



**HAL**  
open science

# RMS balancing and planning under uncertain demand and energy cost considerations

Xavier Delorme, Audrey Cerqueus, Paolo Gianessi, Damien Lamy

► **To cite this version:**

Xavier Delorme, Audrey Cerqueus, Paolo Gianessi, Damien Lamy. RMS balancing and planning under uncertain demand and energy cost considerations. *International Journal of Production Economics*, 2023, 261, pp.108873. 10.1016/j.ijpe.2023.108873 . emse-04079872

**HAL Id: emse-04079872**

<https://hal-emse.ccsd.cnrs.fr/emse-04079872v1>

Submitted on 5 Feb 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# RMS balancing and planning under uncertain demand and energy cost considerations <sup>1</sup>

Xavier Delorme<sup>a</sup>, Audrey Cerqueus<sup>b</sup>, Paolo Gianessi<sup>a</sup>, Damien Lamy<sup>a</sup>

<sup>a</sup>*Mines Saint-Etienne, Univ Clermont Auvergne, INP Clermont Auvergne, CNRS, UMR 6158 LIMOS, F-42023 Saint-Etienne, France (e-mail: delorme@emse.fr, paolo.gianessi@emse.fr, damien.lamy@emse.fr).*

<sup>b</sup>*IMT Atlantique, LS2N, UMR CNRS 6004, F-44307 Nantes, France (e-mail: audrey.cerqueus@imt-atlantique.fr)*

---

## Abstract

In recent years, we have observed rapid changes in the customer demand along with shorter product life cycles. In addition, sustainability concerns about production systems are growing, especially due to energy supply fluctuations in terms of either availability or cost. Among these challenges, energy efficiency is of the utmost importance, and Reconfigurable Manufacturing Systems (RMS), most notably through their scalability feature, could represent a valuable solution: production resources can be reorganized promptly to adapt throughput to external factors, such as uncertain demand or Time-Of-Use prices. Although the aforementioned challenges concern day-to-day management, they should be anticipated at the design stage of the production system, whose behavior might otherwise not meet expectations and hinder the competitiveness of the company. One possibility is to consider the expected performance of such a system from the viewpoint of different productivity and energy-efficiency criteria, through line balancing and future production planning. This can be modeled as a bi-level optimization problem, in which the line balancing of the RMS is the upper level and the configuration planning is the lower level. We consider three criteria, namely the number of workstations, the expected service level and the expected energy cost, taking into account demand uncertainty through scenarios. A three-phase metaheuristic is developed and its performances on instances derived from the literature are discussed. The results show that consistent energy cost savings can be achieved, even with very few configurations.

*Keywords:* Reconfigurable Manufacturing Systems, Scalability, Time-Of-Use energy prices, Scenario-based Optimization, Multi-objective Metaheuristic, Bi-level Optimization

---

## 1. Introduction

With globalization, individual consumption of products has increased drastically in recent decades. This change in the behavior of consumers has led to high volatility in markets demand, making it harder to forecast the throughput requirements of production systems [1]. Meanwhile, the production of goods implies the significant use of several resources, raising sustainability issues [2].

In this regard, and driven by global warming, greenhouse gas emissions and energy consumption have become a worldwide concern. Since the industrial sector (comprising refining, mining, manufacturing, agriculture, and construction) is accountable for a share of more than 50% of the end-use energy consumption [3], companies are increasingly being urged towards energy efficiency. This is also true specifically for what concerns electricity consumption: the International Energy Agency reports [4] that the industrial sector accounted for 22% of total final electricity consumption in 2020, a figure which is expected to grow to 46% in 2050 due to the electrification of several processes. Manufacturing systems (MS) are among the most energy-consuming industrial systems, and are therefore deeply affected by energy-efficiency issues, particularly electricity usage, from design through to management [5]. However,

---

<sup>1</sup>This is an Accepted Manuscript version of the following article, accepted for publication in International Journal of Production Economics: “Xavier Delorme, Audrey Cerqueus, Paolo Gianessi Damien Lamy (2023) RMS balancing and planning under uncertain demand and energy cost considerations, International Journal of Production Economics 261: 108873, DOI:10.1016/j.ijpe.2023.108873”. It is deposited under the terms of the Creative Commons Attribution-Non Commercial License 2023 CC-BY-NC 4.0 (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

MS are confronted with increasingly volatile availability of electric energy, due not only to electric load management policies [e.g. 6], but also to the growing focus on renewable energy sources [7]. This has led to studies into how to use electric energy efficiently in MS, by considering it in the form of various objectives and/or constraints, namely: total energy consumption [8], power peak limitation [9] and Time-Of-Use (TOU) pricing [10].

Designed to cope with market changes, reconfigurable manufacturing systems (RMS), introduced by [11], are usually composed of several workstations connected by a conveyor and/or a gantry to transport the parts. Each workstation comprises several machines in parallel, which may be Computer Numerical Control (CNC) machines, Reconfigurable Machine Tools (RMTs) or other types of resources (e.g. cobots) [12]. This allows an RMS to evolve from configuration to configuration in order to offer the appropriate production setting for each usage [13]. It is therefore standard for manufacturers to use RMS to manage product variety [14], or uncertainty over future demand fluctuations [15]. RMS have also been studied to assess whether they can help dealing with sustainability challenges: for instance, [2] states that RMS must pursue the energy efficiency of products and processes in this respect.

In this paper, we seek to investigate the possibility of using RMS to deal with a TOU energy pricing scheme so as to minimize energy-related production costs. For instance, an energy-aware configuration could be implemented during expensive periods, with a trade-off in productivity that could be balanced, to meet the demand, by using more productive, more energy-consuming configurations during periods with lower energy costs.

However, while energy-related operational costs are usually considered only at an operational level, the choice of the available system configurations is a critical decision at the design stage as it will impact the performance of the RMS through its lifetime [1]. It would thus be beneficial to consider the energy performance of configurations at the design stage. By doing so, once the system is designed, the operational decision of finding the most-fitting configuration planning w.r.t. a given energy pricing profile would actually minimize energy-related production costs. Naturally, in order to anticipate at design stage the possible operational behaviour of the system in the presence of fluctuating energy prices, uncertainty about the future demand needs to be taken into account. Meanwhile, it is known that considering in an integrated way different problems that are usually solved independently can increase the quality of solutions [16, 17].

Although works exist that study the usage of RMS to minimize energy consumption (the recent survey [18] reviews some of them), we are not aware of studies focusing on how to use them to deal with variable energy prices and minimize energy-related costs – a possibility that seems interesting to investigate. Even less considered in the literature, and possibly more interesting, seems the study of how energy costs consideration from the earliest stages of a production system can impact its design, and the possible added value of an integrated approach to the problems of system design and configuration planning.

On this basis, we suggest to tackle jointly the determination of the RMS configuration set at the design level, and the planning of how to use them at operational level, with the aim to obtain a more energy-efficient reconfigurable system. More specifically, we suggest to model both problems as a bi-level optimization problem. A three-step metaheuristic approach to this problem is presented, in which uncertainty concerning the demand to be met is handled through demand scenarios [19]. Three objectives are considered that are not aggregated and have the same level of priority: the installation cost; the expected values of both the energy cost w.r.t. a TOU profile, and the service level, in order not to disregard the production performance. The goal is to determine a set of tradeoff solutions (with regard to the three objectives) which are nondominated in the sense of Pareto dominance. This work is an extension of the conference article [20], with a full definition of the problem considered, extensive numerical experiments and a detailed analysis of the results obtained and their managerial implications.

The remainder of the paper is organized as follows: Section 2 reviews related works in the literature. Section 3 outlines the problem under consideration. Section 4 delves into the details of the proposed approach, Section 5 describes the computational experiments, and Section 6 provides some managerial insights. Finally, Section 7 looks at the conclusions and perspectives of this work.

## 2. Literature review

Reconfigurable manufacturing systems were introduced in the late 1990s [11] with the objective of achieving a tradeoff between productivity and flexibility, and of building a production system at the

frontier of flexible manufacturing systems and dedicated lines.

Dedicated lines are manufacturing systems designed to produce a single model at a high rate. In these systems, a single static configuration is used [21]. Because of this static nature, it is extremely costly to adapt the system to a different model [22] and handling simultaneously multiple model would usually require to implement multiple lines [23].

At the opposite end of the spectrum, flexible manufacturing systems are designed to produce different models within the same system, taking advantage of the similarities between the models to share specific tools and skills [24, 25, 26]. However, the set of models to be produced is static, which usually leads either to consider a very restrictive family of models, thus being at risk to be unable to produce future models, or to consider a very large family, which implies prohibitive investment costs and therefore a non-competitive price per part.

Reconfigurable manufacturing systems are developed to be dynamically reconfigured in order to adapt to changing product specifications. Six core characteristics define an RMS: modularity, integrability, diagnosability, customization, convertibility, and scalability [13]. Among these features, the last three are related to productivity and flexibility issues:

1. the customization feature focuses on the ability to handle multiple models in the system (sometimes referred to as flexibility corridor). However, it is worth mentioning that RMS are not always used in a multiple models setting: single-model RMS are also used and have received attention from researchers (see, e.g., [27, 28, 29]).
2. the convertibility feature corresponds to the ability to change the functionalities of the RMS over time to integrate new models, which enables the designer to consider an initial setting based on a small family of models (or even a single model) without risking to miss on future models [30].
3. the scalability feature is the ability to modulate the throughput volume and is commonly used to meet a changing demand [11].

In the following, we focus on the various topics related to the problem at hand, basing our literature review on advances in RMS and energy efficiency, the notion of scalability, and the link between demand uncertainty and production system design.

Though RMS are designed to increase the responsiveness of manufacturing systems to unforeseen changes [12], their configuration and design is made difficult due to uncertainty in demand. To address this issue, [31] defines a multi-period stochastic programming model of an RMS and considers an uncertain demand. The objective consists in minimizing total costs (such as production, storage, configuration and reconfiguration costs). To ensure the response capability of the production system to future market expansions, demand should be considered at the design stage. This characteristic, which is related to scalability, enables future rapid and incremental throughput upgrades.

The concept of scalability is made possible by adding parallel resources to workstations, or removing them, and is motivated by the need to adapt the production level to the actual demand in terms of product quantity. Considering this characteristic, [32] investigate a model for assessing the performance of different policies (scaling rate, Work In Process level, inventory level and backlog level) in the context of a make-to-order RMS according to different demand scenarios. [33] propose a combinatorial definition of the line balancing problem associated with RMS optimizing its scalability at the design stage. In [34] a scalability planning methodology for RMS is explored. The approach consists in changing the capacity of an existing system through successive reconfigurations in order to minimize the number of machines required to ensure a new throughput. [35] study a production planning system that integrates characteristics of RMS, with a focus on the scalability and convertibility of production capacities. In [36], multi-model and scalable RMS are explored. The objective is to minimize design and reconfiguration costs while respecting a given demand spread over multiple production periods. In their work, it is possible to (i) upgrade or downgrade the RMS depending on forecasted demands for each period, or (ii) select and reconfigure the reconfigurable manufacturing tools for all periods based on longer-term estimates.

Some authors have investigated the potential of RMS to improve energy efficiency. Based on a 6R methodology (Reduce, Reuse, Recycle, Recover, Redesign, Remanufacture) analysis, [2] show that RMS can contribute to this goal. They also underline that scalability is one of the drivers of improvements to energy and resource efficiency and thus of more sustainable production systems, as it provides the right capacity at the required times and thus reduces the wastage of various resources. [37] provide an analysis of the metrics that could be affected by modifications to the six core features of an RMS, and

show that scalability can be an important lever to lower energy usage (total consumption, losses based on inactivity, energy intensity) and energy-related costs.

A few works have explored energy efficiency as a criterion for RMS. [38] introduce the concept of energy-efficient RMS and investigates a discrete event simulation model to evaluate the energy efficiency of systems. [39] study the problem of balancing an RMS so as to minimize the peak of production-related electric power consumption. However, multiple objectives can be optimized in a production system. Thus, several papers consider energy measures in addition to productivity/cost assessments. [40] investigate a multi-objective production planning problem that considers energy consumption, throughput, and inventory holding costs to assess the performance of the planning. A configuration gives rise to a production plan, leading to a total energy consumption. The recent review [18] mentions some other works that investigate energy consumption optimization in RMS.

Although some research projects have investigated energy efficiency in RMS, the operational control of RMS subject to Time-Of-Use (TOU) pricing has attracted little attention in the literature. This topic is more prevalent in production planning and scheduling. For instance, in [41], the objective is to minimize energy consumption costs in the context of a single machine production scheduling problem subject to different production modes (switch on/off, stand-by) while accounting for variable energy prices. In [42] several insights are provided on the proper management of an unrelated parallel machine scheduling problem under TOU pricings with the objective of minimizing the total electricity cost. Recently, in [43], a job-shop scheduling problem is investigated in order to optimize production and economic criteria, namely total production time (makespan) and costs subject to TOU tariffs.

Despite the inclusion of energy-related objectives and/or constraints in recent works, the survey by [7] shows that there is still little research on the use of RMS to improve energy efficiency in production. Meanwhile, as stressed in [44], decisions taken at the design stage of production systems can affect their operational behaviour from the point of view of energy efficiency. It is therefore important to address such issues from the design stage, and from the perspective of different time horizons.

Few works have simultaneously investigated the design and usage of RMS. [17] consider the integration of long-term and short-term decisions in the design and reconfiguration of modular assembly systems. The integrated approach is divided into four sequential steps, namely the definition of configurations, the layout and assembly process of cells, the production planning and the planning of reconfigurations. Once the initial candidate designs are chosen, the equipments and operation sequencing are arranged by an assembly cell configuration tool, followed by production planning based on simulation. To compute the system designs, expected product demands according to low/medium/high demand scenarios are considered. However, it might be difficult, or even intractable to handle all features of an integrated problem at once, so nested optimisation approaches may be required [45]. For instance, in [46], the authors investigate a bi-level approach for process plan generation of RMS. The lower-level has an objective based on energy losses and ensures that the process plan respects parts requirements, while the upper-level checks the feasibility of the obtained solution as to selected machines and tools.

The first studies that have considered a bi-level model to optimize energy costs in the context of RMS are those of [47, 48]. In these works, a balancing problem is investigated in order to obtain a set of configurations that can be used to determine a production plan to match a given demand while minimizing the total energy cost based on a Time-Of-Use (TOU) pricing scheme. However they do not take into account the difference in the available information about the demand that the decision-makers of the two levels (design, planning) have, as well as the multiple criteria to consider in the design stage. In Section 3, we further define the problem and propose a methodology to address it.

