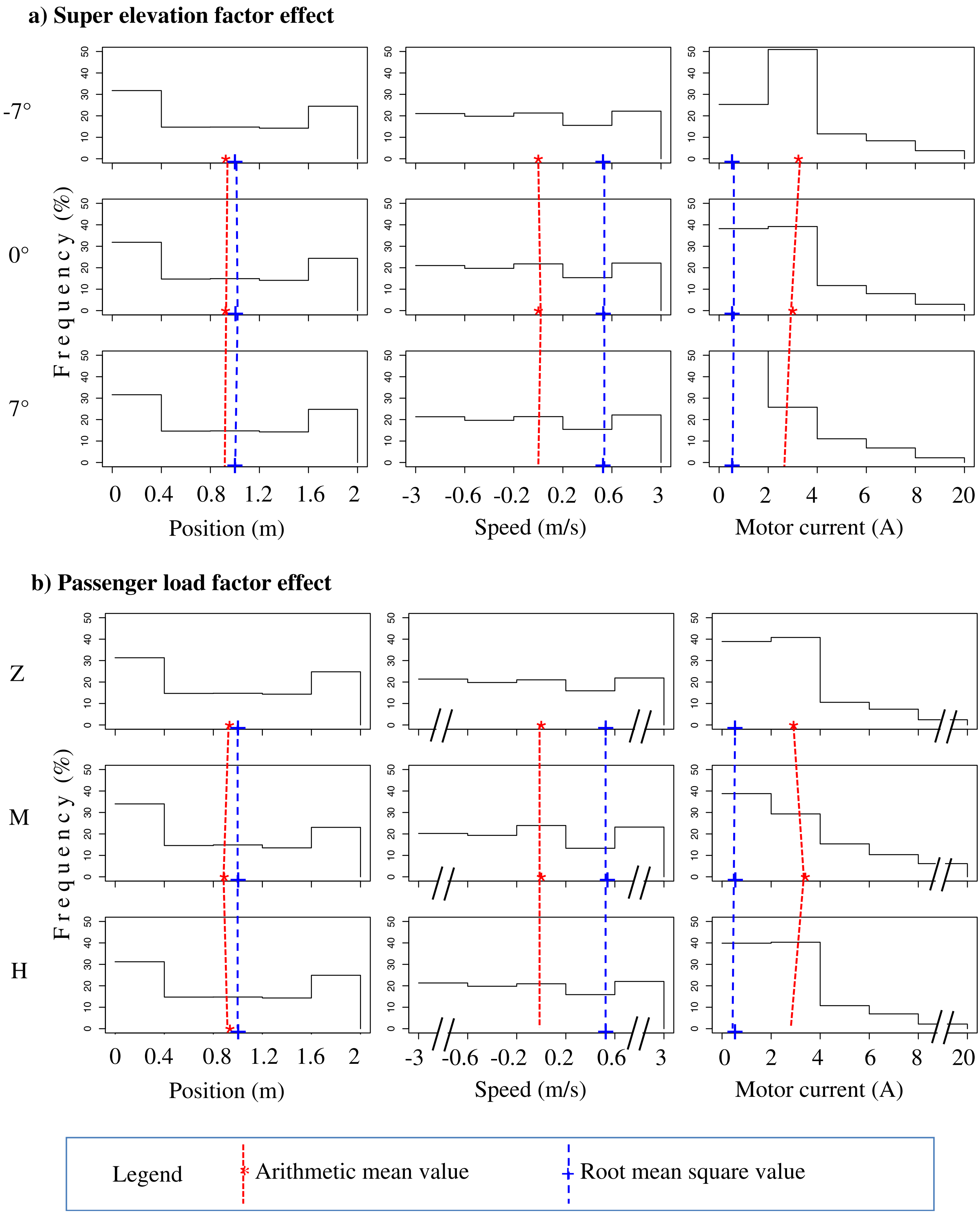


Fig. 9 Example of large and low influences pointed out by MCA (Fig. 7). Case of **a** for the super elevation factor and **b** the passenger load factor

psychologists) usually have a strong desire to assess the factor influences very quickly on indicators computed globally for each trial i (e.g., in our case, one can imagine cycle indicators such as the root mean square value, maximum value or cycle duration, and then with D cycles per trial, the D values are averaged for a given trial i). Given the presence of MS and the possibility of having doubtful data elements, because both the test bench and the door system are new, we chose the descriptive alternative (l=2). In addition, to understand possible links between the MS components, we chose the multidimensional strategy within this descriptive alternative, instead of performing a component-by-component analysis.

