Data-driven decision-making for IT capacity: beyond statistical analyses

Michel Lutz, Xavier Boucher

Data-driven decision-making for IT capacity: beyond statistical analyses

M. Lutz[a] and X. Boucher[b]

[a]TOTAL, Puteaux, France; [b]Department GEO, Université de Lyon, Mines Saint Etienne, Institut FAYOL/UMR EVS, Saint-Etienne, France

This paper reports a research work piece developed in collaboration with the semiconductor wafer production company: STMicroelectronics. This collaborative research programme aimed at implementing statistical methods, so as to improve business decisions focusing on capacity planning for information technology. The current paper presents the specification, and experimentation of a method dedicated to managing the need to integrate rigorously complex contextualisation factors, when developing statistical-based Decision Support Systems (DSS). The key challenge is to increase the end-user acceptance and success rate for DSS developments. To ensure the successful integration of DSS within the user environment, the paper formalises a so-called 'Contextualisation process', integrated within a larger decision-aid development framework. This Contextualisation process is specified with three methodological components, respectively 'Qualitative Contextualisation', 'Statistical Modelling and Formalisation' and 'User Integration'. This approach is applied by STMicroelectronics, to a case study focusing on a project of change management for the infrastructure of the Information System. Based on the demonstration of the case study results, the added-value of the Contextualisation process are discussed and further perspectives for applied research in DSS are drawn.

1. Introduction

This paper presents some results of applied research in the field of Decision Support Systems (DSS) applied to Information Systems (ISs), developed in collaboration with a semiconductor manufacturing plant by the company STMicroelectronics.[1] The overall objective of the research work was to define a decision-aid process, aimed at better managing the coupling between the industrial manufacturing system on the one side and the Information Technology (IT) infrastructure on the other side. This problem of coupling between industrial and information processes is particularly important when managing the frequent changes affecting manufacturing plants in the microelectronics industry (Petruzzi & Garavelli, 2007). In such industrial environments, ISs highly contribute to the efficiency of industrial production (Barkia, Boucher, Le Riche, Beaune, & Girard, 2013; Cardinali, 1992; Gowan and Mathieu,

1

Gowan & Mathieu, 1996; Izza, 2009; Jardim-Goncalves, Grilo, & Steiger-Garcao, 2006), notably for wafer production (Hong, Lim, Ceong, & May, 2012; Izza, Vincent, & Burlat, 2008; Leachman & Hodges, 1996). The performance of ISs depends on the good or bad adaptation of both IT architecture and infrastructure (Earl, 1989). The responsibility of designing and dimensioning these IT architectures and infrastructure relies on decision-makers, called IT managers in the rest of the paper. Such decisions are in the field of IT capacity planning, as mentioned in ITIL best practices (Klosterboer, 2011): IT managers are in charge of ensuring that ISs provide enough functional capabilities and quantitative capacities to support current and future industrial activities of wafer production. At STMicroelectronics, IT Managers constantly check that the present and future workloads of IT resources remain below the maximal available capacities.

In this perspective, the aim of the research was to try and develop a statistical-oriented approach, to provide IT managers with some decision-aid support in anticipating correctly the need for capacity, and then in taking the right decisions in IT technology dimensioning. More precisely, this paper tackles the issue of managing the contextualisation of the DSS to be developed. We underline that the pure statistical development, which is at the heart of the decision-aid, has been previously published in (Lutz, Boucher, & Roustant, 2012) then (Lutz, Boucher, & Roustant, 2013). Thus, the objective of the current paper is not to revisit the building and validation of the statistical model, but to integrate these results within their operational decision-aid context. Based on existing theories and frameworks in decision sciences, a comprehensive decision-aid process is proposed, integrating a specific contextualisation process formalised in this paper, with three key steps: first, the general context of the decision and the decision problem are analysed then formalised; the second stage consists of providing quantitative answers to the decision problem, based on the statistical model mentioned above; the third stage integrates the statistical results in a specific user-oriented decision-aid analysis, with the objective of providing the decision-makers with reliable, contextualised and adapted answers to their decisional needs. Finally, the last goal of the paper is also to demonstrate the validity of the decision-aid results developed in the context of STMicroelectronics application and to analyse the reactions of the final users.

The rest of this article is organised as follows. Section 2 gives more details about the scientific literature concerning the utilisation of statistical methods for IT capacity planning. We also underline the added-value of a general framework for decision-aid proposed by Tsoukiàs (2007). Section 3 presents the methodological contributions of the paper: the specification of a contextualisation process for the development and implementation of DSS dedicated to IT capacity planning, based on the articulation between qualitative and quantitative modelling approaches. This contextualisation framework is then applied to a specific case study at STMicroelectronics. The three key steps of the process are applied in Section 4, and the final results of the statistical developments are demonstrated. The added-value of the overall case study is discussed in Section 4.5. The final conclusions and perspectives highlight broader lessons learned for applied research in DSS.

2. Using statistical methods for IT capacity planning: towards a comprehensive decision-aid process

Information Technologies management are complex and multidimensional subjects requiring different kinds of decision supports (Bozdogan & Zincir-Heywood, 2012): some key

examples can be found with the management of IS failures for (Bozdogan, Zincir-Heywood, & Gokcen, 2013), the management of IT change projects with Alkuraiji, Liu, Oderanti, Annansingh, and Pan (2014), or innovation decision supports developed by Shrestha, Cater-Steel, and Toleman (2016). Capacity management (Bauer, 2015) is a specific area of IT management, which is more specifically based on the use of quantitative approach, called capacity planning approaches. This decision-making domain addresses the design of IT infrastructures (Nelson, 2002) which can be further refined into more specific issues: infrastructure change management planning (Rebouças, Sauve, Moura, Bartolini, & Trastour, 2007), alignment of IT infrastructure on business needs (Bartolini, Stefanelli, & Tortonesi, 2011), or estimation of the optimal capacity required (Kim & Kim, 2014). Recent trends in IT capacity planning take into account the new features of cloud infrastructures (Ghosh, Longo, Xia, Naik, & Trivedi, 2014). Throughout these general references, various types of techniques are used, depending on the precise context: clustering approaches, simulation, optimisation algorithms, etc. In coherence with the orientations developed at STMicroelectronics, the following section puts more specifically the focus on data-driven models for capacity planning.

2.1. The diversity of data-driven IT capacity planning models

Quantitative IT capacity studies are more and more grounded on data analysis methods (Allspaw, 2008; Bartolini et al., 2011; Domanski, 1999; Gunther, 2007; Padonou, Roustant, & Lutz, 2015). Indeed, data-driven approaches are particularly well-adapted to manage complex systems where planning time is limited (Ahituv & Igbaria, 1988; Gunther, 2007). To provide an overview of the multiple data-driven IT capacity planning models, which can be found in the dedicated literature, we propose a classification with two dimensions.

Firstly (dimension 1), the objectives of the statistical models can differ. Mainly, two distinct purposes can be observed:

- *Models dedicated to IT systems performance management*. Their goal is to analyse the system's response times, queuing phenomenon, potential bottlenecks, etc.
- *Models dedicated to IT systems capacity management*. Here, the objective is to understand and anticipate the use of the IT system components: CPU and memory used, database sizes, etc.

Secondly (dimension 2), the methods used to build the statistical models can vary:

- *Models constituted with classic statistical approaches*: usual methods, such as regression or time series modelling are implemented through a human-based and manual process (selected data choice, 'handcrafted' analysis, etc.).
- *Models designed through algorithmic approaches*: automated methods are developed for large-scale analysis of big data-sets.

The mindsets behind both approaches are very different. The first one puts the focus on each variable of the model, on the meanings of the statistical coefficients, etc.: the model should be linked to a 'physical understanding' of the functioning of the system. On the contrary, algorithmic approaches admit a black-box view of the system, rather interested in discovering general patterns in the data available.

Table 1. Data analysis IT capacity planning literature: suggested classification and selected references.