Table 1 summarizes the reviewed scientific contributions and their features, and allows a comparison with this work. The column *problem type* describes the typology of addressed problems, i.e., design, layout, planning or scheduling of production systems. *Energy features* are energy consumption, electricity costs or peak power consideration. These features are generally more common in scheduling and planning problems, and less on works about RMS. Such features appear in the reviewed works as objective functions (O), indicators (I) or constraints (C). *Productivity features* correspond to the time performance of a production system. Economic features aggregate various costs, e.g. exploitation costs or configurations/reconfigurations costs. The column *uncertainty* refers to whether a paper considers it, and its source. *RMS features* columns display the considered characteristic of RMS. The column *method* shows the investigated approaches: Integer Linear Programming (ILP), Mixed-Integer Nonlinear Programming (MINLP) or general Discrete Optimization (DO); Genetic Algorithm (GA), Multistart Evolutionary Local Search (MSxELS) or other types of heuristic approaches, e.g. matheuristic algorithms;

Discrete Event Simulation (DES); System Dynamics (SD); Petri Nets (PN); Constrained K-Shortest Paths (CKSP); Timed Net Condition/Event Systems (TNCES). Finally, the column *objective* displays whether the investigated problem is mono-objective or multi-objective, and in the latter case, whether an aggregated objective function is used. As can be seen, the study at hand addresses a set of features that is underrepresented in the literature.

### 3. Problem description

We consider the design of a paced reconfigurable manufacturing system with uncertain demand taking into account the design cost as well as the future energy cost and service level associated with the future usage of the system. We thus face a *bi-level optimization problem* composed of a line balancing problem as the *upper level*, and a configuration planning problem as the *lower level*.

The line balancing problem occurs at the design stage. A set  $\{1, \dots, n\}$  of operations required to produce a single model of product has to be assigned to a set  $\{1, \dots, m\}$  of stations. Each operation must be assigned once and to exactly one station, with a set  $\mathcal{P}$  of precedence relations and an upper bound  $\bar{n}$  on the number of operations per station, restricting the possible assignments. An operation  $j$  is defined by a processing time  $t_j$  and consumes an energy  $e_j$ , with both values being independent of the station to which the operation is assigned. The RMS is assumed to be a parallel-serial line with crossover [see e.g. 49], i.e., parts move along the stations in a serial manner but several identical resources can be used in parallel at each station. Each configuration  $i$  is thus defined by the number  $r_{i,k}$  of resources of each station  $k$ . In such a system, the workload of a station is the sum of the processing times of operations assigned to it, divided by the number of resources. The line is paced, therefore the takt time  $T_i$  of configuration  $i$  is defined by its bottleneck station, and can be calculated according to (1):

$$T_i = \max_{k \in \{1, \dots, m\}} \left\{ \frac{\sum_{j=1}^n t_j \cdot x_{j,k}}{r_{i,k}} \right\} \quad (1)$$

However, the demand to be met is uncertain and we consider a set  $\mathcal{D}$  of demand scenarios with an associated probability  $\sigma_d$ .

The objective is to determine a set of nondominated assignments  $x$  of operations to stations (in the sense of Pareto dominance) with regard to the three following criteria, and without any priority among them:

1. the number of stations of  $x$ , so as to reduce the cost of the line;
2. the expected value of the service level, which is defined for  $x$  w.r.t demand  $d \in D$  as the percentage of  $d$  that can be met by using the configurations associated with  $x$ ;
3. the expected value of the per-produced-unit energy cost, defined for a planning as the economic cost of the energy consumption associated with the configurations used by the planning when dealing with the demand  $d$ .

The configuration planning problem occurs when the RMS is installed and already operational. At this stage, although the balancing is set and cannot be changed, reconfigurations of the RMS are possible and can be performed by changing the number of resources of each station. Indeed, the basic configuration that can be derived from a balancing  $x$  has only one resource on each station, but other configurations can be obtained from it by incrementing the number of resources of stations. Since the line is paced, its takt time can only be reduced by increasing by one the number of resources of the bottleneck station, potentially leading to a new bottleneck station [33]. This process can be repeated until either the maximum overall number of resources  $\bar{R}$  is reached, or the bottleneck station already has the maximum number  $\bar{r}$  of resources per station. We denote  $\mathcal{C}(x)$  the set of configurations that can be derived from the balancing  $x$ , and  $T_i$  the takt time of the configuration  $i \in \mathcal{C}(x)$ . Since a reconfiguration only implies switching certain resources on or off, we assume that reconfiguration times and costs can be neglected. The planning problem then consists in deciding how to use the configurations of  $\mathcal{C}(x)$  over a time horizon  $H$  to meet the demand associated with a given scenario taking into account a TOU pricing scheme for energy costs. In such a scheme, the time horizon is partitioned in a set  $\Pi$  of periods  $p$  of duration  $v_p$  with associated energy unit cost  $w_p$ . We assume that for any configuration  $i \in \mathcal{C}(x)$ , the energy consumed per unit of time  $g_i$  can be calculated by adding, for each station, the energy needed to process the operations assigned to the station and a residual energy consumption during idle time. The residual

reference	problem type	energy features			productivity features			economic features	uncertainty	RMS features			method	objective
		energy consumption	electricity cost	power peak	time horizon	throughput	service level			customization	scalability	convertibility		
[13]	Design					C		O			X		GA	mono
27	Design							O				X	ILP, CKSP	mono
28	Design							O			X	X	ILP	mono
32	Design							I	demand		X		SD	
33	Design					O		O			X		complete enum.	aggr
36	Design							O		X	X	X	ILP	mono
38	Design	I			I							X	TNCES	
39	Design			O		C							ILP	mono
44	Design			O		C							MSxELS	mono
34	Planning					C		O			X		GA	mono
35	Planning					C		O		X	X	X	DO, PN	mono
40	Planning	O				O		O		X		X	ILP	aggr
41	Scheduling		O		C					X			ILP, GA	mono
42	Scheduling		O		C					X			ILP, heuristic	mono
43	Scheduling		O	C	C			O		X			ILP, heuristic	mono
[17]	Design Planning					C		O	breakdown, demand	X		X	heuristic, DES	mono
[31]	Design Planning							O	demand	X			stochastic ILP	mono
[46]	Design Planning	O								X			MINLP	mono
[47]	Design Planning		O			O					X		matheuristic	aggr
[48]	Design Planning		O			O					X		matheuristic	aggr
this work	Design Planning		O					O	O	demand		X	matheuristic	multi

Table 1: Synthesis of the reviewed literature.

energy consumption of a station  $k$  is considered to be proportional by a factor  $\alpha$  to the product of the idle time, the number of resources  $r_{i,k}$  and the energy consumption of operations assigned to  $k$ . Equation 2 illustrates the computation of  $q_i$ :

$$\begin{aligned} q_i &= \frac{1}{T_i} \sum_{k=1}^m \left( \sum_{j=1}^n e_j \cdot x_{j,k} + \alpha \cdot r_{i,k} \cdot \left( T_i - \frac{\sum_{j=1}^n t_j \cdot x_{j,k}}{r_{i,k}} \right) \frac{\sum_{j=1}^n e_j \cdot x_{j,k}}{\sum_{j=1}^n t_j \cdot x_{j,k}} \right) \\ &= \frac{1}{T_i} \sum_{k=1}^m \left( \sum_{j=1}^n e_j \cdot x_{j,k} \right) \cdot \left( 1 + \alpha \left( \frac{r_{i,k} \cdot T_i}{\sum_{j=1}^n t_j \cdot x_{j,k}} - 1 \right) \right) \end{aligned} \quad (2)$$

No guarantee is given that the RMS will be able to satisfy the given demand, i.e., that the takt time  $\min_{i \in \mathcal{C}(x)} \{T_i\}$  of its most productive configuration will be tight enough. Therefore, the configuration planning problem does indeed have two objectives, which (unlike those of the upper-level problem) are considered in lexicographic order: first, the highest possible percentage of the demand is met; second, the planning to satisfy this maximum demand percentage with minimum per-produced-unit energy-related economic cost is determined.

One aspect that it is of the utmost importance in understanding the problem under study is the reasons behind its bi-level nature. The problem consists in designing an RMS so as to optimize the three aforementioned criteria, two of which require to consider the optimal solution of the operational problem of configuration planning. It is then clear that the two problems are located on two hierarchically and temporally separated decision layers. Design decisions must be taken at a moment in time in which the demand to be fulfilled is not known, forcing the decision maker to proceed by considering demand scenarios; at a later moment, when the system has been designed, another decision maker will have to determine a configuration planning to optimally (w.r.t. service level and energy costs) fulfill a demand that will, by then, be known. Yet, the two decision layers are strongly interconnected: on the one hand the performance of a balancing cannot be assessed without solving the configuration planning problem for each demand scenario, while on the other hand the planning problem cannot be solved if the configuration set is not known in advance. Bi-level optimization is the methodological framework used to formalize and address the overall problem as being made up of two subproblems –an upper-level or *master* problem and a lower-level or *slave* problem– for which such a relation exists. However, the separation between the two problems can affect the way in which elements that are common to the two are addressed. In the particular case of the bi-level problem studied here, this holds for the production demand. Demand is a random variable at the upper level, which is unsurprising since at the design stage it cannot be known: this uncertainty is addressed by means of scenario-based optimization, where every scenario features an assumed demand value. Incidentally, this gives rise to as many slave problems as the scenarios: for each such subproblem, when the production planning must be decided, demand is no longer unknown and can be considered as a deterministic data item for all intents and purposes.

To recap, the problem at study is a bi-level optimization problem, with a multi-objective, scenario-based master problem, and a set of slave problems which are mutually independent, deterministic and have two lexicographically ordered objectives.

Table 2 summarizes all the notations used. The decision variables can be defined as follows:

- $x_{j,k} = 1$  if operation  $j$  is assigned to station  $k$ , 0 otherwise;
- $u_k = 1$  if station  $k$  is used, 0 otherwise;
- $y_{i,p,d} \in [0; 1]$ : percentage of period  $p$  during which production planning uses configuration  $i$  when scenario  $d$  occurs.

Using the aforementioned notations, the bi-level optimization problem at hand can then be expressed as a mixed integer nonlinear program (3). In this model, (3a-3c) represent the three criteria to optimize at the upper level. In the upper-level problem, (3d) and (3e) are classical balancing assignment and precedence constraints, respectively, and (3f) ensures that the upper bound on the number of operations per station is not exceeded and also links variables  $x$  and  $u$ . The constraints (3g) eliminate symmetrical solutions. Note that the decision variables from the lower level  $y_{i,p,d}$  are referred to as  $y_{i,p,d}^*$  in the



upper level since they correspond to the optimal values of these variables for the corresponding lower-level problem. The  $|\mathcal{D}|$  lower-level problems are defined with the criterion to optimize (3j) as well as constraints (3k) and (3l), which ensure that the planning horizon is respected and the demand is met, respectively.

$$\text{Minimize } z_m = \sum_{k=1}^m u_k \quad (3a)$$

$$\text{Maximize } z_{sl} = \sum_{d \in \mathcal{D}} \sigma_d \cdot \frac{\min \left\{ d; \frac{H}{\min_{i \in \mathcal{C}(x)} \{T_i\}} \right\}}{d} \quad (3b)$$

$$\text{Minimize } z_{ec} = \sum_{d \in \mathcal{D}} \sigma_d \cdot \sum_{i \in \mathcal{C}(x)} \frac{v_p \cdot w_p \cdot q_i \cdot y_{i,p,d}^*}{\min \left\{ d; \frac{H}{\min_{i \in \mathcal{C}(x)} \{T_i\}} \right\}} \quad (3c)$$

$$\text{s.t. } \sum_{k=1}^m x_{j,k} = 1, \forall j \in 1, \dots, n \quad (3d)$$

$$\sum_{k=1}^m k \cdot x_{j,k} \leq \sum_{k=1}^m k \cdot x_{g,k}, \forall (j, g) \in \mathcal{P} \quad (3e)$$

$$\sum_{j=1}^n x_{j,k} \leq \bar{n} \cdot u_k, \forall k \in 1, \dots, m \quad (3f)$$

$$u_k \leq u_{k-1}, \forall k \in 2, \dots, m \quad (3g)$$

$$x_{j,k} \in \{0; 1\}, \forall j \in 1, \dots, n, k \in 1, \dots, m \quad (3h)$$

$$u_k \in \{0; 1\}, \forall k \in 1, \dots, m \quad (3i)$$

$$y_d^* = \operatorname{argmin}_{y_d} \sum_{i \in \mathcal{C}(x)} \sum_{p \in \Pi} v_p \cdot w_p \cdot q_i \cdot y_{i,p,d}, \forall d \in \mathcal{D} \quad (3j)$$

$$\text{s.t. } \sum_{i \in \mathcal{C}(x)} y_{i,p,d} \leq 1, \forall p \in \Pi \quad (3k)$$

$$\sum_{i \in \mathcal{C}(x)} \sum_{p \in \Pi} \frac{v_p}{T_i} \cdot y_{i,p,d} \geq \min \left\{ d; \frac{H}{\min_{i \in \mathcal{C}(x)} \{T_i\}} \right\} \quad (3l)$$

$$0 \leq y_{i,p,d} \leq 1, \forall i \in 1, \dots, n, p \in \Pi \quad (3m)$$

#### 4. A three-phase matheuristic approach

To tackle this problem, we developed a multi-objective matheuristic in three phases. Phase 1 aims to generate some balancing and the resulting configurations. Phase 2 focuses on the planning of these configurations, for each balancing, taking into account the demand scenarios and the energy cost profile, and evaluating the three objective functions. Phase 3 aims to filter the balancing solutions to keep only those presenting an interesting tradeoff between the three objective functions.

##### 4.1. Phase 1: generation of the designs

Phase 1 aims to generate interesting balancing solutions and their associated configurations. A Multi-Objective Simulated Annealing (MOSA) is developed [50]. The MOSA focuses on the generation of balancing solutions, during the initialization and determines new solutions at each step, which are evaluated through the derived configurations, obtained using the iterative process presented in the previous section. The configurations vary in terms of the number of resources allocated to each station. The whole set of configurations is considered when evaluating a balancing. We aim to assess it on the

three criteria of the bi-level problem; however, during Phase 1, it is not possible to evaluate exactly the expected energy cost, which depends on the configuration planning. To assess it, we use a proxy function, given by a hypervolume of per-time-unit energy consumption  $q_i$  and takt time  $T_i$  of each configuration  $i$ . Only the non-dominated configurations regarding  $q$  and  $T$  are kept, i.e., those such that there are no configurations derived from the same balancing with a better value regarding both  $q$  and  $T$ . The values of  $q_i$  and  $T_i$  are normalized as follows:  $\tilde{q}_i = \frac{q_i - q_L}{q_U - q_L}$  and  $\tilde{T}_i = \frac{T_i - T_L}{T_U - T_L}$  where  $q_L$ ,  $q_U$ ,  $T_L$  and  $T_U$  are respectively lower and upper bounds for  $q$  and  $T$ . Assuming that configurations are sorted in decreasing order of  $T_i$ , equation 4 shows how the hypervolume is computed:

$$\mathcal{H} = (1 - \tilde{T}_1)(1 - \tilde{q}_1) + \sum_{i \geq 2} (\tilde{T}_{i-1} - \tilde{T}_i) (1 - \tilde{q}_i) \quad (4)$$

This equation can be interpreted as a sum of the geometrical areas of rectangles defined by values  $q_i$  and  $T_i$  of the configurations, as illustrated in Figure 1.