Then, many descriptive tactics are available. For instance, due to the presence of cyclical data, the characterization phase (φ1) could have started using Fourier description. However, this technique is much less descriptive than a magnitude histogram description because Fourier supposes a model based on a sinus curve. For this reason, the magnitude histogram approach was used to characterize each cycle. The number S of space windows was chosen as a compromise between a low value, which simplifies the multidimensional analysis but reduces much the information loss inherent to changing from a continuous scale to an ordinal scale, and a high value, which risks complicating the analysis overmuch and may yield windows with very low frequencies. Given this data characterization approach and our multivariate analysis perspective, the scale homogenization phase (φ2) is not required any more (all data pieces are relatives frequencies). Then the remaining DAP (φ3, φ4, φ5) was continued using two ways, yielding DAPs I and II. The first one was chosen essentially to show the time influence but also for the possibility of highlighting outliers: φ3, the data table building phase, yields a matrix where the row corresponds to the cycle and the column to the space window of each MS component.

The second one was chosen to show the differences between the trials and also for the possibility to underscore the factor influences: φ3 yields a two entry table where the row is the trial (thus a set of cycles) which corresponds a given combination (α,β,χ,δ,ε) for the five considered factors.

Since it was necessary to show (1) the similarity/dissimilarity between the statistical units (either cycles or trials), (2) the connections between the MS components, and (3) the relationships between the results for (1) and (2), phase φ4 consisted in a matrix approximation procedure based on singular value [28] decomposition instead of a classification procedure, such as hierarchical clustering, since this technique mainly produces results either of type (1) or type (2). Finally, given the presence of frequency data, the MCA [26] was chosen instead of the usual PCA [25]. MCA output being rather complex (1) series of row points and column points projected onto planes spanned by main axes and (2) series of tables that aid interpretation, a final phase (φ5) that present results more simply was required.

DAP I made it possible to highlight the intra- and inter-trial differences, which mainly came from space windows 1 and 2 for the current component and space window 3 for the position and speed components. DAP II made it possible to highlight the differences between trials only coming from the motor current space windows. Because the role played by the position and speed space windows for positioning the main axes was lacking, a new MCA with more space windows (S increased from 5 to 8) was performed. Again, the position and speed space windows played no major role, confirming that, globally, the system control "black box" was quite well designed: In order to maintain the position and speed programs, the motor current is increased due to the motor current being adapted based on learning feedback from the previous cycles.

Finally, the last phase of DAP II allowed to show the good performance of a space windowing-based characterizing method instead of usual characterizing method. For instance, two histograms are different while two RMS values may be identical (see [25] for information comparison for several time data characterizing techniques). If one must acknowledge that histogram data analysis (e.g., with multiple correspondence analysis) is much complex and longer that standard indicator analysis (e.g., with analysis of variance), this approach much be seen as better starting point for data mining, more particularly if both the test bench and the train door systems are new.

6 Conclusion

The methodology described in this paper is based on an experimental approach that uses design of experiments for reliability growth and determination of the operating limits of the passenger access system. The methodology's innovation comes mostly from using the design of experiments approach to simulate the system behavior in a complex disturbed environment. We developed an original test bench (scale 1:1) and a multivariate descriptive statistical procedure, which guarantee realistic delays and costs allowing a rapid return on investment for the company. Our experimental approach produced satisfactory results that justify the reliability growth technique introduced by Bombardier during the early design phase.

Most tools required to perform exploratory experiments and data analyses being now present, let us evoke some perspective for our research. First, the experimental bench tool can be improved so that electronics/informatics devices can be tested (more and more, transport system problems come not only from the mechanics). Then, thanks to a tool as SAS-JMP and our procedure performed to reduce the time required to jump from one trial to another (Section 3.2.3), the DoE can be achieved more quickly. Nevertheless, one must keep in mind that, from the full experimental design, many optimal experiment designs can be suggested, given the wide possibility of criteria to be optimized and, once these criteria have been chosen, the larger number of real designs that may be performed (since several experimental designs may give rather identical values for the criteria). Finally, the data processing and analysis tools been now developed (here from Matlab and *R*), it is possible to check rather quickly the data and to find the main trends within the multifactor multivariate database, so that the output of the descriptive statistical analysis becomes the input of the inference analysis tools (e.g., from SAS-JMP).

Thanks to all the tools mentioned above, future research prospects may include modeling the degraded behavior to provide information about the lifetime of the passenger access system. This model will be built with the data issued from DoE and from the traceability of failures during commercial operations.

Acknowledgments This research has been supported by Bombardier Transport, the International Campus on Safety and Inter-modality in Transportation, the Nord/Pas-de-Calais Region, the European Community, the Regional Delegation for Research and Technology, the Ministry of Higher Education and Research, and the National Center for Scientific Research. The authors gratefully acknowledge the support of these institutions.