		Dimension 1 **Objective**	
		Performance	*Capacity*
Dimension 2 **Methods**	*Classic statistics*	Kelly (2005)	Allspaw (2008); Ahituv and Igbaria (1988); Browning (1994); Gunther (2007); Jain (1991)
	Algorithms	Aguilera et al. (2003); Jiang et al. (2008)	Jiang et al. (2008)

As summarised in Table 1, dimensions 1 and 2 are used to classify the literature dedicated to IT capacity planning. For instance, Kelly (2005) uses a linear regression model to explain the performance of a software application, using the volumes of transaction calls as predictors. Aguilera, Mogul, Wiener, Reynolds, and Muthitacharoen (2003) use generic data-mining algorithms to discover causal paths in complex IT architecture, useful for performance improvements. These distinctions apply to capacity management too: Ahituv and Igbaria (1988) use linear regression modelling to anticipate IT components' activities, Browning (1994) uses time series theory, whereas Jiang, Chen, and Yoshihira (2008) develop automated invariant algorithms usable for IT capacity planning.

2.2. Limits in the operationalisation of statistical models

Despite the large range of data analysis approaches for IT capacity planning, some clear limitations on the operational implementations in professional fields can be raised (Bartolini et al., 2011). First of all, it should be mentioned that, when our research started at STMicroelectronics, data analysis was almost unused in the company to manage IT capacity. Indeed, data were collected, sometimes visualised, but neither analysed with statistical nor algorithmic approaches. Banerjee and Igbaria have already pinpointed such a lack of analysis in professional practices in 1993. Moreover, in spite of a lot of industrial frameworks designed to support the management of IT systems (22 major frameworks according to Alford et al., 2006), no concrete guidelines for data analysis implementation are provided for professionals, even if their use is recommended. This limitation is widely discussed in Lutz et al. (2013). Consequently, it can be said that analysing data is never a straightforward task for IT managers. Klosterboer (2011) summarises this challenge: 'You may not be able to make a direct correlation between the business trends and your IT capacity, but you should at least look at whatever your organisation publishes and see how it compares your component and service trends'. Moreover, constructing a model is not sufficient. As mentioned by Bartolini et al. (2011), investigating the human component in decision-making is an important step. This means being able to perform 'what-if scenario analysis', useful for managers. Thus, we highlight below two major challenges.

First, Table 1 illustrates that the data-focused literature does not propose any general framework to design an IT capacity model (first challenge). Rather, a collection of tools and heterogeneous solutions are found. All are idiosyncratically-designed, according to specific application backgrounds. Such local contexts wholly influence models' forms and variables. This is fairly comprehensible. Indeed, regarding the complexity of today's IT systems and the tremendous diversity of management situations, models as well as methods and applications are necessarily diversified and overabundant. Consequently, IT operations managers

are confronted with the following question: when conducting capacity planning studies, what kind of model and method should we select? Answering this question strongly depends on the internal ability to understand the application background of a data-driven tool. Before jumping into data manipulations, managers should thoroughly understand, clarify and, if necessary formalise, what precise decision-making aid is expected. Understanding the decisional context is a crucial preliminary exercise, essential in the process of designing a statistical IT capacity planning model.

A second challenge can be raised. Table 1 pinpoints many types of models. Even if we assume that the operation managers are able to select a proper form and build a relevant model, is it satisfactory enough? Beyond the development of a validated statistical model, comes a further question concerning the usability of the model, within its operational decision-making context. Generating a good model is of course a good and necessary start. But this first result, mainly based on statistical skills, remains insufficient: how to convert the mathematical model into an operational item, fully usable to help making *in situ* decisions? This final utilisation of a model also depends on a broad understanding of the decisional environment and processes, including the awareness of the preliminary application context. Complementary competences and methods, not focusing on statistical modelling, are required.

To deal with both challenges, it appears necessary to step back from the data-analysis work and analyse the whole decision-aid need. Thus, a comprehensive and fully operational IT capacity decision-aid process is required. This process should not only be focused on the statistical modelling phase. Rather, it should embrace all the stages of the decision-making, including the analysis of the application context and the usability of the mathematical model. This issue provides the scientific problematic of this paper: How could we conceptualise a contextualisation method, which would be integrated within a comprehensive and fully operational IT capacity decision-aid process, based on data analysis?

2.3. Towards a contextualisation process for decision-aid development

It has been explained that the question of statistical modelling is not sufficient for effective and concrete decision-making. A more comprehensive process is required, including all the stages of the decision: application context analysis, utilisation of the model. Even if a lot of scientific literature is available on decision-aid development, very few methodologies really integrate both the mathematical and quantitative modelling required for a concrete DSS development and the qualitative analysis on the collective rationalities and action logics, as required by Roy (1993, 2006), Le Moigne in Alcatras, Gianfaldoni, and Paché (2004) and Bouyssou, Dubois, Pirlot, and Prade (2006). Most of the approaches only deal with a partial vision. Tsoukiàs in Bouyssou, Dubois, et al. (2006), Bouyssou, Marchant, Pirlot, Tsoukiàs, and Vincke (2006) and Tsoukiàs (2007) proposes one of the rare methodologies for such integration. His work formalises a valuable theoretical framework derived from the basic assumption that a decision is based on an evaluation model. This model is linked to the problems and preferences of the people involved in the decision-making. Understanding these problems and preferences is an integral part of the decision-aid. Four artefacts can be used to analyse and sequence the decision-aid:

- *Representation of the problem situation* [P^2], which should help to clarify the position of the decision-makers and the role of the analyst in the decision process. It can be represented with the following items:
  - People involved [*A*].
  - Objects, challenges, stakes of all people in the decision process [O].
  - Resources brought by people to commit to the objects (knowledge, financial resources, decision-making power…) [S].
- *Problem formulation* [τ]: based on the representation of the problem situation, a restricted description of the studied reality is expressed, in order to reduce it to a clearly defined decision problem. A decision-aid method could then be applied to this circumscribed piece of reality. Tsoukiàs proposes to describe this artefact as follows:
  - Potential alternatives (actions) to be considered in the context of the representation of the problem situation [*A*].
  - Perspectives under which the potential actions should be observed, analysed, compared [V].
  - Decision problem statement, expected ways in order to evaluate the potential alternatives (choice, sorting, description…) [Π].

When approved by the different stakeholders, it means that the context of the decision is clarified. It helps switching from a fuzzy understanding to a rigorous description of the decision problem. A formal decision-aid can therefore be applied.

- *Evaluation model* [M], which is the usual acceptation of a decision-aid model, as found in many scientific approaches: operations research, game theory, artificial intelligence… and of course, statistics. It can be described with the following:
  - Alternatives to which the evaluation model will be applied [*A*].
  - Dimensions (attributes) to consider the alternatives [D].
  - Scales of each item of the various dimensions considered [E].
  - Criteria to evaluate the alternatives, in order to take into account the preferences of the decision-maker [H].
  - Uncertainty associated with the available information, related to the dimensions of the criteria to evaluate [U].
  - Procedure to aggregate the preferences, measures and uncertainties, in order to obtain synthetic information about the elements associated with the alternatives [R].
- *Final recommendation*: the evaluation model will lead to a result [Θ] and the final recommendation is its translation into natural language. This is a return to reality, providing an operational answer for decision-aid, consistent with the organisational context of the decision problem and usable by managers.

By working with the decision-makers and the people that may provide valuable input to decisions, the analyst constructs a common representation of all artefacts of the decisional process. This analyst, in charge of supervising the development of the final decision-aid required, could be an internal or external consultant, a project manager, a researcher …

This theoretical framework proposed by Tsoukias provides a general and clear description of the different stages of any decision process and can be potentially applied to analyse many situations. The objective of this paper is to fill the gap between this theoretical proposal and the operational context of DSS development, specifically for the field of IT capacity
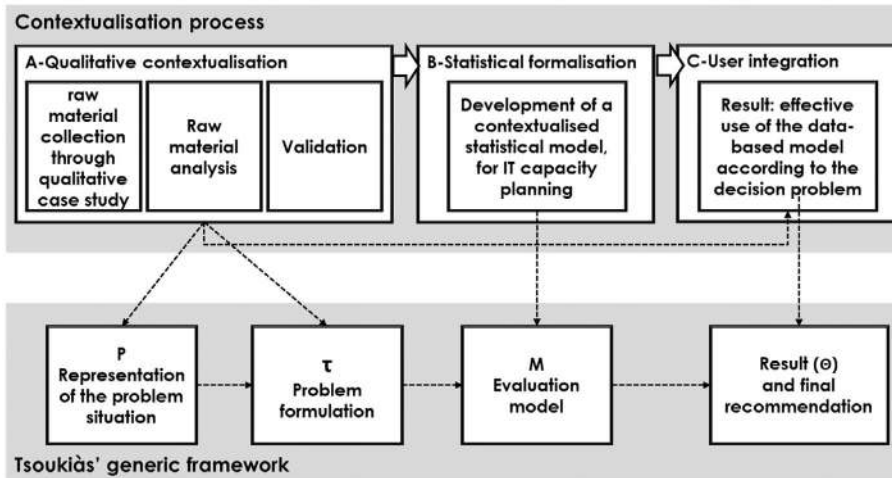
Figure 1. Overview of the decision-aid process.

planning decisions. In this perspective, Section 3 develops several methodological contributions articulated with Tsoukias's framework. Section 4 will illustrate an application to support concrete IT capacity planning decisions at STMicroelectronics.