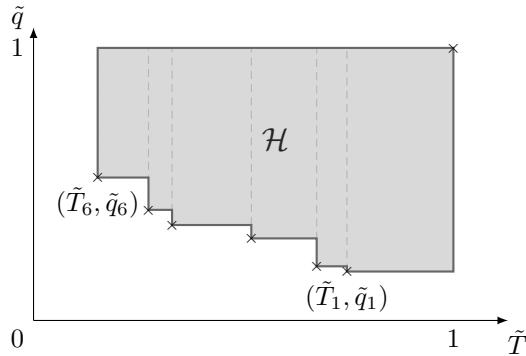


Figure 1: Computation of the hypervolume for a set of configurations

At each step of the MOSA, a new balancing is randomly chosen in the neighborhood of the current balancing. Here we consider a neighborhood composed of all feasible balancing solutions that can be obtained by the current one by moving exactly one operation  $j$  from its current workstation to another one. A feasible balancing has a maximum of  $m$  stations, a maximum of  $\bar{n}$  operations per station and complies with precedence constraints. The station to which the operation is assigned in the neighbor can be a new station with only this operation on it, and if a station becomes empty it is closed.

The new balancing and the current one are compared according to the three objectives of this phase (the number of stations, the expected service level and the hypervolume  $\mathcal{H}$ ). The probability of acceptance is the product of the probabilities of acceptance computed independently for each of these three objectives. On the same principle as a single objective simulated annealing, the probability of accepting a solution according to a given objective is 1 if it is better than the value of the current balancing, otherwise it is a function  $exp(\delta f^p/t)$  of the difference  $\delta f^p$  between the objective values of the current balancing and the new balancing, where  $t$  is the temperature (decreasing according to a given geometrical descent factor  $\beta$  every  $L$  steps of the MOSA). The expected service level and the hypervolume objective function take values in  $[0, 1]$ , and the objective on the number of stations is normalized to also take value in  $[0, 1]$ , by dividing it by the maximum number of stations  $m$  allowed in the system.

During the execution of the MOSA, an archive is used to keep in memory the set  $\mathcal{B}$  of non-dominated balancing solutions, which is returned to Phase 2 at the end of the execution of the MOSA (i.e., after a predetermined number of steps).

In order to get a better coverage of the Pareto front, the MOSA has a multi-start. Two types of greedy initializations are performed: a directed one (using a weighted sum objective function during the construction of the initial solution) and a random one. The two initialization methods work on the same principle: we iteratively build a balancing, and each step consists in assigning one operation to the current station or to the next one (opening a new station). To respect precedence constraints, at each step the operation is chosen among the list of operations for which all predecessors have been assigned (called assignable operations). The procedure ensures that the maximal number of stages and the maximum number of operations per station are complied with. The difference between the directed

and random initialization methods lies in the probability of choosing each of the assignable operations and in the probability of opening a new station. In the random method, each assignable operation has the same probability of being chosen and the probability of opening a new station is randomly chosen at the beginning of the execution. In the directed method, the probability of choosing an operation and its station (current or new one) is proportional to a weighted sum of three values: the number of stages, the takt time (representing the expected service level) and the energy consumption  $Q$  of the partial balancing (representing the expected energy cost). To cover the Pareto front, different starts take different weights in this weighted sum.

#### 4.2. Phase 2: configuration planning

Phase 2 consists in optimally planning, via Linear Programming (LP), the usage of the configurations of  $\mathcal{C}(x)$  for each  $x \in \mathcal{B}$  over a period  $H$  to satisfy each demand scenario  $d \in \mathcal{D}$ . However, since the cases represented by the  $|\mathcal{D}|$  different scenarios are completely independent, they give rise to as many Linear Programs (LPs).

In each of these LPs, the demand must be met over a time horizon  $H$ . However, a balancing  $x$  has a maximum attainable demand  $\frac{H}{\min_{i \in \mathcal{C}(x)} \{T_i\}}$ , corresponding to the productivity of the configuration of  $x$  that features the lowest takt time. The demand  $\Delta$  that must be met when solving the Linear Problem attempts for scenario  $d \in \mathcal{D}$  is the minimum between  $d$  and this maximum attainable demand. Lower-level real variables  $y_{i,p} \in [0, 1]$ ,  $p \in P$ ,  $i \in \mathcal{C}(x)$ , represent the percentage of period  $p$  spent producing with configuration  $i$ . Since lower-level LPs are independent and each refer to a demand scenario  $d \in \mathcal{D}$  which is a constant w.r.t. it, to simplify the notation we purposefully omitted the  $d$  index to  $y$  variables, which are to all intents and purposes the same as for model (3a)-(3m).

The LP model is:

$$\text{Minimize} \quad \sum_{i \in \mathcal{C}(x), p \in \Pi} v_p \cdot w_p \cdot q_i \cdot y_{i,p} \quad (5a)$$

$$\text{s.t.} \quad \sum_{i \in \mathcal{C}(x), p \in \Pi} \frac{v_p}{T_i} \cdot y_{i,p} \geq \Delta \quad (5b)$$

$$\sum_{i \in \mathcal{C}(x)} y_{i,p} \leq 1, \quad \forall p \in \Pi \quad (5c)$$

$$0 \leq y_{i,p} \leq 1, \quad \forall i \in \mathcal{C}(x), p \in \Pi \quad (5d)$$

Since  $v_p \cdot y_{i,p}$  is the lapse of time of TOU period  $p$  during which production uses configuration  $i \in \mathcal{C}(x)$ ,  $q_i \cdot v_p \cdot y_{i,p}$  is the associated energy consumption, and term (5a) is the overall energy cost to minimize. Constraint (5b) enforces satisfaction of demand  $\Delta$ , since  $\frac{v_p}{T_i} \cdot y_{i,p}$  is the corresponding produced quantity derived from takt  $T_i$ . Finally, (5c) forbids production in period  $p$  from exceeding its duration  $v_p$ : nonetheless, production can last less than  $v_p$ , in which case  $\sum_i y_{i,p} < 1$ .

It is worth noting that since some demand values  $d \in \mathcal{D}$  could not be accomplished by the configurations of set  $\mathcal{C}(x)$ , the assessment of the suitability of a balancing  $x$  could not be done based on the overall energy cost of objective function (5a). This is why the energy-related term (3c) of the upper-level problem concerns the per-produced-unit energy cost.

#### 4.3. Phase 3: Filtering non-dominated balancing

After Phase 2, each balancing has been evaluated w.r.t. the three objectives. More specifically, for each balancing, the value of the proxy function used during Phase 1 can be replaced by the per-produced-unit energy cost computed in Phase 2. Dominated balancing solutions w.r.t. the final three objectives can be discarded, i.e., those s.t. there exists another balancing with better values for all three objective functions. This yields a final set of design solutions that form the final Pareto front of the studied three-objective problem.

## 5. Computational results

In this section we describe the computational experience that we conducted to assess the performances of our approach.

### 5.1. Description of the computational experience

Four TOU cost profiles were considered, named EP1 to EP4: the first two are taken from, respectively, [51] and [52]; the last two are derived from energy contracts negotiated by two manufacturing companies with their electric energy providers. Profiles EP1 to EP4 are shown in Figure 2. In order to facilitate reading, and without loss of generality, time periods are ordered in the figure from left to right from the least to the most expensive one: this can be done because the problem does not take into account the time at which each period occurs, but only its duration. Each TOU cost profile covers a timespan  $H$  of 24 hours; tariffs are expressed in cost units per used energy unit.

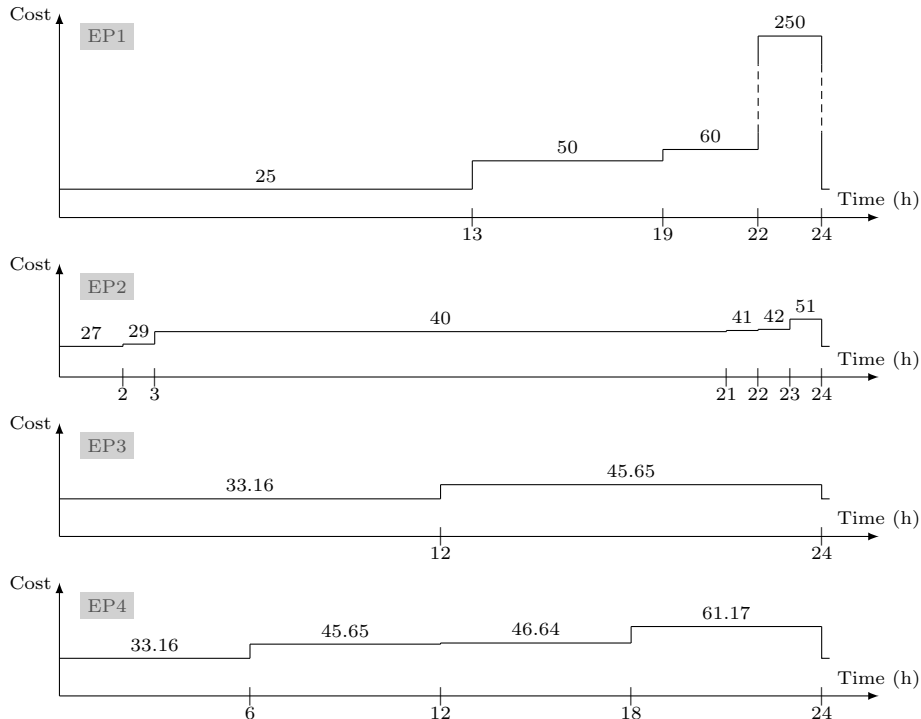


Figure 2: Detail of the considered TOU profiles.

We considered a total of 46 instances, derived from as many Simple Assembly Line Balancing Problems of type 1 (SALBP-1). A core line balancing problem in Combinatorial Optimization, SALBP-1 can describe various industrial environments. SALBP-1 consists in optimally assigning the tasks of an assembly line to a set of workstations, whose number has to be minimized, under a given maximum takt time limit and given task processing times and precedence constraints [21]. In this paper, we consider the whole set of 25 SALBP-1 instances from [53] and a representative sample of 21 of the instances of [54]. For each instance, we consider 20 demand scenarios, whose values are generated uniformly in a symmetrical interval centered on a reference value derived from the SALBP-1 instance; the interval width is  $\pm 25\%$  to preserve feasibility of the upper demand value of the range. Probabilities are then associated to demand scenarios following a discretized Gaussian probability distribution. This choice is motivated by the fact that the considered demand value represents what, in a manufacturing system, would be the accumulation of the demands of its many clients, which could be modeled by as many independent random variables. Then, based on the Central Limit Theorem, it is reasonable to consider that the random variable representing the aggregated demand follows a Gaussian distribution. More specifically, we choose a discretized distribution since the demand represents a number of manufactured items.

For each Scholl instance, the reference value is set on the median of the demand values considered in the group of original instances of the same name, which differ in the target takt time; for Otto instances, the value is set on the value obtained as  $\lfloor \frac{H}{T} \rfloor$ , with  $T = 1000$  being the takt common to all the original instances. In order to ensure that there is a feasible solution for all scenarios, we used IBM CPLEX 12.7.1 to solve, for each instance, the SALBP-1 corresponding to the largest demand scenario (using a 3

hours time limit). Out of the 25 instances from [53], three were disregarded, namely *Mertens*, *Bowman* and *Scholl*, because no feasible solution was found for them. For all Otto instances a feasible solution was found, thus yielding a final set of 43 instances: 22 from the Scholl set, with a number  $n$  of operations ranging from 9 to 148, and 21 from the Otto set, all with  $n = 20$ . Finally, for each instance, we associated each operation with an energy consumption  $e$  value proportional to its processing time  $t$  by a random value picked uniformly between 5 and 50 and generated independently for each operation.

Since the approach contains a random component in the MOSA, 10 replications of the matheuristic are run per instance and cost profile, and aggregated or averaged results are shown. In the MOSA, the descent factor  $\beta$  for the temperature is 0.98 and the initial temperature  $t$  is 10. In each replication, the MOSA is started 30 times: 15 directed starts, with the weights of the weighted sum varying by steps of 0.2, and 15 randomized starts. The total number of iterations and the length  $l$  of the steps are proportional to the number of operations  $n$  in the instance (resp.  $1250n$  and  $5n$ ). For each instance,  $\bar{n}$  is set to 40% of  $n$ ,  $\bar{r}$  to 3 and  $\bar{R}$  to  $n$ . All tests were conducted on an Intel Xeon E5-2660 v3 2.6 Ghz machine with 62.65Gb RAM.

### 5.2. Overall performance analysis of the proposed approach

In Table 3, the quality of the yielded Pareto fronts is evaluated via a hypervolume measure, averaged over all energy cost profiles. First, for each instance and cost profile, an overall front is obtained by merging those resulting from the replications. The merged front is actually the best known solution set. The hypervolume of the merged front is then computed. The indicator  $\overline{\%mrg}$  is the ratio of the hypervolume of the front of a replication to the hypervolume of the merged front, and  $\sigma$  is the standard deviation of hypervolume values for the replications. The average number of solutions in the front  $\overline{ndes}$ , as well as the average running time  $\overline{tmh}$  in seconds, are also reported.

Hypervolume figures show that the proposed method is stable in terms of the output Pareto front of balancing solutions: this can be seen by looking at the standard deviation of hypervolume values, which is never greater than 0.026 and 0.062 for Otto and Scholl instances, respectively. The yielded Pareto fronts also seem to be made up of good trade-off solutions, as the  $\overline{\%mrg}$  metric is never less than 97.1% and 91.7%, for Otto and Scholl instances, respectively. This also confirms that the method is stable w.r.t. the random component of the MOSA.

From a computational point of view, the matheuristic executes quickly, always taking less than 8 seconds for Otto instances, and in general less than 3 minutes for Scholl instances, except for Barthol2 which, due to being one of the instances with the highest number of operations (148) can take up to 8 minutes of computation. The computation times are essentially independent of the cost profile, which only impacts Phase 2 whose running times are negligible.

### 5.3. Analysis of the diversity of solutions obtained in the Pareto front

Tables 4 and 5 aim to analyze the diversity of solutions obtained in the Pareto front. For each instance and cost profile, each replication gives rise to a Pareto front of nondominated solutions, each solution being associated with a balancing and its performances w.r.t. the three objectives  $z_m$ ,  $z_{sl}$  and  $z_{ec}$  (see equations (3a) to (3c)). The minimum and maximum values of  $z_m$ ,  $z_{sl}$ ,  $z_{ec}$  are averaged over the replications and the respective relative ratios (denoted by  $\text{rlr}()$ ), i.e. the average maximum over the average minimum, are computed. The values  $\text{rlr}(z_m)$ ,  $\text{rlr}(z_{sl})$ ,  $\text{rlr}(z_{ec})$  compose the three first columns of each energy cost profile in Tables 4 and 5, while the fourth is the average number of solutions in the Pareto front.

The values obtained with each cost profile are quite similar for all four indicators, except for  $\text{rlr}(z_{ec})$ , which was to be expected since the main impact is the difference in costs between the time periods. The other indicators exhibit slight variations because the fronts yielded by the MOSA are not identical and the evaluation of these solutions and the filtering of Phase 3 depend on the energy profile.