References

1. Hobbs GK (2000) Accelerated reliability engineering: Halt and Hass. Wiley series in quality and reliability engineering. Wiley, Hoboken. ISBN 0-471-97966-X
2. Charki A, Laronde R, Guerin F, Bigaud D, Coadou F (2011) Robustness evaluation using highly accelerated life testing. Int J Adv Manuf Technol 56:1253–1261
3. Xiong C, Zhu K, Yu K (2006) Statistical survival analysis with applications. In: Pham H (ed) Handbook of engineering statistics. Springer, Berlin, pp 347–365. ISBN 1-85233-806-7
4. Nelson W (2005) A bibliography of accelerated test plan. IEEE Trans Reliab 54(2):194–197
5. Nelson W (2005) A bibliography of accelerated test plan. IEEE Trans Reliab 54(3):370–373
6. Elsayed EA, Liao HT (2004) Reliability prediction and testing plan based on an accelerated degradation rate model. Int J Mater Prod Technol 21(5):402–422
7. Goupy J (1996) Unconventional experimental designs theory and application. J Chemom Intell Lab Syst 33:3–16, Elsevier
8. Paul RA, Tsai WT, Chen Y, Fan C, Cao, Z (2006) End-to-End Testing and evaluation of High-Assurance Systems. In: Hoang Pham Editor. Handbook of Engineering Statistics, Springer, 443-474. ISBN 1-85233-806-7
9. SAS Institute Inc (2009) JMP 8 design of experiments guide, 2nd edn. SAS Institute, Cary. ISBN 978-1-60764-303-6, 269p
10. Özsoy S, Celik M, Kadioglu FS (2008) An accelerated life test approach for aerospace structural components. Eng Fail Anal 15:946–957
11. Mohammadiam HS, Aït-Kadi D (2010) Design stage confirmation of lifetime improvement for newly modified products through accelerated life testing. Reliab Eng Syst Saf 95:897–905
12. Pascual F, Meeker WQ, Escobar LA (2006) Accelerated life test models and data analysis. In: Pham H (ed) Handbook of engineering statistics. Springer, Berlin, pp 397–425. ISBN 1-85233-806-7
13. Elsayed A. Elsayed (2003) Accelerated life testing. In: Pham H (ed) Handbook of reliability engineering. Springer, Berlin, pp 415–428. ISBN 1-85233-453-3
14. Tang L.C.(2003) Multiple-steps step stress. In: Pham H (ed) Handbook of reliability engineering. Springer, Berlin, pp 441–455. ISBN 1-85233-453-3
15. Xiong C (2003) Step-stress accelerated life testing. In: Pham H (ed) Handbook of reliability engineering. Springer, Berlin, pp 457–469. ISBN 1-85233-453-3
16. Mohammadiam HS, Aït-Kadi D, Coulibaly A, Mutel B (2009) A contribution to accelerated testing implementation. In: Martorell S et al (eds) Reliability, risk and safety: theory, methods and applications. Taylor & Francis Group, London. ISBN 978-0-415-48513-5
17. Nelson W (1990) Accelerated testing: statistical models, tests plans, data analyses. Wiley, New York
18. Goupy J, Creighton L (2007) Introduction to design of experiments with JMP examples, 3rd edn. Cary: SAS. ISBN 978-1-59994-422-7
19. Lin DK (2003) Industrial experimentation for screening. In: Khattree R, Rao CR (eds) Handbook of statistics. Elsevier, Amsterdam, pp 2–73. ISBN 0-444-50614-4
20. Fedorov VV (1972) Theory of optimal experiments. Academic, New York. ISBN 0122507509
21. Gauchi JP (2005) Plan d'expériences optimaux: un exposé didactique. Modulad 33:139–162
22. Dette H, Wong WK (1998) Bayesian D-optima designs on a fixed number of design points for heteroscedatic polynomial models. Biometrika 85:869–882
23. Goupy J (2002) Modélisation par les plans d'expérience. Technique de l'ingénieur 2002;R 275
24. Turgis F, Copin R, Loslever P, Cauffriez L, Caouder N (2010) Design of a testing bench for simulating tightened-up operating conditions of train's passenger access. In: Bris R, Guedes Soares C, Matorell S (eds) Reliability, risk and safety: theory and applications. Taylor & Francis Group, London. ISBN 978-0-415-55509-8
25. Loslever P (2001) Obtaining information from time data statistical analysis in human component system studies. Inf Sci Methods Perform 132:133–156
26. Benzecri JP (1992) Correspondence analysis handbook. Marcel Dekker, New York
27. Jobson JD (1992) Applied multivariate data analysis. In: Regression and experimental design, vol 1. Springer, New York
28. Jobson JD (1992) Applied multivariate data analysis. In: Categorical and multivariate methods, vol 2. Springer, New York