3. Methodological contributions of the paper

Considering Tsoukias' framework as a generic basis to develop decision-aid systems, the objective of this paper is to develop an additional methodological layer, dedicated to support the operational application of the framework within the context of real industrial IT capacity planning situations. As synthesised in Figure 1 we call this new methodological layer the 'Contextualisation Process', constituted by three methodological components 'Qualitative contextualisation (A)', 'Statistical modelling and formalisation (B)' and 'User integration (C)'. This approach aims at answering the key challenges highlighted in Section 2.2, for the field of IT capacity planning.

The key components of this contextualisation process are introduced below. They are further developed in Section 4, illustrated by a case study (Figure 2).

3.1. Qualitative contextualisation

The 'Qualitative contextualisation' step consists of using a qualitative case study building procedure, in order to generate the representation of the problem situation [P] and the problem formulation [τ]. The objective is to avoid risks, arising when developing statistical tools, of jumping too fast to mathematical approaches: this usual mistake leads to designing mathematical models fitting to the mathematical analyst's habits and preferences, rather than meeting concrete managerial decision-aid needs. To draw a clear and concrete definition of [P] and [τ], the method uses a qualitative case study building procedure. According to Thietart et al. (2007) and Coutelle (2005), case study methodologies are well-adapted to understanding a social phenomenon (decisions about IT capacity planning) in its natural
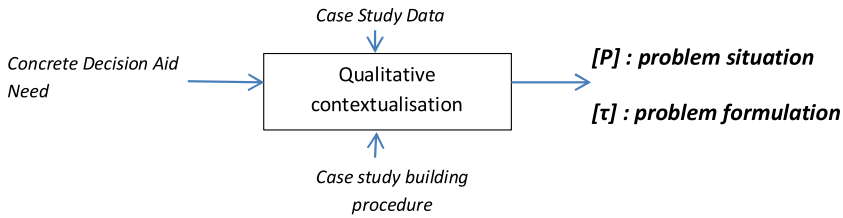
Figure 2. Objectives of the qualitative contextualisation step.

environment (a microchip production plant). Referring to Thietart et al. (2007), Coutelle (2005) and Wacheux (1996), a qualitative procedure has been structured into three steps:

- *Step 1: Raw material collection*. Interviews with people involved in the decision process (decision-makers, people that may provide valuable inputs to decisions) are conducted by the analyst. A semi-structured interview method is used, to favour discussion without being too directive. Information collection is based on an interview guide, based on generic structure, designed to collect precise expressions of the problem situation and the problem formulation. Interviews are recorded and transcribed in *verbatim* reports. During the interviews, additional materials such as emails or meeting minutes may be collected as well.
- *Step 2: Information analysis.* All interviewees may have different viewpoints about the decision problem. A whole picture of the decisional situation, articulating these multiple point of views, has to be drawn up by organising and structuring all information collected. The main themes raised during the interviews are identified, categorised, and the categories are linked to Tsoukiàs' theory and formalisms. This analysis, often named 'reflexive' analysis (Lejeune, 2010) helps in extracting and synthesising a proposal of collective meaning of the decisional situation. At the end of this step, a full synthesis of the case studied can be proposed, including a complete description of [P] and [τ].
- *Step 3: Validation*. A common acknowledgement of the study has to be obtained amongst all the interviewees. Here again the approach is qualitative: through a validation workshop, the synthesis is communicated and discussed, potentially updated and finally approved by decision-makers. This final output should describe as well as possible the reality perceived by all people implicated in the decision-making.

3.2. Statistical modelling and formalisation

Based on the IT capacity problem formulation resulting from the qualitative contextualisation, this second component of the contextualisation process consists of developing a usable quantitative model answering the requirements of the problem formulation. A smooth transition between the problem formulation [τ] and the evaluation model [M] should be ensured. Building [M] is not just a matter of mathematical and modelling skills. It is necessary to have a deep understanding of the context of the data analysis to be performed and to build an appropriate evaluation model. [M] is built upon two types of contextualised information: (i) large amounts of data concerning the functioning of IT resources and
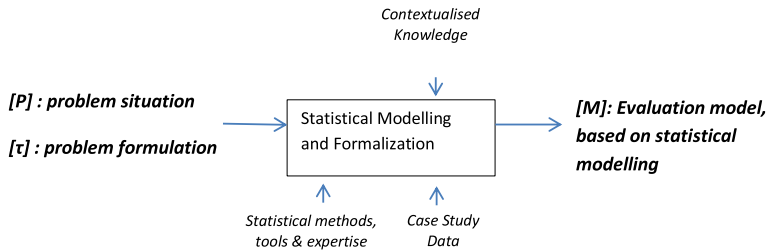
Figure 3. Objectives of the statistical modelling and formalisation step.

processes which are part of the case study IS; (ii) specific expertise of the actors (called 'experts') which is integrated within the quantitative modelling procedure (Figure 3).

Confronted with a vast amount of available data and variables, a structured modelling procedure is required, to efficiently converge towards useful and practical models. A bottom-up macroscopic process is suggested. It starts from the predicted IT resources and focuses only on the main influencing variables. Two modelling sequences are presented to find out these influencing variables: exploration and explanation. These sequences are divided into several modelling steps, mixing qualitative and quantitative work. Qualitative work involves the context experts, through task groups. According to the sequence, quantitative work uses exploratory statistics or forecasting methods.

Sequence 1: exploration

Step 1.1. Initialisation

Data related to a specific variable, for which a capacity model has to be built, are extracted, eventually corrected, and visualised.

Step 1.2. Map of the IT application layer

IT applications and sets of software, which are supported by the IT resources considered, are listed and mapped with the help of the experts.

Step 1.3. Expressing the variability of the application layer

The assumption here is that, for an IT capacity problem, the application activity is the principal forecaster for the IT resource activity. As application layers might encompass hundreds of interrelated variables, it is crucial to have a synthetic overview of this activity, to grasp the main sources of variability and summarise relationships between application variables. Multivariate analyses are used for this purpose.

Sequence 2: Explanation

Step 2.1. Finding explanatory business-related variables

The outcome of step 1.3 is the construction of synthetic variables, expressing the application layer variability (usually called 'components' or 'latent variables', according to the kind of multivariate analyses used). Such synthetic variables can be considered as underlying unobserved constructs, consequences of the plant's industrial activities. These synthetic variables are discussed and interpreted with the experts. When potential explanations are found, measurable explanatory variables (via the capacity management IS or mined in the production activity databases) should be connected. Such variables should be as much as possible foreseeable, as well as understood by production managers.
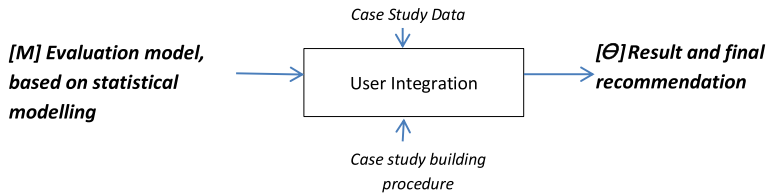
Case Study Data

[M] Evaluation model, based on statistical modelling

User Integration

[Ө] Result and final recommendation

Case study building procedure

Figure 4. Objectives of the user integration step.

Step 2.2. Predictive modelling

The identified variables (IT resources and measurable explanatory variables) are quantitatively connected, through a statistical model.

Step 2.3. Validation and improvement

The model is presented to the experts and confirmed. Possible improvements are discussed.

For the interested reader, a detailed formalisation and application of this statistical modelling approach has been presented in Lutz et al. (2012).