The average number of solutions in the front is strongly correlated with the amplitude ratios  $\text{rlr}(z_m)$ ,  $\text{rlr}(z_{sl})$ ,  $\text{rlr}(z_{ec})$ . For Otto instances, if we compute the correlation coefficient for all the pairs among  $\text{rlr}(z_m)$ ,  $\text{rlr}(z_{sl})$ ,  $\text{rlr}(z_{ec})$ , we get no less than 0.918 (which occurs between  $\text{rlr}(z_m)$  and  $\text{rlr}(z_{sl})$ ). Similarly, the correlation coefficients between the three relative ratio values and  $\overline{ndes}$ , the average number of solutions in the Pareto front, are respectively 0.962, 0.871 and 0.954. Moreover,  $\overline{ndes}$  is in most cases either greater than or equal to 15.7, or strictly less than 2.2. This seems to suggest that, in general and regardless of the cost profile, Otto instances lead to a Pareto front with either very few or a good variety

$1, \dots, n$	set of operations
$1, \dots, m$	set of stations
$\bar{n}$	maximum number of operations assigned to the same station
$\mathcal{P}$	set of precedences between operations ( $(j, g) \in \mathcal{P}$ means that operation $j$ is a predecessor of operation $g$ )
$t_j$	processing time of operation $j$
$e_j$	total energy consumed during the processing of operation $j$ (i.e., the integral of the power consumption profile of operation $j$ over its processing time)
$\Pi$	set of periods for the TOU energy cost
$\mathcal{D}$	set of demand scenarios
$\sigma_d$	probability of occurrence of scenario $d$ , defined such that $\sum_{d \in \mathcal{D}} \sigma_d = 1$ (e.g., equal to $\frac{1}{ \mathcal{D} }$ in case of equiprobable distribution)
$H$	time horizon for the planning
$v_p$	duration of period $p$
$w_p$	energy cost during period $p$
$\alpha$	residual energy consumption factor during idle time
$\mathcal{C}(x)$	set of available configurations using the balancing $x$
$r_{i,k}$	number of resources used in configuration $i$ for station $k$
$\bar{r}$	maximum number of resources per station
$\bar{R}$	maximum number of resources for the whole line
$T_i$	takt time of the RMS when using configuration $i$
$q_i$	energy consumed per unit of time when configuration $i$ is used

Table 2: Notations used for the problem data

	hypervolume		ndes	tmh (s)		hypervolume		ndes	tmh (s)
	%mrg	$\sigma$				%mrg	$\sigma$		
Otto025	97.9%	0.012	19.8	7.35	Jaeschke	96.3%	0.000	4.5	1.53
Otto050	97.8%	0.019	2.2	4.22	Jackson	97.4%	0.000	3.0	1.95
Otto075	98.8%	0.000	1.7	3.13	Mansoor	98.1%	0.000	1.0	1.40
Otto100	97.3%	0.016	16.7	6.95	Mitchell	97.9%	0.018	4.7	4.59
Otto125	98.3%	0.014	3.2	4.93	Roszieg	97.8%	0.017	8.8	8.13
Otto150	98.7%	0.004	1.7	3.08	Heskiaoff	98.7%	0.014	3.1	6.23
Otto175	98.4%	0.001	18.7	7.00	Buxey	98.6%	0.008	18.9	12.41
Otto200	99.0%	0.000	12.2	5.99	Sawyer	97.6%	0.037	14.7	13.30
Otto225	98.8%	0.000	1.7	3.20	Lutz1	98.8%	0.009	14.8	14.41
Otto250	98.3%	0.000	15.8	7.11	Gunther	98.8%	0.015	16.2	16.09
Otto275	98.8%	0.000	4.6	4.96	Kilbridge	96.6%	0.027	12.3	15.89
Otto300	98.9%	0.000	2.1	4.20	Hahn	96.9%	0.034	4.6	16.24
Otto325	98.5%	0.001	16.9	7.05	Warnecke	96.6%	0.025	15.1	60.10
Otto350	98.9%	0.000	4.4	5.04	Tonge	97.4%	0.018	12.5	74.78
Otto375	98.5%	0.004	1.5	3.22	Wee-mag	96.3%	0.010	16.9	98.53
Otto400	97.1%	0.026	16.9	7.30	Arcus1	91.7%	0.062	8.1	75.19
Otto425	98.5%	0.014	9.6	5.58	Lutz2	95.4%	0.043	13.2	127.94
Otto450	98.5%	0.004	1.6	3.18	Lutz3	93.5%	0.043	10.6	116.27
Otto475	97.9%	0.012	15.7	6.92	Mukherje	94.3%	0.049	9.3	133.98
Otto500	98.4%	0.014	9.2	6.32	Arcus2	96.7%	0.032	11.1	160.91
Otto525	98.7%	0.004	1.5	3.26	Barthol2	95.3%	0.030	11.9	440.57
					Barthold	96.3%	0.029	4.8	166.03

Table 3: Quality and features of replications w.r.t. best known solutions, averaged over all energy cost profiles.

	EP1			EP2			EP3			EP4			
	$\text{rlr}(z_m)$	$\text{rlr}(z_{al})$	$\text{rlr}(z_{ec})$	$\text{rlr}(z_m)$	$\text{rlr}(z_{al})$	$\text{rlr}(z_{ec})$	$\text{rlr}(z_m)$	$\text{rlr}(z_{al})$	$\text{rlr}(z_{ec})$	$\text{rlr}(z_m)$	$\text{rlr}(z_{al})$	$\text{rlr}(z_{ec})$	$\text{ndes}$
Otto025	3.81	1.23	1.73	20.4	1.18	1.18	18.4	1.21	1.21	1.18	1.21	1.22	21.8
Otto050	1.13	1.00	1.00	1.4	1.06	1.06	2.9	1.00	1.00	1.00	1.00	1.00	2.9
Otto075	1.10	1.00	1.00	1.3	1.04	1.04	2.0	1.13	1.00	1.00	1.00	1.00	2.0
Otto100	3.22	1.24	1.73	17.0	1.16	1.16	15.7	3.56	1.21	1.18	1.21	1.21	19.6
Otto125	1.29	1.00	1.00	1.8	1.08	1.08	4.2	1.48	1.00	1.02	1.00	1.00	4.2
Otto150	1.07	1.00	1.00	1.2	1.04	1.04	2.2	1.10	1.00	1.00	1.00	1.00	2.2
Otto175	3.67	1.22	1.77	19.6	1.16	1.16	17.1	3.67	1.21	1.16	1.21	1.21	21.4
Otto200	1.37	1.00	1.10	10.4	1.11	1.11	13.9	1.47	1.00	1.06	1.00	1.00	13.9
Otto225	1.03	1.00	1.00	1.1	1.05	1.05	2.1	1.03	1.00	1.00	1.00	1.00	2.3
Otto250	3.73	1.25	1.85	15.8	1.12	1.12	15.4	3.73	1.17	1.15	1.17	1.17	18.5
Otto275	1.40	1.00	1.04	3.6	1.06	1.06	5.5	1.40	1.00	1.04	1.00	1.00	5.5
Otto300	1.00	1.00	1.00	1.0	1.06	1.06	3.0	1.00	1.00	1.00	1.00	1.00	3.3
Otto325	3.87	1.37	1.84	20.3	1.31	1.31	15.1	4.03	1.35	1.30	1.35	1.35	19.1
Otto350	1.53	1.00	1.06	3.6	1.07	1.07	5.1	1.53	1.00	1.04	1.00	1.00	5.1
Otto375	1.03	1.00	1.00	1.1	1.05	1.05	2.0	1.03	1.00	1.00	1.00	1.00	1.9
Otto400	3.76	1.26	1.68	19.5	1.22	1.22	13.9	3.76	1.25	1.23	1.25	1.25	18.8
Otto425	1.55	1.00	1.10	7.8	1.09	1.09	11.3	1.61	1.00	1.06	1.00	1.00	11.3
Otto450	1.07	1.00	1.00	1.2	1.04	1.04	2.0	1.03	1.00	1.00	1.00	1.00	2.1
Otto475	3.39	1.27	1.76	15.4	1.18	1.18	15.7	3.52	1.26	1.18	1.26	1.26	18.2
Otto500	1.94	1.00	1.31	8.5	1.12	1.12	7.2	2.07	1.00	1.13	1.00	1.00	12.5
Otto525	1.03	1.00	1.00	1.1	1.04	1.04	2.0	1.00	1.00	1.00	1.00	1.00	2.0

Table 4: Comparison of the different energy cost profile regarding the amplitude ratios of the three objectives  $z_m$ ,  $z_{al}$ ,  $z_{ec}$  and the average number of balancing solutions in the Pareto front, for Otto instances.

	EPI			EP2			EP3			EP4		
	rlr( $z_m$ )	rlr( $z_{s1}$ )	ndes	rlr( $z_m$ )	rlr( $z_{s1}$ )	ndes	rlr( $z_m$ )	rlr( $z_{s1}$ )	ndes	rlr( $z_m$ )	rlr( $z_{s1}$ )	ndes
Jaeschke	1.67	1.00	3.0	2.67	1.00	6.0	1.67	1.00	3.0	2.67	1.00	6.0
Jackson	1.33	1.00	2.0	2.00	1.00	4.0	1.33	1.00	1.03	2.00	1.00	4.0
Mansoor	1.00	1.00	1.0	1.00	1.00	1.0	1.00	1.00	1.00	1.00	1.00	1.0
Mitchell	1.53	1.00	4.0	1.91	1.00	5.2	1.56	1.00	4.1	2.13	1.00	1.13
Roszig	2.18	1.07	7.4	3.33	1.07	9.9	2.18	1.07	1.13	3.33	1.07	10.6
Heskiaoff	1.32	1.00	2.0	2.19	1.00	4.1	1.44	1.00	1.04	2.06	1.00	3.9
Buxey	3.07	1.29	18.9	4.61	1.17	17.3	3.19	1.17	1.14	4.65	1.17	22.6
Sawyer	2.91	1.37	18.0	4.18	1.15	14.0	3.06	1.15	1.13	4.18	1.15	15.8
Lutz1	2.84	1.24	14.9	3.94	1.08	14.3	2.97	1.14	1.15	4.00	1.18	17.5
Gunther	3.25	1.35	24.7	3.78	1.13	12.4	3.41	1.13	1.13	3.81	1.13	15.0
Kilbridge	1.83	1.05	14.4	2.56	1.06	11.5	1.48	1.04	1.07	2.54	1.03	13.4
Hahn	1.51	1.00	3.7	1.98	1.01	5.0	1.78	1.01	1.06	1.84	1.01	5.0
Warnecke	5.33	2.13	16.4	5.39	2.13	13.7	5.35	2.13	1.98	5.44	2.13	15.2
Tonge	4.89	1.59	11.4	6.09	1.59	13.3	5.09	1.59	1.53	6.09	1.59	14.4
Wee-mag	5.70	3.52	18.3	5.70	3.51	15.9	5.70	3.51	3.21	5.70	3.51	16.6
Arcus1	1.85	1.11	7.2	2.28	1.10	8.4	1.85	1.11	1.13	2.30	1.11	9.7
Lutz2	4.37	2.09	13.8	4.35	2.09	12.4	4.37	2.09	2.07	4.39	2.09	13.3
Lutz3	2.83	1.34	9.9	3.72	1.33	10.6	2.98	1.33	1.39	3.78	1.33	12.4
Mukherje	3.04	1.42	7.8	4.46	1.42	10.3	3.09	1.42	1.40	4.55	1.42	11.1
Arcus2	3.66	1.70	9.6	5.86	1.78	11.9	4.51	1.81	1.58	5.85	1.78	13.4
Barthol2	4.83	2.43	12.4	4.83	2.43	11.2	4.83	2.43	2.32	4.83	2.43	11.9
Barthold	1.58	1.06	6.4	1.78	1.04	4.3	1.68	1.04	1.08	1.96	1.04	4.7

Table 5: Comparison of the different energy cost profile regarding the amplitude ratios of the three objectives  $z_m$ ,  $z_{s1}$ ,  $z_{ec}$  and the average number of balancing solutions in the Pareto front, for Scholl instances.



of solutions, and in the second case, the ranges of values of the three ratios are of equal amplitude, with no consistent preference or disadvantage of one against the others. These correlations are in general weaker for Scholl instances: between 0.793 ( $\text{rlr}(z_m)$  and  $\text{rlr}(z_{sl})$ ) and 0.993 ( $\text{rlr}(z_{ec})$  and  $\text{rlr}(z_{sl})$ ) for pairs of objective relative ratios; between 0.507 and 0.746 between each objective and  $\overline{\text{ndes}}$ . Besides, the values of  $\overline{\text{ndes}}$  are much more scattered. However, all these facts could result from Scholl instances being much more heterogeneous in terms of features, notably the number  $n$  of operations.

It is nonetheless noteworthy that the correlation between  $\text{rlr}(z_{ec})$  and  $\text{rlr}(z_{sl})$  appears to be strong for both Otto and Scholl instances, suggesting a relation between service level and economic cost. This can be explained as follows. As seen in Section 3, the lower-level planning problem tries first to satisfy the highest percentage of the demand  $d \in \mathcal{D}$  associated with a given scenario, and then to satisfy this percentage with the lowest possible per-produced-unit energy cost. Section 4.2 then highlighted that any given balancing  $x$  has a maximum attainable demand  $\frac{H}{\min_{i \in \mathcal{C}(x)} \{T_i\}}$ : as long as  $d$  is much lower than this limit,  $d$  can be easily met ( $z_{sl} = 1$ ) without producing during the costliest TOU period  $\bar{p}$ . However, as  $d$  approaches the maximum attainable demand, producing during period  $\bar{p}$  becomes necessary, causing  $z_{ec}$  to increase dramatically (in particular in cases such as EP1 where the most expensive period is relatively short but with a strong cost ratio compared to less costly periods). Finally, when  $d \geq \frac{H}{\min_{i \in \mathcal{C}(x)} \{T_i\}}$ , we have  $z_{sl} \leq 1$  and the highest possible energy cost  $z_{ec}$ , due to the entire timespan being spent producing with the most productive configuration of  $x$ , which in most cases is not the least consuming. Hence, we either have values  $z_{sl} = 1$  and low  $z_{ec}$  values, or  $z_{sl} = 1$  and near-maximum  $z_{ec}$  values, or  $z_{sl} < 1$  and maximum  $z_{ec}$  values, which can be empirically considered as the reason for the observed correlation.

For the number of stations, the relative ratio on Otto instances is always less than 4.55, and on average 2.42; the figures are respectively less than 6.1 and 3.3 for Scholl instances. Indeed, a greater number of stations enables one to increase the overall productivity, at the cost of an increase in energy consumption, which can be worthwhile for producing only in lower price periods to satisfy the demand. Conversely, having a low number of stations can lead to a low per-unit energy consumption, but the demand may not be entirely met.

At each instance, cost profile and replication, at least one solution with a service level equal to 1 is yielded, the worst one having on average a level greater than 73.7% for the 21 Otto instances, and 28.5% for the 22 Scholl instances.

#### 5.4. Analysis of the usage of configurations during the planning phase

In this section, we aim to analyse the number of configurations used to meet all the demands. For each instance and cost profile  $\Pi$ , and for each replication and balancing solution  $x$  of the Pareto front, we compute the expected value of the number of configurations  $|\{i \in \mathcal{C}(x) : (\exists p \in \Pi) y_{i,p,d} > 0\}|$  used to meet a demand scenario  $d$ , as well as the number of all the configurations used to meet at least one of them,  $|\{i \in \mathcal{C}(x) : (\exists p \in \Pi, d \in \mathcal{D}) y_{i,p,d} > 0\}|$ . These figures are denoted in the following, respectively, as  $E[\#\text{uc}]$  and  $\text{tot} \#\text{uc}$  for conciseness.  $E[\#\text{uc}]$  can be seen as an indicator of the number of reconfigurations required during the time horizon over which a demand scenario must be met;  $\text{tot} \#\text{uc}$  indicates, for a balancing solution  $x$ , how many of the configurations of  $\mathcal{C}(x)$  are actually needed to face every possible scenario.

Figure 3 illustrates how  $E[\#\text{uc}]$  and  $\text{tot} \#\text{uc}$  are computed. It depicts, in terms of percentage of the time horizon used to attain a given demand ( $\%H$ , reported on the vertical axis) how one of the nondominated solutions  $x$  of one of the replications for cost profile EP1, instance Mukherje, reacts to all possible demand values from  $d=150$  to  $d=500$ .

Discrete values  $\mathcal{D} = \{277, 286, \dots, 461\}$  represent the demand scenarios of the corresponding discretized Gaussian probability distribution, as shown by the boxes associated with them, but the figure covers a wider set of demand values to assess the behavior of  $x$  beyond the considered scenarios. RMS  $x$  has  $|\mathcal{C}(x)| = 19$  configurations, but –as we will show– only three are used all along the demand range  $150 \leq d \leq 500$ . Hence, for the sake of simplicity, we will consider  $\tilde{\mathcal{C}}(x) = \{c_0, c_1, c_2\}$ . Configuration features  $T_i$  (takt time) and  $q_i T_i$  (per-produced-unit energy consumption) are reported in the figure:  $c_2$  is the most productive one (having by construction the lowest takt time, see Section 3);  $c_0$  is the least energy-consuming one; finally,  $c_1$  has intermediate features, i.e., it is less consuming and less productive than  $c_2$ , but also more consuming and more productive than  $c_0$ . On the left, TOU periods are indexed from the least to the most expensive  $p=0$  to  $p=3$ , as depicted by the red strips on the right. Demand  $d=150$  only requires  $c_0$ , as its takt time  $T_0$  is largely sufficient to meet  $d$  in 58.6% of period  $p=0$ . The

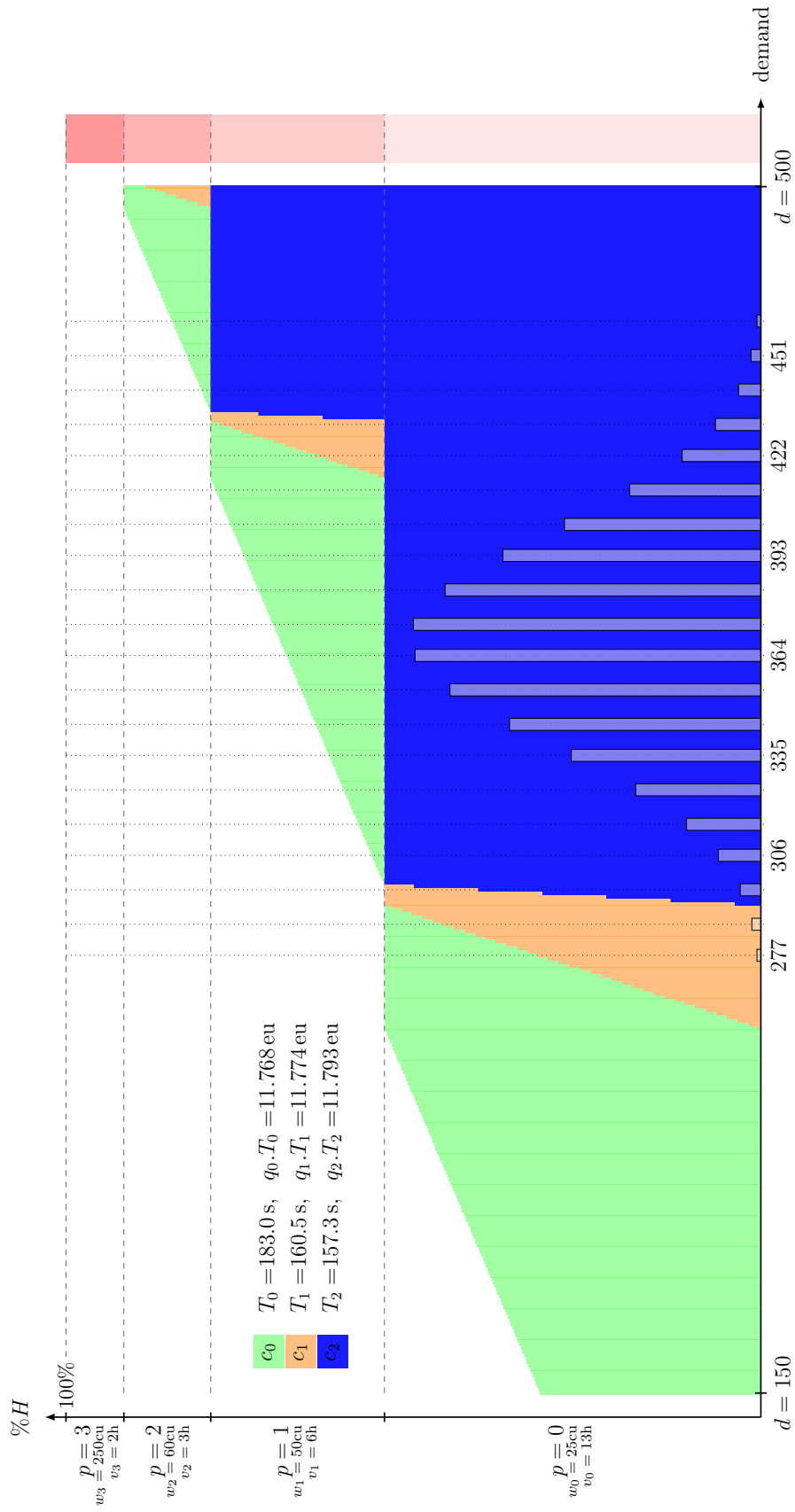


Figure 3: Behavior of an RMS solution for Mukherje instance, cost profile EPI, faced with a wide, continuous range of demand values.

	EP1				EP2				EP3				EP4			
	$E[\#uc]$		tot #uc		$E[\#uc]$		tot #uc		$E[\#uc]$		tot #uc		$E[\#uc]$		tot #uc	
	min	max	min	max	min	max	min	max	min	max	min	max	min	max	min	max
Otto025	1.00	2.22	1.0	3.0	1.00	2.47	1.0	3.0	1.00	2.08	1.0	2.4	1.00	2.14	1.0	2.7
Otto050	1.30	1.67	1.3	1.7	1.17	2.08	1.7	2.9	1.20	1.31	1.3	1.5	1.00	1.70	1.0	1.8
Otto075	1.70	1.90	1.7	1.9	1.37	1.87	1.9	2.8	1.06	1.09	1.2	1.3	1.00	1.40	1.0	1.5
Otto100	1.00	2.09	1.0	3.0	1.00	2.38	1.0	3.0	1.00	2.07	1.0	2.4	1.00	2.08	1.0	2.8
Otto125	1.47	1.66	1.5	1.8	1.00	1.85	1.0	2.6	1.00	1.36	1.0	1.4	1.00	1.53	1.0	1.8
Otto150	1.60	1.80	1.7	1.8	1.38	2.09	2.0	3.0	1.00	1.05	1.0	1.1	1.11	1.60	1.2	1.8
Otto175	1.00	2.13	1.0	3.0	1.00	2.31	1.0	3.0	1.00	2.06	1.0	2.5	1.00	2.02	1.0	2.5
Otto200	1.03	1.76	1.2	3.0	1.00	2.05	1.0	3.0	1.00	2.00	1.0	2.5	1.00	2.00	1.0	2.3
Otto225	2.32	2.36	2.8	2.9	1.45	2.09	2.4	3.0	1.82	1.82	2.7	2.7	1.08	1.80	1.1	1.8
Otto250	1.00	2.09	1.0	3.0	1.00	2.25	1.0	3.0	1.00	2.02	1.0	2.6	1.00	2.01	1.0	3.0
Otto275	1.18	1.79	1.8	3.0	1.09	2.21	1.3	3.0	1.04	1.98	1.3	2.2	1.00	1.91	1.0	2.0
Otto300	1.93	1.93	2.3	2.3	1.11	2.06	1.5	3.0	1.49	1.49	1.5	1.5	1.00	1.70	1.0	2.1
Otto325	1.00	2.03	1.0	3.0	1.00	2.21	1.0	2.9	1.00	2.00	1.0	2.3	1.00	2.14	1.0	2.7
Otto350	1.05	1.75	1.2	2.6	1.00	2.00	1.0	3.0	1.01	1.86	1.1	1.9	1.00	1.97	1.1	2.2
Otto375	1.93	2.02	2.0	2.1	1.20	1.78	2.0	3.0	1.02	1.02	1.1	1.1	1.04	1.71	1.3	1.8
Otto400	1.00	2.18	1.0	3.0	1.00	2.16	1.0	2.9	1.00	2.00	1.0	2.3	1.00	2.08	1.0	2.9
Otto425	1.09	1.89	1.1	2.9	1.00	2.23	1.0	3.0	1.00	1.90	1.0	2.2	1.00	2.06	1.0	2.4
Otto450	1.09	1.19	1.1	1.2	1.04	1.55	1.2	2.6	1.82	1.91	2.7	2.8	1.08	1.80	1.1	2.0
Otto475	1.00	2.17	1.0	3.0	1.00	2.45	1.0	3.0	1.00	2.00	1.0	2.5	1.00	2.00	1.0	3.0
Otto500	1.01	2.03	1.1	2.9	1.00	2.10	1.0	2.8	1.00	2.00	1.0	2.2	1.00	2.05	1.0	2.1
Otto525	1.72	1.82	2.0	2.1	1.35	1.95	2.0	2.7	1.03	1.03	1.3	1.3	1.19	1.64	1.3	1.9

Table 6: Average usage of configurations by proposed RMS solutions, Otto instances.

usage of  $c_0$  in  $p=0$  increases up to a demand value of  $d = \lfloor \frac{13.3600}{183} \rfloor = 255$ , after which the optimal solution is to partly use  $c_1$  so as to avoid producing during  $p=1$ . This is due to the cost ratio  $\frac{w_1}{w_0} = 2$  being much greater than  $\frac{q_1 \cdot T_1}{q_0 \cdot T_0} \approx 1$ , the ratio between the per-unit energy consumptions of  $c_1$  and  $c_0$ . Using  $c_1$  gradually replaces  $c_0$  up to a value  $d=291$ , after which, based on the same principle, the usage of  $c_1$  is gradually replaced by  $c_2$ . Demand  $d=297$  is the last one that can be covered entirely by producing during the least expensive period only, and this by using the most productive configuration: from  $d=298$  on, a residual demand exists that is met by using  $c_0$  during period  $p=1$ . It is easy to see that this scheme repeats up to  $d=500$ . In this example, the number of configurations used is 2 for each demand scenario, except  $d=422$  and the following ( $d=431$ ), for which 3 are required along the same timespan. Note that scenarios requiring two configurations do not always use the same two, e.g.,  $d=277$  uses  $c_0$  and  $c_1$ , while  $d=335$  uses  $c_0$  and  $c_2$ . The associated probability values range from 0.13% for  $d=271$  and  $d=461$  to 12.43% for  $d=364$  and 12.48% for  $d=373$  (the following): the corresponding weighted sum gives a value of 2.045 for  $E[\#uc]$ . As for  $\text{tot} \#uc$ , its value is 3, as the subset of the configurations actually used for at least one scenario has only 3 of the original 19 members.

Tables 6 and 7 provide an analysis of how the RMS solutions proposed by the matheuristic use their configuration set, on Otto and Scholl instances, respectively. Table 6 reports, for each Otto instance and cost profile, the average over the 10 replications of the minimum and maximum values of  $E[\#uc]$  (denoted  $E[\#uc]_{\min}$  and  $E[\#uc]_{\max}$ , respectively) and  $\text{tot} \#uc$  (denoted  $\text{tot} \#uc_{\min}$  and  $\text{tot} \#uc_{\max}$ , respectively) among the Pareto front solutions of each replication. Average intervals per instance and cost profile are thus obtained for the two figures. The same goes for Table 7 for Scholl instances.

As regards  $E[\#uc]$ , we notice in both Tables 6 and 7 for many instances an average minimum value,  $E[\#uc]_{\min}$ , equal to 1: for these instances, all replications yield at least one balancing solution of the front for which every demand scenario can be dealt with by only one configuration, and thus no reconfiguration is required. Many of these instances occur: 16 out of 22 and 11.5 out of 21, on average over all cost profiles, for Scholl and Otto instances, respectively. If in particular for the same instance we also have  $\text{tot} \#uc_{\min}$  equal to 1 (e.g. Otto025 for every cost profile), then the same balancing uses the same configuration for all scenarios. On the other hand,  $E[\#uc]_{\min}$  being strictly greater than 2 for an instance and cost profile might indicate that every balancing solution of the front of every replication requires at least two configurations for every possible demand scenario. Finally, we notice that  $E[\#uc]_{\max}$  is always strictly less than 3. As for  $\text{tot} \#uc_{\max}$ , we note that it is rarely strictly greater than 3.