3.3. User integration

A final transition has to be achieved between [M] and [Ө] to ensure the end of the contextualisation process. Even when a model is properly built, its use for decision-making is rarely straightforward. What should be done with the model? Analyse its parameters? Use it for simulations? These answers can only be provided with a deep understanding of the problem formulation [τ]. The analyst should carefully favour this linkage and ensure that any result coming from the use of a quantitative model effectively fits a decisional issue raised during the contextualisation work. The approach developed to correctly manage the integration of use will be developed in the case study of Section 4 (Figure 4).

As aforementioned, using numerical methods to make decisions related to IT capacity planning is a challenging task. Without any surprise, STMicroelectronics IT managers were struggling building such quantified approaches to support their daily activities. That is why the presented research work was initiated. Our objective was therefore to build an operational but scientific decision-aid, thanks to the rigorous contextualisation process described above, designed specifically for STMicroelectronics purpose. Section 4 demonstrates the application and results of this methodological approach, to give more details on the three components A, B, C and show their concrete application for a microelectronics industry IS. More general lessons learned for DSS applied research are also discussed.

4. The contextualisation process applied to an industrial IS case study

Based on the STMicroelectronics case study, this section illustrates the concrete application of the three components of the contextualisation process: 'Qualitative contextualisation (A)', 'Statistical Modelling and Formalisation (B)' and, finally, 'User integration (C)'.

STMicroelectronics microchip plants are highly automated and most manufacturing operations are computer controlled or aided. Consequently, IT capacity failures mean production downtime or loss of efficiency. It is, thus, fundamental to plan carefully IT system capacities and the criticality of the ISs emphasises the essential decisional challenges for STMicroelectronics IT managers: How can they ensure an adequate future capacity of the information technologies, able to provide continuous support for changing manufacturing activity? In Lutz et al. (2012), it was explained that sizing and IT investment planning were performed on the basis of qualitative estimates provided by IT experts. Despite a clear internal demand to base these judgements on rigorous quantitative models, STMicroelectronics was not able to organise a complete quantitative study.

The case study focuses on an earlier major STMicroelectronics project affecting the IS: a major change in the production Manufacturing Execution System (MES) IT infrastructure. Many levels of the IT architecture were impacted: software upgrades, new Oracle database versions, change of the physical servers. This project was highly business-critical. Indeed, as the MES system represents the core of the industrial IS, any poor definition of the new infrastructure could have destabilised the manufacturing processes. The project was complex and long term: it was completed in a period of around two years and involved several business and IT teams, with sometimes conflicting interests. Many decision problems occurred successively. As proposed in this paper, applying the suggested decision-aid process will help to clarify this complicated decisional context.

Statistical formalisation is part of the decision-aid process proposed in this paper. This statistical study on IT capacity analysis has been performed on a delimited subsystem of the overall semiconductor manufacturing IS: this subsystem focuses on the MES. The MES is the core of the production activity. It supports two major applications: one used to manage all the business rules (BR) applied during the production processes (more than 500 BRs, which could be considered as sub-programmes) and one used to manage the maintenance of production equipment. This functional layer is supported by a multilayer hardware infrastructure, including application and database servers. At the upper level of the model, the business activity variables commonly used are the number of moves (corresponding to the transition of an operation of a production process to another) and the number of wafer-outs (corresponding to the number of production processes completed).

The first stage consists of analysing qualitatively the decision processes, throughout the IS project mentioned, using a set of expert interviews. To take into account the a posteriori dimension of the study (the decision processes are studied for a 'past' project of the company), the interview guide has been slightly adapted for the raw material collection (qualitative case study/step 1). To rediscover past decision processes, Nutt's recommendations (Nutt, 1993, 1998) were followed. It consists of a systematic use of questions such as 'what happened before/next?'. Six people were interviewed: this was as much as was necessary to saturate the case analysis. The collected material was then manually analysed (qualitative case study/step 2). We discovered that the project was broken down into a succession of interdependent decisions. These different stages of the project were structured through a
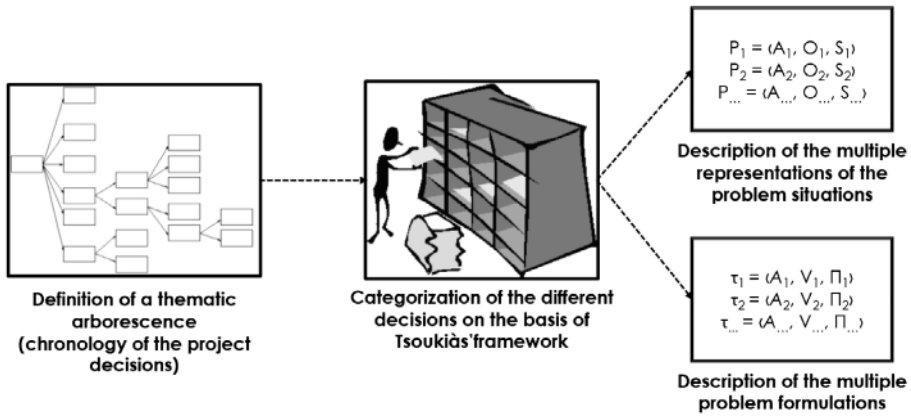
Figure 5. Analysis of the raw material collected, applied to a retrospective case study.

thematic arborescence. Consequently, it was possible to provide a chronological view of the multiple decision purposes. This thematic arborescence was then categorised, with the aim of describing the multiple representations of problem situations and problem formulations. This work is summarised in Figure 5.

After this analysis, a final workshop was organised with relevant people to validate the analysis (qualitative case study step 3). When done, a full description of the decisional process was clearly described and made available for quantitative decision-aiding. The case study constitution helped in understanding with accuracy all the different steps of the whole decision process. The MES project was defined by two principal phases:

- *Project initiation*: to identify the need of the project, define a rough appropriate technical solution and the project management strategy.
- *Project realisation*: to define in detail the IT solution, implement and test it, then perform the move to production.

Several specific decision problems were identified within those phases:

- *Project initiation decision problems*:
    - Identification of distinct risks of IT capacity failure.
    - Definition of a change scenario for the IT architecture and infrastructure: software, middleware and hardware upgrades.
    - Financial support, procurement and legal concerns.
    - *Project realisation decision problem*:

    - Project management (development, testing, training, release plan, etc.).
    - Hardware set-up, installation, migration scenarios.
    - Move to production.

All of these problems occurred throughout the two-year MES project. Of course, each of these steps involves distinct decisional concerns and not all of them concern IT capacity nor require statistical approaches. Nevertheless, all were discussed by task groups. As a result, the conclusion underlined that the most challenging ones were 'Identification of any risk of IT capacity failure' and 'Definition of a change scenario for the IT architecture and

Table 2. 'Definition of a change scenario for the IT architecture and infrastructure': problem situation description.

[A] People involved	Managers of the IT architecture team	Technical experts of the IT architecture team	MES application team leaders	MES application team technical experts
[O] Objects, challenges	Ensure a global IT architecture, throughout the factory, to furnish an IT service fulfilling present and future industrial needs		Ensure an MES application architecture, to furnish an IT service fulfilling present and future industrial needs	
[S] Resources	Final decision-maker recognised by the company; Technical expert support; Relationships with applications team leaders; IT capacity management database	Technical knowledge (Oracle, UNIX, hardware…); IT capacity management database	Technical expert support	Technical knowledge (application); Business application databases

infrastructure', especially to size properly the future MES hardware servers (memory, CPU …). Not surprisingly, they are both related to IT capacity planning.

4.2.1. Representation of the problem situation [P]

Each of the decision problems identified above could lead to a representation of the problem situation, then a deeper formalisation. In this paper, we will restrict the scientific formalisation and contribution to the decision situation 'Definition of a change scenario for the IT architecture and infrastructure', by showing how a quantitative decision-aid can be provided to size the CPU of the future MES servers. However, let us mention that numerical solutions were also proposed for 'Identification of any risk of IT capacity failure': an interested reader may refer to Padonou et al. (2015).

A more precise formal description of the contextualisation artefacts for 'Definition of a change scenario for the IT architecture and infrastructure' has to be provided. The qualitative case study presents an informal communication process between different stakeholders, in order to size the IT components. Managers of the IT architecture team are responsible for the decision. They work with technical experts of their team (Oracle, UNIX, hardware experts …), but also with close interaction with the MES application team leaders, also supported by their own experts. Each team brings its own resources, but also its own stakes. For instance, the application team will be focused on application performance, whereas the IT architecture team will be focused on maintaining the global consistency of the whole IT architecture throughout all the factory (for instance, the hardware components that will guarantee the best response times of MES application are not necessarily the best choices to ensure the overall consistency and performance of the entire industrial IS). Table 2 provides an overview of the different items of the problem situation for the decisional step 'Definition of a change scenario for the IT architecture and infrastructure'.