If we look at the gap between  $E[\#uc]_{\min}$  and  $E[\#uc]_{\max}$ , with all instances (Otto and Scholl) considered, we find values of 0.77 for EP1, 1.03 for EP2, 0.71 for EP3 and 0.94 for EP4. Indeed, EP1 and EP3 have similar structures: EP3 is divided into two very large periods (12h each) with an energy cost ratio between the two not greater than 1.38, and EP1 has 13 hours with the lowest cost and 9 more where the cost rises according to a ratio not greater than 2.4. On the other hand, EP2 and EP4 have a wide central interval with a constant (or nearly constant) value and much shorter extreme intervals whose costs differ significantly. This seems to explain both why the gap between  $E[\#uc]_{\min}$  and  $E[\#uc]_{\max}$  is narrower with EP1 and EP3, and why for these two the average value of  $E[\#uc]_{\max}$ , all instance considered, is less than 2: presumably, in this cases the balancing solutions try to meet as much demand as possible of the demand during the cheapest period with the least consuming configuration, and only for

	EP1				EP2				EP3				EP4			
	$E[\#uc]$		tot #uc		$E[\#uc]$		tot #uc		$E[\#uc]$		tot #uc		$E[\#uc]$		tot #uc	
	min	max	min	max	min	max	min	max	min	max	min	max	min	max	min	max
Jaeschke	1.00	1.62	1.0	2.0	1.00	2.00	1.0	3.0	1.00	1.90	1.0	2.0	1.00	2.00	1.0	2.0
Jackson	1.65	2.49	3.0	3.0	1.00	2.00	1.0	2.0	1.97	2.00	2.0	2.0	1.00	2.00	1.0	2.0
Mansoor	1.97	1.97	2.0	2.0	1.86	1.86	3.0	3.0	1.50	1.50	2.0	2.0	2.00	2.00	2.0	2.0
Mitchell	1.13	2.03	1.3	2.9	1.07	2.08	1.1	3.0	1.09	1.70	1.2	2.2	1.00	1.80	1.0	1.8
Roszieg	1.01	1.97	1.1	2.9	1.00	2.00	1.0	3.0	1.00	1.94	1.0	2.0	1.00	2.00	1.0	2.0
Heskiaoff	1.25	1.66	1.6	2.0	1.00	2.04	1.0	2.8	1.10	1.72	1.1	2.0	1.00	1.90	1.0	2.1
Buxey	1.00	2.06	1.0	3.0	1.00	2.04	1.0	2.9	1.00	2.00	1.0	2.3	1.00	2.01	1.0	2.8
Sawyer	1.00	1.96	1.0	2.8	1.00	2.11	1.0	2.6	1.00	2.00	1.0	2.2	1.00	2.12	1.0	2.6
Lutz1	1.00	2.13	1.0	3.0	1.00	2.17	1.0	2.8	1.00	2.00	1.0	2.1	1.00	2.04	1.0	2.8
Gunther	1.00	2.03	1.0	3.0	1.00	2.10	1.0	2.6	1.00	2.00	1.0	2.2	1.00	2.01	1.1	2.5
Kilbridge	1.00	1.81	1.1	2.7	1.10	2.19	1.1	2.9	1.08	1.63	1.2	2.0	1.00	1.86	1.0	2.2
Hahn	1.00	1.43	1.0	2.0	1.00	1.70	1.0	1.9	1.00	1.62	1.1	2.2	1.00	1.76	1.0	2.1
Warnecke	1.00	2.44	1.0	3.2	1.00	2.26	1.0	2.8	1.00	2.02	1.0	3.0	1.00	2.39	1.0	2.8
Tonge	1.00	2.27	1.0	2.6	1.00	2.28	1.0	2.6	1.00	2.06	1.0	2.6	1.00	2.28	1.0	2.7
Wee-mag	1.00	2.46	1.0	2.8	1.00	2.39	1.0	2.6	1.00	2.02	1.0	2.7	1.00	2.12	1.0	2.8
Arcus1	1.00	1.76	1.0	2.0	1.00	2.03	1.0	2.2	1.00	1.73	1.0	2.1	1.00	2.00	1.0	2.3
Lutz2	1.00	2.08	1.0	2.7	1.00	2.08	1.0	2.4	1.00	2.00	1.0	2.6	1.00	2.27	1.0	2.6
Lutz3	1.00	2.20	1.0	3.0	1.00	2.33	1.0	2.8	1.00	2.05	1.0	2.9	1.00	2.37	1.0	3.1
Mukherje	1.00	2.13	1.0	2.7	1.00	2.52	1.0	2.6	1.00	1.92	1.0	2.5	1.00	2.48	1.0	2.9
Arcus2	1.00	1.92	1.0	2.1	1.00	2.40	1.0	2.4	1.00	2.08	1.0	2.3	1.00	2.12	1.0	2.4
Barthold	1.00	2.03	1.0	2.9	1.10	2.14	1.1	2.5	1.10	2.01	1.1	2.8	1.10	2.55	1.1	2.8
Barthold	1.06	1.85	1.2	2.1	1.20	2.00	1.2	2.0	1.07	1.83	1.1	2.0	1.24	2.01	1.3	2.1

Table 7: Average usage of configurations by proposed RMS solutions, Scholl instances.

the highest demand scenarios possibly resort to the most productive, costlier one, since the cost ratio is not severe. For EP2 and EP4, on the other hand, when the demand to be met is high, before switching from one of the cheapest to the most productive configuration, intermediate ones (i.e. more productive than the cheapest ones, and less consuming than the most productive) are used, before resorting to the most productive only to avoid production during the costliest TOU period. Of course, EP1 differs from EP3 in that it includes a short, much more expensive TOU period of 2h, which explains the higher gap between  $E[\#uc]_{\min}$  and  $E[\#uc]_{\max}$ .

Lastly, we note that the gap between the values of  $E[\#uc]_{\max}$  and  $\text{tot \#uc}_{\max}$ , when all instances and cost periods are considered, is on average 0.49 and never greater than 1.25 overall. This seems to suggest that in most cases, of all the configurations  $\{|i \in \mathcal{C}(x) : (\exists p \in \Pi, d \in \mathcal{D}) y_{i,p,d} > 0\}$  that a balancing solution  $x$  uses to meet all demand scenarios, only a few or even none are used sporadically, while most of them are used to cope with most demand scenarios.

### 5.5. Comparison with dedicated lines

In order to evaluate the impact of using RMS rather than classical serial lines, the results obtained by our approach for each of the 43 instances have been compared with 5 dedicated lines with one resource on each workstation. These dedicated lines correspond to 5 different system paradigms that could be adopted in industry to deal with the uncertain nature of the demand, i.e., designed to meet to the demand associated with the quantiles 50%, 75%, 90%, 95%, 100% of the distribution. The first one aims to meet the median demand and is widely used in industry to achieve a good service level while keeping the number of stations relatively low. The last one considers the worst-case scenario, designing a highly productive system, possibly at the cost of a large number of stations. In the following, we refer to the 5 paradigms as DL50% to DL100%. For all the instances and the 5 quantiles considered, an optimal SALBP-1 solution is sought with IBM CPLEX within a time limit of 3 hours.

To provide the reader with a visual understanding of the proposed comparison, Figure 4 illustrates it on one of the instances, by using the same graphic convention as in Figure 3. The uppermost subfigure depicts how one of the nondominated solutions of one of the replications for cost profile EP1, instance Mitchell, reacts to all possible demand values from  $d = 1800$  to  $d = 6000$ . The demand scenarios of the discretized Gaussian probability distribution are also shown. Only 2 configurations are needed to meet the full range of demands, with the least-consuming one (green color) being sufficient to cover all scenarios in the quantile 80% of the distribution (i.e. up to demand scenario  $d = 4384$ ). The RMS solution is capable of efficiently using very few of its configurations to reduce production time and keep energy-related costs low while meeting the demand. The middle subfigure refers to dedicated line DL100%, which can by definition satisfy all demand scenarios and guarantee to have  $z_{sl} = 1$ , but at a much higher energy cost, since with equal demand, a much higher portion of the time horizon is required to satisfy it, and the most expensive period can be needed. Additionally, higher demand values cannot be reached within the time horizon, giving rise to backorders. These limits are even more evident for dedicated line DL50%, depicted by the lowermost subfigure: in this case, backorders can occur also for the demand values associated with some of the scenarios  $d \in \mathcal{D}$ , causing  $z_{sl}$  to be strictly less than 1. The concerned scenarios are highlighted in the figure.

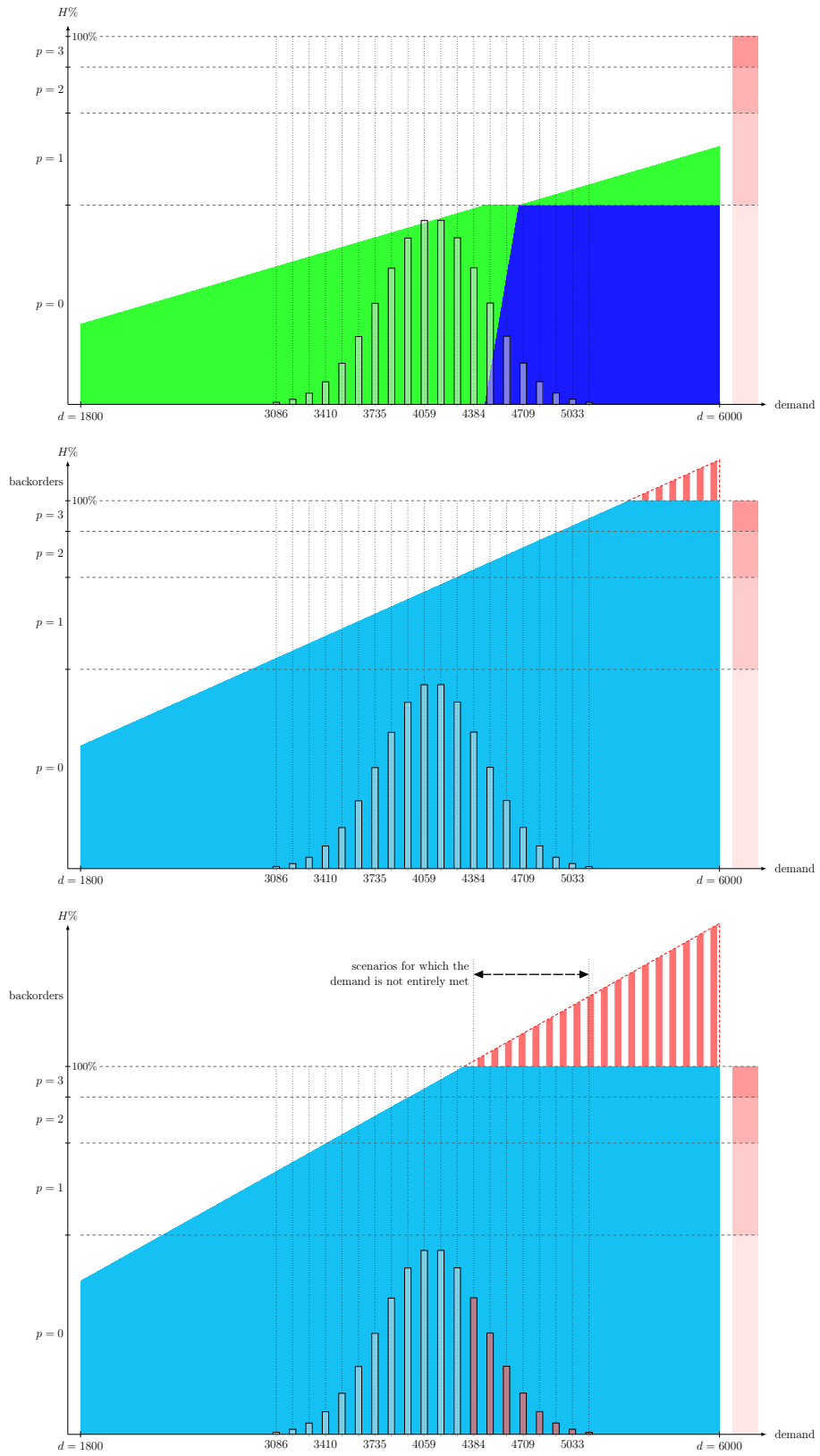


Figure 4: Comparison of an RMS nondominated solution (top) with dedicated lines DL100% (middle) and DL50% (bottom), Mitchell instance, cost profile EP1.

Tables 8 and 9 show the result of the comparison between RMS solutions obtained by our approach and dedicated line solutions yielded by the SALBP-1 solver. Table 8 compares the performances of an RMS yielded by the matheuristic and the 5 dedicated lines. The RMS for the comparison is chosen as follows: for each replication, the solution from the Pareto front obtained by MOSA which is the nearest (in terms of normalized distance) to the ideal point (given by the best values of each of the objective functions) is considered; the values of the three objectives  $z_m$  to  $z_{ec}$  of the 10 solutions obtained –one per replications– are averaged, yielding a reference RMS solution denoted RMS\*. In Table 8, for each criterion, five columns are reported, detailing the relative performances (percentage variations) of RMS\* and DL100% to DL75% compared to the performance of the dedicated line, DL50% (not shown) designed to respond to the demand associated with the 50% quantile of the distribution. For each instance, the executions on the four energy profiles are averaged. The complementary average values, computed for each energy profile over all the instances, are given at the bottom of the same table.

Table 9 is built similarly to Table 8 for the expected value of the used percentage of the time horizon,  $E[\%H]$ , except that all quantile DL and RMS\* are shown and values are not relative to DL50% but absolute values instead.

	$z_m$ (min)					$z_{sl}$ (max)					$z_{ec}$ (min)				
	RMS*	DL 100%	DL 95%	DL 90%	DL 75%	RMS*	DL 100%	DL 95%	DL 90%	DL 75%	RMS*	DL 100%	DL 95%	DL 90%	DL 75%
Otto025	-39.77	54.55	36.36	27.27	9.09	2.44	2.56	2.46	2.26	1.44	-15.43	-4.88	-1.77	-0.97	0.14
Otto050	-20.00	25.00	25.00	0.00	0.00	0.91	0.91	0.91	0.61	-0.20	-19.58	-7.38	-7.38	-2.30	0.95
Otto075	0.00	33.33	33.33	33.33	33.33	2.56	2.56	2.56	2.26	2.05	-21.38	-8.92	-8.92	-1.87	1.87
Otto100	-39.32	54.55	18.18	9.09	9.09	2.12	2.15	2.04	1.84	1.02	-16.74	-6.90	-3.86	-3.15	-0.17
Otto125	-33.00	20.00	0.00	0.00	0.00	1.11	1.11	1.01	0.91	0.20	-19.07	-7.61	-3.74	-2.97	-0.18
Otto150	0.83	33.33	33.33	33.33	33.33	2.56	2.56	2.56	2.46	2.46	-21.44	-8.16	-8.16	-3.43	-3.43
Otto175	-31.75	60.00	30.00	30.00	20.00	2.52	2.56	2.46	2.26	1.64	-16.40	-6.05	-3.63	-1.95	-0.32
Otto200	-29.17	16.67	16.67	0.00	0.00	2.14	2.15	2.15	1.84	1.02	-19.79	-9.49	-9.49	-4.38	-0.72
Otto225	0.00	33.33	33.33	33.33	0.00	2.56	2.56	2.56	2.36	1.64	-22.26	-9.76	-9.76	-1.77	-2.03
Otto250	-25.25	70.00	40.00	40.00	30.00	2.56	2.56	2.46	2.26	1.64	-17.63	-3.98	-1.47	-0.38	0.29
Otto275	-29.50	40.00	20.00	20.00	20.00	2.56	2.56	2.46	2.26	2.05	-18.08	-7.48	-4.34	-2.65	-1.37
Otto300	-25.00	25.00	25.00	0.00	0.00	2.56	2.56	2.56	2.26	2.26	-20.27	-7.68	-7.68	-3.67	-3.57
Otto325	-54.82	21.43	21.43	14.29	0.00	0.63	2.56	2.56	2.26	1.44	-17.27	-8.46	-8.46	-2.77	-1.78
Otto350	-25.00	40.00	20.00	20.00	20.00	2.56	2.56	2.46	2.46	1.64	-18.40	-8.61	-3.98	-4.24	-0.84
Otto375	0.83	33.33	33.33	33.33	0.00	0.91	0.91	0.91	0.91	0.00	-21.41	-5.92	-5.92	-3.55	-0.43
Otto400	-44.17	33.33	25.00	25.00	16.67	2.46	2.56	2.46	2.26	1.44	-16.26	-7.92	-2.64	-1.50	1.02
Otto425	-25.00	33.33	16.67	0.00	0.00	0.50	0.50	0.40	0.20	0.00	-18.59	-5.75	-0.92	-1.48	-1.01
Otto450	0.83	33.33	33.33	33.33	0.00	2.15	2.15	2.15	2.04	1.02	-21.92	-8.38	-8.38	-2.37	-0.96
Otto475	-42.73	63.64	36.36	27.27	9.09	0.73	2.56	2.46	2.26	1.44	-17.63	-6.14	-3.62	-1.66	-0.43
Otto500	-48.75	12.50	12.50	12.50	12.50	2.56	2.56	2.56	2.56	1.64	-17.75	-9.11	-9.11	-5.40	-0.01
Otto525	0.83	33.33	33.33	33.33	0.00	0.91	0.91	0.91	0.91	0.00	-21.57	-8.25	-8.25	-3.28	-0.10
Jaeschke	-57.14	14.29	14.29	0.00	0.00	0.10	0.10	0.10	0.00	0.00	-16.01	-5.91	-5.91	0.00	0.00
Jackson	-50.00	16.67	16.67	16.67	0.00	0.30	0.30	0.30	0.30	0.00	-16.58	-5.51	-5.51	-5.51	0.00
Mansoor	-25.00	0.00	0.00	0.00	0.00	0.30	0.30	0.30	0.30	-0.50	-21.47	-8.65	-8.65	-4.74	2.46
Mitchell	-32.50	16.67	0.00	0.00	0.00	1.21	1.21	1.11	1.11	0.91	-19.97	-9.88	-6.61	-6.61	-3.41
Roszieg	-33.44	25.00	25.00	0.00	0.00	1.00	1.01	1.01	0.81	0.81	-17.02	-7.69	-7.69	-3.68	-3.68
Heskiaoff	-36.50	20.00	20.00	20.00	20.00	1.94	1.94	1.94	1.73	1.12	-16.81	-8.24	-8.24	-2.19	0.84
Buxey	-30.00	20.00	20.00	10.00	10.00	1.79	1.94	1.94	1.73	1.43	-15.87	-9.19	-9.19	-3.56	-1.60
Sawyer	-28.00	20.00	10.00	10.00	10.00	1.87	1.94	1.83	1.73	1.43	-16.28	-9.09	-5.18	-3.47	-1.82
Lutz1	-38.61	22.22	11.11	11.11	0.00	1.98	2.15	1.94	1.84	1.23	-16.98	-8.47	-4.50	-2.90	-3.14
Gunther	-37.75	30.00	20.00	10.00	0.00	1.17	1.52	1.42	1.32	0.91	-14.81	-7.30	-3.71	-3.46	-1.83
Kilbridge	-35.94	12.50	0.00	0.00	0.00	2.39	2.46	2.36	2.25	1.64	-16.85	-8.85	-5.63	-4.85	-1.67
Hahn	2.50	60.00	20.00	20.00	20.00	3.19	3.20	2.99	2.68	2.17	-18.94	-6.53	-4.00	-2.34	-0.84
Warnecke *	-61.52	26.09	17.39	*17.39	4.35	-18.39	3.73	3.63	*3.63	3.01	-27.29	-8.74	-4.76	*-4.76	-2.34
Tonge	-41.25	18.75	12.50	6.25	6.25	-6.81	2.46	2.36	2.25	1.64	-19.68	-8.64	-5.00	-4.17	-1.56
Wee-mag	-84.92	3.33	1.67	1.67	0.00	-44.13	1.11	1.01	1.01	0.51	-40.97	-9.15	-4.73	-4.73	-1.56
Arcus1	-6.54	30.77	15.38	15.38	7.69	1.10	3.31	3.10	3.10	2.38	-16.46	-9.63	-5.92	-5.92	-2.31
Lutz2	-63.09	29.41	17.65	8.82	0.00	-24.35	0.81	0.81	0.71	0.00	-27.58	-8.85	-6.75	-3.52	0.00
Lutz3	-36.53	22.22	16.67	11.11	5.56	-6.91	2.15	2.04	1.94	1.43	-19.28	-8.45	-4.84	-3.97	-1.81
Mukherje	-30.13	26.32	15.79	10.53	10.53	-5.95	2.46	2.36	2.25	1.64	-18.55	-8.12	-4.68	-3.64	-0.97
Arcus2	-47.08	*22.22	11.11	11.11	0.00	-19.08	*1.94	1.83	1.83	0.00	-31.25	*-10.39	-6.52	-6.52	0.00
Barthol2	-60.83	*25.64	*15.38	10.26	5.13	-25.86	*2.25	*2.15	2.04	1.33	-28.86	*-8.22	*-4.98	-3.63	-1.30
Barthold	-16.88	16.67	8.33	8.33	0.00	1.40	2.56	2.46	2.26	1.64	-16.63	-8.93	-5.15	-3.74	-2.01
EP1	-34.64	29.04	19.35	14.51	7.27	0.15	1.98	1.91	1.77	1.18	-39.09	-26.14	-20.68	-14.52	-6.24
EP2	-28.59	29.04	19.35	14.51	7.27	-3.21	1.98	1.91	1.77	1.18	-8.67	0.48	0.68	1.15	1.05
EP3	-34.88	29.04	19.35	14.51	7.27	-2.90	1.98	1.91	1.77	1.18	-14.38	-1.56	-0.56	0.64	0.92
EP4	-28.50	29.04	19.35	14.51	7.27	-2.57	1.98	1.91	1.77	1.18	-17.21	-4.53	-2.66	-0.64	0.40
Overall	-31.65	29.04	19.35	14.51	7.27	-2.13	1.98	1.91	1.77	1.18	-19.84	-7.94	-5.80	-3.34	-0.97

Table 8: Comparison of the reference RMS solution and the dedicated lines for the three objectives ( $z_m$ ,  $z_{sl}$ ,  $z_{ec}$ ), using DL50% as reference.

Some slightly degenerating cases can occur. A dedicated line associated with a quantile is obtained by solving an SALBP-1 whose target takt time is obtained by dividing the time horizon by the target demand and rounding down. In doing so, in case of high values for target demand, rounding can lead to similar target takt time values for different quantiles, and the corresponding dedicated lines may be the same. Moreover, SALBP-1 aims at minimizing the number of stations, and the actual takt time of the optimal solution may be strictly less than the upper bound, resulting in a dedicated line performing better than expected.

In Tables 8 and 9, starred results refer to dedicated lines which are not optimal SALBP-1 solutions, or whose optimality has not been proved within the imposed time limit of 3 hours; in the case of instance Warnecke, this occurs for the solution of the reference line DL50%.

Table 8 shows that when computing the average over the 43 instances, the relative performance of

	$E[\%H]$					
	RMS*	DL	DL	DL	DL	DL
		100%	95%	90%	75%	50%
Otto025	56.28	78.82	86.67	88.89	93.00	94.98
Otto050	42.52	77.62	77.62	89.29	93.30	92.32
Otto075	35.30	75.53	75.53	88.60	90.44	95.47
Otto100	53.73	77.62	87.04	89.19	92.71	95.25
Otto125	47.86	78.02	87.34	88.36	91.82	92.41
Otto150	35.21	78.52	78.52	87.04	87.04	96.05
Otto175	53.18	79.02	87.14	89.29	92.12	96.05
Otto200	46.12	77.22	77.22	88.60	93.20	94.95
Otto225	39.40	74.64	74.64	88.36	92.42	95.85
Otto250	49.67	79.22	87.34	89.19	91.73	95.66
Otto275	51.33	78.92	86.84	89.09	90.54	95.85
Otto300	44.10	78.42	78.42	89.09	89.19	95.27
Otto325	64.56	79.02	79.02	89.29	93.00	95.85
Otto350	50.29	76.93	87.34	87.24	91.92	96.05
Otto375	33.04	78.92	78.92	83.08	91.92	92.32
Otto400	57.33	77.42	87.44	88.70	93.20	95.37
Otto425	45.75	79.22	87.34	89.19	90.24	90.84
Otto450	36.62	77.22	77.22	87.34	93.00	95.05
Otto475	58.06	77.13	85.98	89.29	93.20	95.56
Otto500	59.58	77.62	77.62	84.35	92.02	95.37
Otto525	33.97	74.04	74.04	83.38	92.22	92.32
Jaeschke	55.93	75.00	75.00	87.38	87.38	87.38
Jackson	55.55	80.00	80.00	80.00	89.69	89.69
Mansoor	41.22	77.39	77.39	83.82	92.64	89.94
Mitchell	47.43	76.18	85.66	85.66	90.11	93.90
Roszieg	51.97	77.78	77.78	88.68	88.68	93.34
Heskiaoff	54.83	79.63	79.63	88.68	92.62	95.16
Buxey	54.27	77.78	77.78	88.68	91.15	95.19
Sawyer	52.64	77.78	86.04	88.68	91.15	95.19
Lutz1	54.66	79.96	88.23	90.08	92.19	95.58
Gunther	58.48	79.63	86.94	88.68	91.90	94.63
Kilbridge	55.30	79.70	87.18	88.37	92.72	96.03
Hahn	43.94	76.74	87.26	89.44	91.67	96.18
Warnecke *	89.35	78.81	87.15	* 87.15	92.14	* 97.30
Tonge	75.69	79.86	87.02	89.02	92.64	96.04
Wee-mag	100.00	79.47	87.06	87.06	91.66	93.63
Arcus1	60.58	77.52	86.26	86.26	91.85	96.20
Lutz2	91.65	75.00	81.25	87.38	92.84	92.84
Lutz3	73.74	79.32	86.45	88.40	92.00	95.55
Mukherje	73.17	79.86	87.02	89.02	92.64	96.04
Arcus2	82.67	*83.24	90.81	90.81	97.80	97.80
Barthol2	93.29	*79.42	* 86.47	89.00	92.77	95.72
Barthold	59.50	79.71	87.38	89.36	92.78	96.06
EP1	58.98	78.16	83.14	87.87	91.84	94.61
EP2	54.90	78.16	83.14	87.87	91.84	94.61
EP3	58.36	78.16	83.14	87.87	91.84	94.61
EP4	52.86	78.16	83.14	87.87	91.84	94.61
Overall	56.27	78.16	83.14	87.87	91.84	94.61

Table 9: Time usage comparison of the reference RMS solution and the 5 dedicated lines .

dedicated lines in terms of number of stations and expected value of service level is predictably not impacted by the cost profile . Table 9 allows the same conclusion to be drawn as to time usage.

The reconfigurable line RMS\* has in general (-2.13% on average over all cost profiles and instances) a lower expected value of service level w.r.t. the line DL50%. The other dedicated lines have an expected service level that is slightly better than DL50%. Again, this is not completely surprising, since the dedicated lines are designed based on a service level guarantee, while RMS\* results from best-compromise solutions over the three objective functions. Moreover, since the highest possible demand value of the distribution is 25% greater than the median, it is easy to see that by construction even DL50% cannot have an expected service level lower than 90%. This can further explain the seemingly unfavorable trend of RMS\* as to service level. However, if we look attentively at Table 8, we notice that for most of the instances (35 out of 43, considering the average values over cost profiles) RMS\* actually also outperforms DL50% in terms of expected value of service level. In the remaining cases, the average design nearest to the ideal point performs worse in terms of  $z_{sl}$ , but the method has seemingly preferred to put  $z_m$  and  $z_{ec}$ , their improvement being much more dramatic than in other cases.

RMS\* behaves much better than DL50% in terms of expected per-produced-unit energy cost (up to nearly 40% better for EP1, averaged over all instances), and also in terms of number of workstations (up to nearly 35% better, again for EP1), which RMS\* seems to favor as opposed to the service level. For what concerns this two aspects, it must be pointed out that RMS\* is also significantly better than the other dedicated lines: while on the one hand this can be expected for the number of stations (which can only worsen in dedicated lines DL75% to DL100% w.r.t. DL50%), on the other hand this once again proves the effectiveness of reconfigurable systems, which –thanks to their flexibility– can afford consistent energy cost savings, all with a consistent reduction in the number of workstations.

More generally, we can conclude that RMS\* can lose on average 2.13 to 4.03% (from DL50% to DL100%) in terms of expected service level compared to dedicated lines, but can save on average 12.93 to 19.84% (from DL100% to DL50%) in the expected per-produced-unit energy cost, while using, on average, 31.65 to 47.03% (from DL50% to DL100%) fewer workstations.

A joint analysis of Tables 8 and 9 highlights the fact that RMS\* can outperform dedicated lines so

significantly because of the dramatic reduction in the time spent producing (on average -21.89 to -38.84% from DL100% to DL50%), as already observed in the discussion of Figure 4. Reduced production time, made possible by the flexibility provided by its set of configurations, is then a lever not only to increase the service level, but also to avoid production as much as possible during the costlier periods of the TOU profile, and as a consequence to reduce the per-produced-unit energy cost.

## 6. Managerial insights and practical implications

Several insights can be derived from the above results.

First, and probably most relevant given the motivations behind this study, the presented results show that RMS can help achieve significant savings in terms of per-produced-unit energy cost. Compared to dedicated lines, the average gain can range from almost 13% to nearly 20%. Obviously, the potential gain can vary according to the considered energy cost profile but RMS always lead to more cost-efficient solutions.

Another interesting aspect highlighted here is the variety of planning solutions made possible by RMS to reduce production-related energy costs. Indeed, when the energy cost is considered in a production system, the two most common levers are: overdimensioning the system (or reconfiguration to increase the throughput in the case of RMS); or deliberately choosing not to meet the demand entirely, in order to avoid producing during the most expensive TOU periods. The results of this study show that a wide set of trade-off solutions exist between these two options and can be found by the proposed approach. The more complex the case (instance) under study, the greater the possibility for the decision maker to fine-tune the main features –service level, per-unit energy cost, number of workstations– and choose the most fitting solution. Incidentally, the matheuristic is capable of converging and finding a fitting solution set in relatively short computation time, fully compatible with the needs of designing a production system. On the other hand, the problem, which is by nature multi-objective, contemplates very few nondominated solutions, and the decision maker can hopefully achieve consistent energy cost savings and the service level at the same time.

As noted in the conclusion of Section 5.5, the general behavior of the RMS solutions yielded by the proposed matheuristic is to reduce energy-related economic costs by cleverly using its configurations so as to considerably reduce the time spent producing and to avoid costlier TOU periods. In doing so, the RMS solutions rarely seem to need more than 2-3 configurations to face a generic demand scenario over a given time horizon. This seems to be a reliable indicator of the small number of reconfigurations required over the same timespan. Moreover, computational experiments suggest that although a balancing solution has a wide set of possible configurations, the number of them actually used to meet *all* the considered demand scenarios is also, in general, no larger than 2-3, potentially indicating that the same restricted subset is always used regardless of the demand.