Two resources should be noticed: the IT capacity management database of the IT architecture team and the business application databases of the MES application team. They correspond to the two kinds of data sources presented and used in Lutz et al. (2012). The data analyses presented in this paper come from the same sources.

In the scope of this paper, we consider the following IT capacity decision: sizing the CPU of the future MES servers. The underlying decision problem can be better formalised thanks to the qualitative case study, which helps to describe the different components of Tsoukiàs' problem formulation.

[A] are the potential alternatives to be considered. Basically, it relies on the following question: given a horizon h, will it be necessary or not to upgrade the MES servers and, if yes, to what extent? A proper answer ensures an adequate sizing, to avoid any lack of capacity (IT service disruption), but also overcapacity (unnecessary extra cost, technological risk due to too recent IT components …).

To observe, analyse and compare the potential alternatives [V], the qualitative case study helps showing that managers are trying to evaluate an 'increasing factor' of the need in IT resources, at the horizon h. Let us denote it I_h. In absolute terms, I_h can vary from one to the infinite (steady to unlimited IT resource need). In concrete terms, I_h is defined as the ratio of:

- The estimate of the IT resource need, when the evaluation is performed ($\widehat{Y}_0$).
- The estimate of the IT resource need at the horizon h ($\widehat{Y}_h$).

These estimates are expected to be provided by a forecasting model designed to predict Y over time t ($\widehat{Y}_t$), according to explanatory variables $X_{i,t}$ (i ranged from 1 to n), where $X_{i,t}$ has to depict the business/industrial activity of the plant.

To summarise, [V] is evaluated according to:

$$I_h = \frac{\widehat{Y}_h}{\widehat{Y}_0}$$

with $\widehat{Y}_t = f(X_{i,t})$ and where $X_{i,t}$ should not be trivial and understood by STMicroelectronics' top management, to ease communication between IT operational managers (IT architecture team, application teams) and high-level executives. Consequently, IT decisions could be taken on the basis of strategic business-focused scenarios thanks to variables understood all over the company.

Through the qualitative case study, we discovered evidence illustrating the difficulty of evaluating I_h. During the MES project, a 2.5-fold increase estimation was provided by IT architecture team experts, without using any formal method: 'The factor 2.5 was proposed a bit haphazardly, we did not use sophisticated tools nor existing models.'[3] As a consequence, many people lacked confidence in this informal proposition, not founded on a scientific or numerical method. As mentioned by an application team leader: 'He said "2.5", but it seemed to be pulled out of a hat'. Such unease, which did not facilitate the negotiation between the different stakeholders, must be deplored.

Thus, to overcome this difficulty, the decision problem statement [Π] can be expressed as follows: *how to constitute a forecasting model to quantify I_h, e.g. to predict the increase in the need for IT resources, according to high-level business variables?* This final decisional issue does not only deal with the creation of a forecasting model including STMicroeletronics-specific variables, but also with the use of this model to provide a final output as expected by IT management (the I_h estimate). This is a very local decision problem statement, never encountered in such a form in the IT capacity planning literature. This emphasises the two added-values of our work: (1) bring an ad hoc support to STMicroelectronics operational

managers, directly useful in their daily activities; (2) propose a new contribution to the IT capacity planning literature.

4.3. Statistical formalisation phase

4.3.1. Statistical model

As mentioned in the definition of the problem statement, a forecasting model is expected. Let us call it F_t. This model will later be embedded in an evaluation model that will allow to quantify I_h, by furnishing estimates for $\widehat{Y}_0$ and $\widehat{Y}_h$ at $t = 0$ and $t = h$. The modelling process that was used to build F_t model has been fully presented in Lutz et al. (2012). That is why, for the sake of brevity, this part of our work will not be developed in the scope of this paper. However please note that the statistical modelling process exactly applies the procedure synthesised in Section 3.2: two main sequences, exploratory and explanatory, based upon qualitative (experts' knowledge) and quantitative analyses (statistical methods), to take the best advantage possible from the mass of data stored in STMicroelectronics manufacturing IS (IT capacity management and business applications databases, as mentioned in the representation of the problem situation).

F_t is different from the model presented in Lutz et al. (2012), because it fits to a different purpose. Indeed, Lutz et al. (2012) introduced a model to describe the day-to-day activity of the MES CPU. This model is not adapted to a mid-term forecast as required for a two-year long project, such as the MES upgrade described through the qualitative case study:

- *A daily observation is too microscopic* and introduces the modelling of pointless temporal phenomena (weekend dummy variables, AR(1) residual process).
- *Some variables (number of production lots prepared, clogging abstract variables) are not meaningful* to describe mid-term scenarios understood by top-level management, as required by the qualitative case study description.

Consequently, a complementary model was built to overcome these weaknesses. The application of the modelling process proposed in Lutz et al. (2012) for the specific purpose of the MES project results in the simpler following model:

$$F_t : Y_t = \alpha_1 X_{1,t} + \alpha_2 X_{2,t} + \alpha_3 X_{3,t} + \varepsilon_t$$

where:

- t is the time unit, given as a week aggregation;
- Y_t is the MES server CPU consumed per time unit in seconds;
- $X_{1,t}$ is the average work in progress (WIP) of the plant per week. This gives an estimate of the global workload of the plant: processed and pending microelectronic wafers throughout the manufacturing system;
- $X_{2,t}$ is the effective production activity within the plant, given in average number of moves per week: a move is a production operation;
- $X_{3,t}$ is the percentage of automation of the production process (most production operations can indeed be performed either in a manual or automated way);
- ε_t are the F_t residuals, independent and identically distributed (i.i.d.) according to a law $N(0, \sigma^2)$.
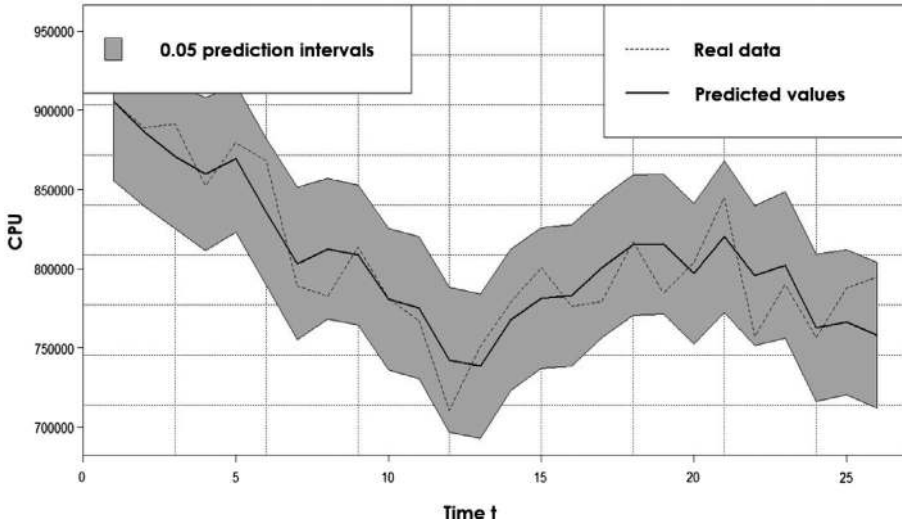
Figure 6. F_t: real data, predicted values and prediction intervals.

This linear model is computed by least-square estimation through a 26-week observation period. It is characterised by a good adjusted-R^2 and meets all the hypothesis of the linear model. It is positively validated by cross-validation (see Lutz et al., 2012, for more details) and can therefore be used with confidence for forecasting. The estimated point model is given by the formula:

$$F_t:Y_t \approx 14.02X_{1,t} + 183.54X_{2,t} - 6414.53X_{3,t} + \varepsilon_t$$

ε_t i.i.d. according to N(0, 20,958.52^2).

We can see that both increases in terms of WIP and move increase the CPU consumption. On the contrary, using the automated processes decrease the CPU activity: this is because another hardware server is required for automated tasks.

Model real data, predicted values and prediction intervals are plotted in Figure 6.

All real data are contained in the prediction intervals. This is consistent: with a 95% confidence level and for 26 observations, a maximum of one Y_t value could be expected outside the prediction intervals. Thus, the model approaches fairly well the MES system's real behaviour.

In a further step, this statistical model is still required to be embedded in an evaluation model [M], as required by Toukiàs' method. Indeed, the statistical model itself does not directly provide the final information expected by IT managers. They do not expect a mathematical relationship, but an effective and directly usable answer based on the quantification of I_h.

4.3.2. Evaluation model [M]

Let us show how F_t can be used to formalise the final evaluation model [M], specifically designed to answer the contextualised decisional problem described above. We recall that an evaluation model is given by: M = ‹A, D, E, H, U, R›. As mentioned earlier, the alternatives [A] are considered through all possible values of I_h (potentially, one to infinite). Indeed, the

Table 3. Formalisation of [M]: dimensions to consider the alternatives of the decision.

Dimensions [D]	Metrics [E]
Y_t: the MES server CPU activity	Seconds CPU consumed per week (s/w)
$X_{1,t}$: global workload of the plant	WIP per week (wip/w)
$X_{2,t}$: effective production activity	Moves per week (mo/w)
$X_{3,t}$: automation of the production process	Per cent (%)

sizing of the new physical servers (CPU architecture, number of CPU, number of cores, etc.) are selected according to the I_h values: the technical considerations that should result from I_h to the choice of the final server are outside the scope of the evaluation model considered within this paper, focused on I_h estimate.

The dimensions [D] and their metrics [E] are F_t variables, as recalled in Table 3.

[H], the criteria to evaluate [A] depends on the values of [D] at the forecasting horizon h. These values express the hypothesis of the decision-maker, regarding a forecasted activity of production. At the time of the studied MES project, the manufacturing management anticipated an overall production increase expressed in terms of wafer-outs, by a factor of 2.25. A wafer-out corresponds to a finished microelectronic wafer that leaves the factory. Comparatively to Table 3 variables, it means that an in-process wafer will be taken out of the volume of WIP after the completion of several moves (several hundred moves, depending on the complexity of the component). The challenge for the IT manager is to link this information to the F_t variables[4]: an increase in the number of wafer-outs, could be induced by a plant throughput intensification for a steady WIP, but also by more general growth of the manufacturing activity intensity, and it also depends on the automation level of the factory. Thus, some hypothesis has to be formalised, within forecast scenarios. Based on collaborative work with experts (task groups), j scenarios are built; the variations of the variables $X_{i,h}^j$ at the horizon h are estimated:

- *Optimistic scenario (j = O)*: the anticipated increase is the consequence of a de-bottlenecking of the plant. The global workload is steady, but the effective production increases. The automated production mode is deployed as expected by manufacturing executives (+30 points);
- *Realistic scenario (j = R)*: the anticipated increase is the consequence of an increase of the global workload, coupled to an equivalent increase of the effective production (which means the production throughput is kept constant). The automated production mode is deployed as expected by manufacturing executives;
- *Pessimistic scenario (j = P)*: the global workload increases, as much as the effective production. In addition, the complexity of the processed components grows (+11% of moves to produce one wafer) and the automated production mode is not deployed;
- *Implicit scenario (j = I)*: it expresses the hypothesis implicitly taken by experts when I_h was estimated during the MES project: increase of the global workload, coupled to an equivalent increase of the effective production, without being able to quantify the consequences of the automated production mode deployment.

Table 4 provides an overview of these scenarios. It describes the variations of $X_{i,h}^j$ at the horizon h compared to time $t = 0$.

The uncertainty of this evaluation model has to be taken into account:

Table 4. Forecast scenarios.

Scenario	Variation compared $t = 0$		
j	$X^j_{1,h}$	$X^j_{2,h}$	$X^j_{3,h}$
O	Steady	x 2.25	+30 points
R	x 2.25	x 2.25	+30 points
P	x 2.25	x 2.5	Steady
I	x 2.25	x 2.25	Steady

Table 5. Numerical $X^j_{i,h}$ values used for simulations.

j	$X^j_{1,h}$	$X^j_{2,h}$	$X^j_{3,h}$
O	50,000	4500	65
R	112,500	4500	65
P	112,500	5000	35
I	112,500	4500	35

- *Preferences of the decision-makers* [H]: using different forecasting scenario is a way to take the uncertainty into account. No mathematical method is required to handle this uncertainty. It is supposed that the values of the predictors are known;
- *The uncertainty associated with the available information* [U] is directly expressed by F_t/I_h: the decision will be made on the basis of a probabilistic model.

Last but not least, we need the aggregation procedure [R], which will output the relevant synthetic information required to make a decision. This procedure is given by using the forecasting model F_t aforementioned, as follows:

$$I_h = \frac{\widehat{Y}_h}{\widehat{Y}_0} = \frac{\widehat{\alpha}_1 X^j_{1,h} + \widehat{\alpha}_2 X^j_{2,h} + \widehat{\alpha}_3 X^j_{3,h} + \varepsilon_h}{\widehat{\alpha}_1 X^j_{1,0} + \widehat{\alpha}_2 X^j_{2,0} + \widehat{\alpha}_3 X^j_{3,0} + \varepsilon_0}$$

(α_i are supposed to be known: $\widehat{\alpha}_i$ estimations are used).

With these final elements, the model M = ‹A, D, E, H, U, R› is available for use.

4.4. User integration step : result [Φ] and final recommendation

This final step illustrates how [M] can be used to produce a result and an effective final recommendation, efficient within the operational usage context.

I_h is equal in law to the ratio of two normal laws (due to ε_h and ε_0). It corresponds to a Cauchy-like law (see Cedilnik, Košmelj, and Blejec (2004) and Qiao, Wood, Lai, and Luo (2006) for more details about this law). However, in the scope of our study, we will not try to evaluate the theoretical parameters of this law. The calculation of I_h for different values of $X^j_{i,h}$ will be analysed through simulations. To achieve such a result, the following procedure is applied :

- *At t = 0, the $X^j_{i,0}$ industrial activity is depicted with the expert*, on the basis of available information: $X^j_{i,0}$ = {50,000, 2000, 35};
- *Based upon $X^j_{i,0}$ and values from* Table 4, $X^j_{i,h}$ values are calculated (Table 5).

I_h density probabilities are evaluated through simulation on the basis of Table 5 values: for each scenario, I_h is calculated 10^6 times. The result of this work is presented to the
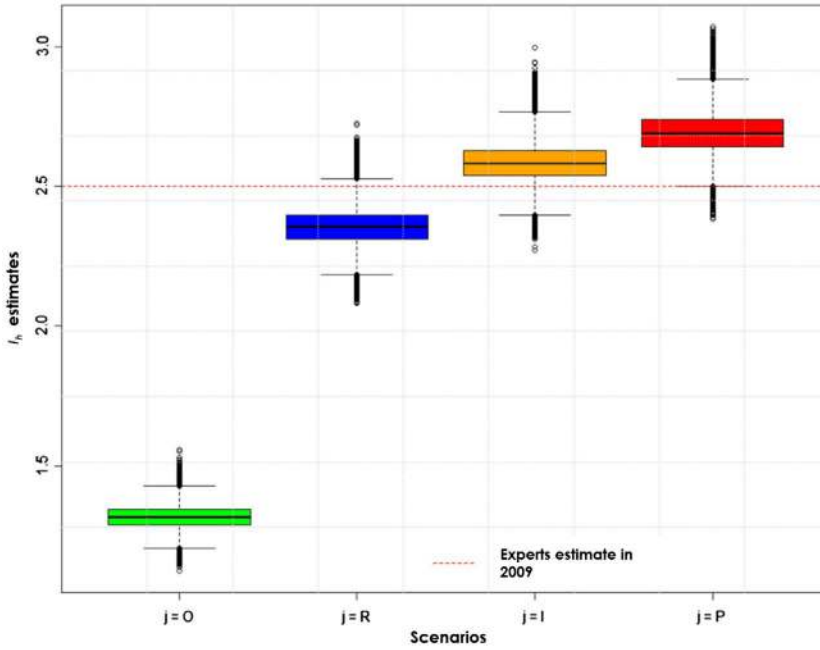
Figure 7. Visualisation of I_h simulation result [Φ].

decision-maker by using a visual representation: a boxplot is used, which enables the display of the main characteristics of the distribution, such as median, quartiles and outliers (see for instance Correa & González, 2002; for more details). This output is plotted in Figure 7. By comparison, the intuitive estimate provided by experts in 2009 is also indicated.