Some final conclusions can also be drawn concerning dedicated lines. The results have shown that if such a line is designed based on worst-case demand, the gain in terms of service level compared to one based on the median demand is only about 2%, at a cost of almost 30% more workstations. However, opting for the former solution is still of interest if energy cost is the main driver, as the savings in that respect can rise up to 26% in the event of an unbalanced TOU energy profile.

All these findings highlight the need for Product Lifecycle Management (PLM) systems to integrate the capabilities of RMS to deal efficiently with energy concerns.

## 7. Conclusion and perspectives

In this article, we addressed a scenario-based bi-level optimization problem consisting of balancing the operations and planning the configurations of Reconfigurable Manufacturing Systems (RMS), by optimizing three objectives: the number of stations, the expected energy cost per produced unit and the expected service level in the face of uncertain demands. This new problem corresponds to a challenge companies are more and more forced to cope with as the prices of energy and the volatility of energy sources increase, and highlights an innovative way to use RMS. We developed a multi-objective matheuristic, composed of a Multi-Objective Simulated Annealing (MOSA) for the balancing level and a Linear Program (LP) for the planning level.



Experiments were conducted to analyze the performances of the method and the characteristics of the solutions. The results obtained prove the applicability of our approach with very low computational times, although additional experiments on industrial cases would help to better assess its performances.

We also compared the RMS solutions generated with those corresponding to dedicated lines in order to show the potential gain involved in using RMS. Though the solutions used for RMS were not necessarily optimal and those for dedicated lines were, significant savings were observed. Managerial insights into the use of RMS to optimize the three objectives can also be provided. In particular, they highlight the large range of possible tradeoffs between overdimensioning the system and choosing to produce less when aiming to reduce the energy cost of production. The results suggest that when a strategic decision is made to use RMS, substantial gains can be achieved in terms of energy costs. Moreover, even already implemented RMS (whether for scalability purposes in a mono-product setting, or for convertibility reasons in view of a future multi-product evolution) can also be used to achieve consistent savings in energy-related production costs.

Since the approach proposed in this paper is a metaheuristic, some improvements seem possible. For example, the experiments have shown that only a small subset of the configurations generated are used in the production plan: it would thus be interesting to filter configurations in the earliest steps of the algorithm to better evaluate the different balancing solutions. In addition, the development of an integrated model could yield an exact method for the studied bi-level problem. Such a model would certainly involve a huge number of variables, which would prevent large instances from being solved, but an approach based on Column Generation could work, allowing a better assessment of the gains achievable with RMS by comparing optimal solutions for both types of systems.

Further works could investigate certain variations of the problem explored here. One future direction could be the integration of an upper bound on the number of reconfigurations allowed during the planning horizon. Also, we assumed that the decision maker at the planning level would prioritize the service level over the energy cost, but it would be interesting to study cases where different trade-offs are sought. Furthermore, situations where the uncertainty surrounding the demand remains at the planning level would require the use of a reactive approach, which could take advantage of the low number of configurations to consider. Finally, it would certainly be worth studying an extension of the proposed problem and approach in a multi-product setting, be it mixed-model or multi-model, and assessing whether RMS can also be beneficial in terms of energy-related cost savings also when reconfigurations serve not only for scalability, but also for convertibility purposes. In this case, flexible manufacturing systems would be used as comparison basis (taking the role of dedicated lines in this study).

## Acknowledgements

This work is supported by the project ANR-21-CE10-0019 RECONFIDURABLE.

## References

- [1] Koren Y, Gu X, Guo W. Choosing the system configuration for high-volume manufacturing. *International Journal of Production Research* 2018;56(1-2):476–90.
- [2] Koren Y, Gu X, Badurdeen F, Jawahir IS. Sustainable living factories for next generation manufacturing. *Procedia Manufacturing* 2018;21:26–36.
- [3] U.S. Energy Information Administration . International energy outlook 2019 with projections to 2050. Tech. Rep.; U.S. Department of Energy; 2019. URL: <https://www.eia.gov/outlooks/ieo/pdf/ieo2019.pdf>.
- [4] International Energy Agency . World energy outlook 2021. Tech. Rep.; IEA Publications; 2021. URL: <https://www.iea.org/reports/world-energy-outlook-2021>.
- [5] Giret A, Trentesaux D, Prabhu V. Sustainability in manufacturing operations scheduling: A state of the art review. *Journal of Manufacturing Systems* 2015;37:126–40.
- [6] Andiappan V. Optimization of smart energy systems based on response time and energy storage losses. *Energy* 2022;258. doi:10.1016/j.energy.2022.124811.

- [7] Battaia O, Benyoucef L, Delorme X, Dolgui A, Thevenin S. Sustainable and energy efficient reconfigurable manufacturing systems. In: *Reconfigurable Manufacturing Systems: From Design to Implementation*. Springer; 2020, p. 179–91.
- [8] Zhang R, Chiong R. Solving the energy-efficient job shop scheduling problem: a multi-objective genetic algorithm with enhanced local search for minimizing the total weighted tardiness and total energy consumption. *Journal of Cleaner Production* 2016;112(4):3361–75.
- [9] Bruzzone AA, Anghinolfi D, Paolucci M, Tonelli F. Energy-aware scheduling for improving manufacturing process sustainability: A mathematical model for flexible flow shops. *CIRP annals* 2012;61(1):459–62.
- [10] Luo H, Du B, Huang GQ, Chen H, Li X. Hybrid flow shop scheduling considering machine electricity consumption cost. *International Journal of Production Economics* 2013;146(2):423–39.
- [11] Koren Y, Heisel U, Jovane F, Moriwaki T, Pritschow G, Ulsoy G, et al. Reconfigurable manufacturing systems. *CIRP Annals* 1999;48(2):527–40.
- [12] Koren Y, Gu X, Guo W. Reconfigurable manufacturing systems: Principles, design, and future trends. *Frontiers of Mechanical Engineering* 2018;13(2):121–36.
- [13] Koren Y, Wang W, Gu X. Value creation through design for scalability of reconfigurable manufacturing systems. *International Journal of Production Research* 2017;55(5):1227–42.
- [14] ElMaraghy H, Schuh G, ElMaraghy W, Piller F, Schönsleben P, Tseng M, et al. Product variety management. *CIRP Annals* 2013;62(2):629–52.
- [15] Koren Y. The emergence of reconfigurable manufacturing systems (rmss). In: *Reconfigurable Manufacturing Systems: From Design to Implementation*. Springer; 2020, p. 1–9.
- [16] Battaia O, Dolgui A, Guschinsky N, Levin G. Integrated configurable equipment selection and line balancing for mass production with serial–parallel machining systems. *Engineering Optimization* 2014;46(10):1369–88.
- [17] Manzini M, Unglert J, Gyulai D, Colledani M, Jauregui-Becker JM, Monostori L, et al. An integrated framework for design, management and operation of reconfigurable assembly systems. *Omega* 2018;78:69–84.
- [18] Yelles-Chaouche AR, Gurevsky E, Brahim N, Dolgui A. Reconfigurable manufacturing systems from an optimisation perspective: a focused review of literature. *International Journal of Production Research* 2021;59(21):6400–18.
- [19] Liu X, Yang X, Lei M. Optimisation of mixed-model assembly line balancing problem under uncertain demand. *Journal of Manufacturing Systems* 2021;59:214–27.
- [20] Delorme X, Cerqueus A, Gianessi P, Lamy D. Design and planning of configurations in RMS to minimize the energy cost facing uncertain demand. In: *International Working Seminar on Production Economics 2022 (IWSPE)*. Innsbruck, Austria; 2022,.
- [21] Battaia O, Dolgui A. A taxonomy of line balancing problems and their solution approaches. *International Journal of Production Economics* 2013;142(2):259–77.
- [22] Makssoud F, Battaia O, Dolgui A. Multi-objective approach and model for transfer line reconfigurations. In: *Reconfigurable Manufacturing Systems: From Design to Implementation*. Springer. ISBN 978-3-030-28782-5; 2020, p. 193–208. doi:10.1007/978-3-030-28782-5\_10.
- [23] Özcan Ug, Çerçioğlu H, Gökçen H, Toklu B. Balancing and sequencing of parallel mixed-model assembly lines. *International Journal of Production Research* 2010;48(17):5089–113. doi:10.1080/00207540903055735.
- [24] Priore P, de la Fuente D, Puente J, Parreño J. A comparison of machine-learning algorithms for dynamic scheduling of flexible manufacturing systems. *Engineering Applications of Artificial Intelligence* 2006;19(3):247–55. doi:10.1016/j.engappai.2005.09.009.

- [25] Toader FA. Production scheduling in flexible manufacturing systems: A state of the art survey. *Journal of Electrical Engineering, Electronics, Control and Computer Science* 2017;3(1):1–6.
- [26] Yadav A, Jayswal S. Modelling of flexible manufacturing system: a review. *International Journal of Production Research* 2018;56(7):2464–87. doi:10.1080/00207543.2017.1387302.
- [27] Dou J, Dai X, Meng Z. Graph theory-based approach to optimize single-product flow-line configurations of RMS. *The International Journal of Advanced Manufacturing Technology* 2009;41(9):916–31.
- [28] Moghaddam SK, Houshmand M, Fatahi Valilai O. Configuration design in scalable reconfigurable manufacturing systems (RMS); a case of single-product flow line (SPFL). *International Journal of Production Research* 2018;56(11):3932–54.
- [29] Lahrichi Y, Deroussi L, Grangeon N, Norre S. A Balance-First Sequence-Last Algorithm to design RMS A Matheuristic with performance guaranty to balance Reconfigurable Manufacturing Systems. *Journal of Heuristics* 2021;27:107–32. doi:10.1007/s10732-021-09473-1.
- [30] Koren Y, Shpitalni M. Design of reconfigurable manufacturing systems. *Journal of manufacturing systems* 2010;29(4):130–41.
- [31] Cui F, Geng N, Jiang Z, Zhou X. A stochastic programming approach for configuration optimization of reconfigurable manufacturing systems. In: *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE; 2020, p. 1152–6.
- [32] Deif AM, ElMaraghy HA. Assessing capacity scalability policies in RMS using system dynamics. *International Journal of Flexible Manufacturing Systems* 2007;19(3):128–50.
- [33] Cerqueus A, Delorme X. Evaluating the scalability of reconfigurable manufacturing systems at the design phase. *International Journal of Production Research* 2023;doi:10.1080/00207543.2022.2164374.
- [34] Wang W, Koren Y. Scalability planning for reconfigurable manufacturing systems. *Journal of Manufacturing Systems* 2012;31(2):83–91.
- [35] Hees A, Schutte CS, Reinhart G. A production planning system to continuously integrate the characteristics of reconfigurable manufacturing systems. *Production Engineering* 2017;11(4):511–21.
- [36] Moghaddam SK, Houshmand M, Saitou K, Fatahi Valilai O. Configuration design of scalable reconfigurable manufacturing systems for part family. *International Journal of Production Research* 2020;58(10):2974–96.
- [37] Huang A, Badurdeen F, Jawahir IS. Towards developing sustainable reconfigurable manufacturing systems. *Procedia Manufacturing* 2018;17:1136–43.
- [38] Zhang J, Khalgui M, Boussahel WM, Frey G, Hon C, Wu N, et al. Modeling and verification of reconfigurable and energy-efficient manufacturing systems. *Discrete Dynamics in Nature and Society* 2015;doi:10.1155/2015/813476.
- [39] Delorme X, Gianessi P. Designing reconfigurable manufacturing systems to minimize power peak. *IFAC-PapersOnLine* 2022;55(10):1296–301.
- [40] Choi YC, Xirouchakis P. A holistic production planning approach in a reconfigurable manufacturing system with energy consumption and environmental effects. *International Journal of Computer Integrated Manufacturing* 2015;28(4):379–94.
- [41] Shrouf F, Ordieres-Meré J, García-Sánchez A, Ortega-Mier M. Optimizing the production scheduling of a single machine to minimize total energy consumption costs. *Journal of Cleaner Production* 2014;67:197–207.
- [42] Che A, Zhang S, Wu X. Energy-conscious unrelated parallel machine scheduling under time-of-use electricity tariffs. *Journal of Cleaner Production* 2017;156:688–97.

- [43] Masmoudi O, Delorme X, Gianessi P. Job-shop scheduling problem with energy consideration. *International Journal of Production Economics* 2019;216:12–22.
- [44] Lamy D, Delorme X, Gianessi P. Line balancing and sequencing for peak power minimization. *IFAC-PapersOnLine* 2020;53(2):10411–6.
- [45] Oduguwa V, Roy R. Bi-level optimisation using genetic algorithm. In: *Proceedings 2002 IEEE International Conference on Artificial Intelligence Systems (ICAIS 2002)*. IEEE; 2002, p. 322–7.
- [46] Khezri A, Benderbal HH, Benyoucef L, Dolgui A. Diagnosis on energy and sustainability of reconfigurable manufacturing system (rms) design: A bi-level decomposition approach. In: *2020 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. IEEE; 2020, p. 141–5.
- [47] Cerqueus A, Gianessi P, Lamy D, Delorme X. Balancing and configuration planning of RMS to minimize energy cost. In: Lalic B, Majstorovic V, Marjanovic U, von Cieminski G, Romero D, editors. *Advances in Production Management Systems. Towards Smart and Digital Manufacturing*; vol. 592 of *IFIP Advances in Information and Communication Technology*. Springer International Publishing; 2020, p. 518–26.
- [48] Gianessi P, Cerqueus A, Lamy D, Delorme X. Using reconfigurable manufacturing systems to minimize energy cost: a two-phase algorithm. *IFAC-PapersOnLine* 2021;54(1):379–84.
- [49] Freiheit T, Shpitalni M, Hu SJ. Productivity of paced parallel-serial manufacturing lines with and without crossover. *Journal of Manufacturing Science and Engineering* 2004;126(2):361–7.
- [50] Ulungu EL, Teghem J, Fortemps P, Tuyttens D. MOSA method: a tool for solving multiobjective combinatorial optimization problems. *Journal of Multicriteria Decision Analysis* 1999;8(4):221.
- [51] Braithwait S, Hansen D, O’Sheasy M. Retail electricity pricing and rate design in evolving markets. *Tech. Rep.*; Edison Electric Institute; 2007.
- [52] Abikarram JB, McConky K, Proano R. Energy cost minimization for unrelated parallel machine scheduling under real time and demand charge pricing. *Journal of Cleaner Production* 2019;208:232–42.
- [53] Scholl A. Data of assembly line balancing problems. *Tech. Rep.*; Darmstadt Technical University, Institute for Business Studies (BWL); 1993. URL: <https://assembly-line-balancing.de/salbp/benchmark-data-sets-1993>.
- [54] Otto A, Otto C, Scholl A. Systematic data generation and test design for solution algorithms on the example of SALBPGen for assembly line balancing. *European Journal of Operational Research* 2013;228(1):33–45.