Using this result, the following recommendation could have been provided in 2009 by using the proposed evaluation model:

'Considering a realistic industrial scenario, and based on the first and fourth quartiles of the simulation, I_h should be within the range [2.31, 2.40].'

This final recommendation was provided as a result of to the application of the full decision-aid process proposed in this paper and ends the application of the whole process. Before moving to a conclusion, let us discuss this achievement.

4.5. Discussion: STMictroelectronics case study and further lessons

The scenarios O and P are too far from the implicit scenario used by experts in 2009 to be discussed in detail here. However, they demonstrate the added value of the evaluation model proposed, which enables the characterisation of several industrial scenarios to express detailed and relevant hypothesis. As a consequence, a thorough study of the considered system can be performed and I_h can be evaluated with full background knowledge.

The recommendation based on the scenario R suggests that I_h is comprised between the interval [2.31, 2.40]. In 2009, with an I_h value equals to 2.5, experts slightly overestimated the factor of increase value. However, one can think that this choice, a little bit higher than the median value of the simulation (2.35), would have been a good choice. Indeed, this small overestimation protects the new IT system from any unexpected overload of the plant. This

means that the simulation would have been a good support for IT experts to justify quantitatively their estimates. This would have protected them from the criticisms of their informal way of preparing the capacity plans (see above).

Nevertheless, this result can be moderated. In 2009, experts tried rather to estimate the scenario I. According to the simulation study, I_h first and fourth quartiles are, in this case, contained in the interval [2.54, 2.63], with values that can grow up to 2.77 for the boxplot upper limit. This means that, for the scenario I, experts underestimated the increase factor.

Thus, the estimate provided by experts in 2009 was fairly good, but could have been better controlled. Indeed, experts provided a good forecast for the realistic scenario R, whereas they thought they were anticipating the implicit scenario I. If IT managers had had the simulation results in 2009, they might have built their IT capacity plans in a different manner. Actually, the realistic scenario implies a success in the deployment of the automated production mode, which is not the case of the implicit one. Consequently, if the deployment of this new production mode fails for any unforeseen reason, the new MES servers may encounter a lack of capacity and the production activity may be jeopardised. Hence, the interest of using a quantitative approach for decision-making is fully proven: it helps formalising and challenging the experts' intuitive propositions and, thus, provides better and more reliable risk management tools for IT decision-makers.

But this case study can also bring more general feedback concerning the deployment of DSS within such industrial companies. The key focus underlined for this paper was the importance to manage a contextualisation process, making it possible to increase the added-value of the DSS development: so, what can we learn on the impact of managing contextualisation?

First, it is interesting to note that it is not the statistical model in itself that brings value to the experts. Business experts do not really care about the model: working with them to build and validate it is a kind of 'necessary evil'. The real business value of our work was to clearly express users' question [τ] and bring clear answer [Φ].

Given all the current hype around data science and big data, it is important to keep this in mind. Statistical or machine learning modelling is often not the big deal, as it requires 'only' high-level technical and scientific skills. Value will only be brought to an organisation if such skills are made to serve a valuable business question. This means that data analytics in companies are not Kaggle competitions.[5] Too much money is spent on 'Proofs Of Concepts', only launched to test the ability to build models with business data. Once finished, such Proofs Of Concepts will be abandoned because they do not meet any clear business concern. That is why the key challenge is rarely about creating models. The point is really about asking the right questions then exploiting efficiently the developed models, so as to provide an accurate and direct answer. That is why we think the proposed decision-aid process could be particularly helpful.

A second important discussion deals with the obsolescence of the model. The precision and reliability of the I_h forecast provided by the model, directly depends on the data used to build the forecasting model. An industrial context like STMicroelectronics production plant, constantly evolves at the level of technology, product flows, data flows. To maintain a real ability to support decision-makers on IT capacity management, the model itself should be updated quite systematically. This underlines two important feedbacks and requirements for such DSS:

- The need to monitor and put under control the quality of the statistical models implemented, then to develop an automated or semi-automated approach in charge of systematic updating of the DSS based on data-mining and knowledge extraction;
- The need to maintain in the company a sufficient level of competencies on data mining techniques as well as on the various DSS models implemented industrially, so as to be able to manage such constant obsolescence risks.

5. Conclusion

In this paper a contextualisation process supporting a comprehensive decision-aid building method was specified, dedicated to generate data-driven answers to IT capacity management issues. The first phase of the contextualisation process aims at anticipating and analysing the expected use of a quantitative decision-aid. The second step specifies a way of constructing a quantitative decision-aid model on the basis of this context analysis. This quantitative step is based on the assumption that masses of data stored within modern manufacturing systems can build useful predictive tools, through data analysis. The third ented context of decision-making.

This process was applied to the STMicroelectronics case to confront the decision-aid results with the reality, on a past project of IT capacity transformation. The contextualisation led to the construction of a forecasting statistical model. At this stage of the whole process, the outcome is a statistical model, linking IT resource utilisation to several high-level manufacturing activity-related variables. Through the final step of the process, this model was used to implement a final evaluation model, which generates recommendations for decision-making. This application highlights that the decision-aid process could have brought very valuable insights to the past project decisions. This recommendation is fully contextualised: it leads to a final way to generate a decision-aid answer which is fully specific to STMicroelectronics' internal context (the determination of an increase factor, modelled by the ratio of two normal laws). The contextualisation phase was the only way to conclude to such a fine-tuned decision-aid, which really meets STMicroelectronics managers' needs.

This positive result on the industrial case study both brings reusable results for applied research in DSS and opens new perspectives. First, designed as an additional methodological layer to be integrated within a very generic DSS specification method (i.e. the work by Tsoukias, afore mentioned), the contextualisation process presented in this paper is available and generic enough to be re-used in very different industrial context, when there is a necessity to implement some statistical model. One of the objectives of this research work lies in this genericity: the approach aims at providing the pertinent methods to adapt decision-aid approaches to very distinct industrial contexts, with the ability to address the special features of each context. Within the decision-aid process presented above, the 'Qualitative contextualization' phase answers a generic purpose of capturing and structuring the key features of an organisational context and the associated decision needs. The quantitative methods applied in the subsequent phases ('Statistical Modelling and Formalisation', then 'User Integration') could also be used to fit in a wide panel of contexts; even if alternative quantitative methods could be useful, the overall procedure to integrate the context throughout the quantitative decision-aid building process is fully generic. The added-value of this work is thus to fill one of the gaps limiting the industrial deployment of data mining techniques:

the ability to capture the right user problem and to implement the final decision-aid in a very contextualised form, thus increasing the acceptance rate for the end-user.

However, filling this gap is not sufficient. The case study also puts forth another important challenge for such industrial applications: the crucial need to manage the obsolescence of the knowledge integrated within the statistical models. This confirms important research orientations for the scientific community: (i) on the analysis of the deviations and drifts on the performance of the statistical models and quality of their answers and (ii) on the technical or expertise-based solutions which could support a systematic update of these models, but also (iii) on the way to develop statistical models which would avoid too general and excessively complicated models (Bartolini et al., 2011).

Notes

1. Crolles Production Plant (France), IS Architecture Department in charge of the IT infrastructure and architecture for the production plant.
2. Tsoukiàs uses a formal codification to denote the different parts of the decision-making process.
3. Reported from case study *verbatim* extracts (translated from French).
4. Of course, a straightforward model between Y_t and the number of wafer-outs could be expected. Unfortunately, such a model provides poor predictions. As reported from the qualitative case study: 'it is not easy to say if, in the MES domain, you can observe a correlation coefficient. We do not necessarily have all the relevant data, neither the appropriate statistical tools.'
5. https://www.kaggle.com/ : 'Kaggle is the world's largest community of data scientists. They compete with each other to solve complex data science problem'.

Acknowledgement

Disclosure statement

References

Aguilera, M. K., Mogul, J. C., Wiener, J. L., Reynolds, P., & Muthitacharoen, A. (2003). Performance debugging for distributed systems of black boxes. *Presented in Proceedings of the nineteenth ACM symposium on Operating systems principles, 37*, 74–89.

Ahituv, N., & Igbaria, M. (1988). A model for predicting and evaluating computer resource consumption. *Communications of the ACM, 31*, 1467–1473.

Alcatras, J. R., Gianfaldoni, P., & Paché, G. (2004). *Décider dans les organisations: dialogues critiques entre économie et gestion* [Decide in organizations: Debate between economical and managerial sciences]. Paris: L'Harmattan.

Alford, S., Akker, R., Assaf, R., Bentley, C., van der Bent, J., Breslin, P., … Koch, G. (2006). *Frameworks for IT management: An introduction*. Zaltbommel: Van Haren.

Alkuraiji, A., Liu, S., Oderanti, F., Annansingh, F., & Pan, J. (2014). Knowledge network modelling to support decision-making for strategic intervention in IT project-oriented change management. *Journal of Decision Systems, 23*, 285–302.

Allspaw, J. (2008). *The art of capacity planning*. Sebastopol: O'Reilly.

Barkia, H., Boucher, X., Le Riche, R., Beaune, P., & Girard, M. A. (2013, December 10–13). *Semiconductor yield loss' causes identification: A data mining approach*. 2013 IEEE International Conference on Industrial Engineering and Engineering management (IEEM2013), Bangkok, Thailand, pp. 843–1647.

Bartolini, C., Stefanelli, C., & Tortonesi, M. (2011, May 23–27). *On decision making in business-driven IT management*. 12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011), Dublin, Ireland, pp. 1082–2088.

Bauer, J. F. (2015). Understanding how organizations operate their IT capacity-management processes (Dissertation of the College of Technology). Eastern Michigan University, Ypsilanti, MI.

Bouyssou, D., Dubois, D., Pirlot, M., & Prade, H. (Dir.). (2006). *Concepts et méthodes pour l'aide à la décision 1 – outils de modélisation* [Concepts and methods for decision aid: Modelling tools]. Paris: Hermès.

Bouyssou, D., Marchant, T., Pirlot, M., Tsoukiàs, A., & Vincke, P. (2006). *Evaluation and decision models with multiple criteria – Stepping stones for the analyst*. Berlin: Springer.

Bozdogan, C., & Zincir-Heywood, N. (2012, April 16–20). *Data mining for supporting IT management*. IEEE Network Operations and Management Symposium, Maui, HI, USA, pp. 1378–1385.

Bozdogan, C., Zincir-Heywood, N., & Gokcen, Y. (2013). *Automatic optimization for a clustering based approach to support IT management*. IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), Ghent, Belgium, pp. 1233–2236.

Browning, T. (1994). *Capacity planning for computer systems*. Cambridge: AP Professional.

Cardinali, R. (1992). Information systems – A key ingredient to achieving organizational competitive strategy. *Computers in Industry, 18*, 241–245.

Cedilnik, A., Košmelj, K., & Blejec, A. (2004). The distribution of the ratio of jointly normal variables. *Metodološki Zvezki, 1*(1), 99–108.

Correa, J. C., & González, N. (2002). *Gráficos Estadísticos con R* [Statistical representations with R]. Sede Medellín: Universidad Nacional Sede Medellín.

Coutelle, P. (2005). Introduction aux méthodes qualitatives en science de gestion. In *Cours du CEFAG – Séminaire d'études qualitatives* [Introduction to qualitative methods in management sciences]. (pp. 141–151). Paris: Presses Universitaires de France.

Domanski, B. (1999). Distributed capacity planning. *Enterprise Systems Journal, 14*, 66–71.

Earl, M. J. (1989). *Management for information technology*. Herdfordhsire: Prentice Hall.

Ghosh, R., Longo, F., Xia, R., Naik, V. K., & Trivedi, K. S. (2014). Stochastic model driven capacity planning for an infrastructure-as-a-service cloud. *IEEE Transactions on Service Computing*, 7, 667–680.

Gowan, J. A., & Mathieu, R. G. (1996). Critical factors in information system development for a flexible manufacturing system. *Computers in Industry, 28*, 173–183.

Gunther, N. J. (2007). *Guerrilla capacity planning*. Berlin: Springer.

Hong, S. J., Lim, W. Y., Ceong, T., & May, G. S. (2012). Fault detection and classification in plasma etch equipment for semiconductor manufacturing e-diagnostics. *IEEE Transactions on Semiconductor Manufacturing, 25*, 83–93.

Izza, S. (2009). Integration of industrial information systems: From syntactic to semantic integration approaches. *Enterprise Information Systems, 3*, 51–57.

Izza, S., Vincent, L., & Burlat, P. (2008). Exploiting semantic web services in achieving flexible application integration in the microelectronics field. *Computers in Industry, 59*, 722–740.

Jain, R. (1991). *The art of computer systems performance analysis – Techniques for experimental design, measurement, simulation, and modeling*. New York, NY: Wiley.

Jardim-Goncalves, R., Grilo, A., & Steiger-Garcao, A. (2006). Challenging the interoperability between computers in industry with MDA and SOA. *Computers in Industry, 57*, 679–689.

Jiang, G., Chen, H., & Yoshihira, K. (2008). Profiling services for resource optimization and capacity planning in distributed systems. *Cluster Computing, 11*, 313–329.

Kelly, T. (2005). *Detecting performance anomalies in global applications*. Paper presented at the Second Workshop on Real Large Distributed Systems, San Francisco, CA, pp. 183–189.

Kim, J. S., & Kim, J. B. (2014, May 6–9). *A simulation model for estimating optimal capacity of information system*. International Conference on Information Science & Applications (ICISA'2014), Seoul, Korea, pp. 1–3.

Klosterboer, L. (2011). *ITIL capacity management*. Boston, MA: IBM Press – Pearson.

Leachman, R. C., & Hodges, D. A. (1996). Benchmarking semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing, 9*, 158–169.

Lejeune, C. (2010). Montrer, calculer, explorer, analyser: Ce que l'informatique fait (faire) à l'analyse qualitative [To display, calculate, explore or analyse: The way computer science contributes to qualitative analysis]. *Recherches Qualitatives, 9*, 15–32.

Lutz, M., Boucher, X., & Roustant, O. (2012). Information Technologies capacity planning in manufacturing systems: Proposition for a modelling process and application in the semiconductor industry. *Computers in Industry, 63*, 659–668.

Lutz, M., Boucher, X., & Roustant, O. (2013). Methods and applications for IT capacity decisions: Bringing management frameworks into practice. *Journal of Decision Systems, 22*, 332–355.

Nelson, D. (2002). How to future-proof your IT infrastructure. *IT Professional, 4*, 22–26. doi: 10.1109/MITP.2002.1046640

Nutt, P. C. (1993). The formulation processes and tactics Used in organizational decision making. *Organization Science, 4*, 226–251.

Nutt, P. C. (1998). Framing strategic decision. *Organization Science, 9*, 195–216.

Padonou, E., Roustant, O., & Lutz, M. (2015). Robust monitoring of an industrial IT system in the presence of structural change. *Quality and Reliability Engineering International, 31*, 949–962.

Petruzzi, G. L., & Garavelli, A. C. (2007, May 21). *The strategic value of the "fit" between business processes and IT management: The case of the Italian publishing industry, BDIM '07*. 2nd IEEE/IFIP International Workshop on Business-Driven IT Management, Munich, Germany, pp. 110–111.

Qiao, C. G., Wood, G. R., Lai, C. D., & Luo, D. W. (2006). Comparison of two common estimators of the ratio of the means if independent normal variables in agricultural research. *Journal of Applied Mathematics and Decision Sciences, 2006*, 1–14.

Rebouças, R., Sauve, J., Moura, A., Bartolini, C., & Trastour, D. (2007, May 21–25). *A decision support tool to optimize scheduling of IT changes*. 10th IFIP/IEEE International Symposium on Integrated Network Management, Munich, Germany, pp. 343–652.

Roy, B. (1993). Decision science of decision-aid science? *European Journal of Operational Research, 66*, 184–203.

Roy, B. (2006). Regard historique sur la place de la recherche opérationnelle et de l'aide à la décision en France [Historical analysis of the place of operational research and decision-aid in France]. *Mathématiques et sciences humaines, 175*, 25–40.

Shrestha, A., Cater-Steel, A., & Toleman, M. (2016). Innovative decision support for IT service management. *Journal of Decision Systems, 25*, 486–499.

Thietart, R. A. (2007). *Méthodes de recherche en management* [Research methods for management sciences]. Paris: Dunod.

Tsoukiàs, A. (2007). On the concept of decision aiding process. *Annals of Operations Research, 154*, 3–27.

Wacheux, F. (1996). *Méthodes qualitatives et recherche en gestion* [Qualitative research methods for management sciences]. Paris: Economica.