Computers & Industrial Engineering 87 (2015) 561–568

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

The scheduling of maintenance. A resource-constraints mixed integer linear programming model ☆

Riccardo Manzini *, Riccardo Accorsi, Teresa Cennerazzo, Emilio Ferrari, Fausto Maranesi

Department of Industrial Engineering, Alma Mater Studiorum – Bologna University, Italy

A R T I C L E I N F O

Article history:
Received 18 December 2014
Received in revised form 9 June 2015
Accepted 10 June 2015
Available online 17 June 2015

Keywords:
Scheduling
Reliability
Maintenance action scheduling
Preventive maintenance
Corrective maintenance
Mixed integer linear programming (MILP)

A B S T R A C T

The scheduling of preventive maintenance is crucial in reliability and maintenance engineering. Hundreds of parts compose complex machines that require replacement and/or repairing. Maintenance involves the machine vendor (1), the machine user (2) and the service maintenance provider (3). The vendor and the maintenance service provider have to guarantee a high level of availability and productivity of the machines and maintain their down-time at a minimum even though they are installed worldwide and usually far from the vendor's headquarters and/or the locations of the provider's regional service offices. Moreover, many companies have great profits from maintenance and spare parts management.

This study aims to illustrate an original mixed integer linear programming (MILP) model for the cost-based, reliability-based and resource-constraints scheduling of preventive maintenance actions. The model minimizes the total cost function made of spare parts contributions, the cost of the execution of the preventive actions and the cost of the additional repair activity in case of unplanned failure. The cost of the personnel of the producer and/or the maintenance service provider is also included. Finally, the paper presents a case study in a what-if environment demonstrating the effectiveness and the novelty of this study in real and complex applications.

1. Introduction

Literature classifies maintenance planning and scheduling into two major categories: the scheduled maintenance (1) and the unscheduled maintenance (2). The second deals with emergency breakdowns. The first includes preventive and routine maintenance (1.1), and the scheduled overhauls and corrective maintenance (1.2). The unscheduled maintenance is stochastic in nature. According to Duffuaa and Al-Sultan (1999) "this stochastic nature makes maintenance scheduling a challenging problem".

Many companies produce and distribute worldwide complex production systems and machines. They also offer several maintenance services that include spare parts management, preventive maintenance actions, corrective maintenance actions, warranty management, and training of personnel. Maintenance service is a strategic activity to have a high level of productivity, quality, safety, and reliability of production systems. Furthermore, this can be a very expensive and labor-intensive service but also an opportunity for economic returns by post-sale services. The cost of maintenance can be also significantly affected by logistics decisions, including the number and location of service providers and regional offices, the inventory management of spare parts, and the organization of maintenance crews.

This paper illustrates an original cost-based, reliability-based and capacity-constraints optimization model for the scheduling of the maintenance and repair tasks within a maintenance plan (i.e., task plan).

The maintenance tasks refer to the set of activities necessary to replace a component or a group of components subjected to wear and tear within a generic plant or machine. The group of maintenance tasks including all the repairing and/or replacing activities that a generic machine or a plant require over its own life-cycle is named task plan. Each task to be scheduled usually involve spare parts, personnel (e.g., local personnel or service providers' operators), resources and equipment. The frequency of each task is generally determined by the failure rates (i.e., the curve of failure probability to the machine up time) of the most critical component of the task. The general rule complied by the maintenance service provider in presence of complex components is assuming a constant failure rate (i.e., Assumption 1) corresponding to the average

☆ This manuscript was processed by Area Editor Maged M. Dessouky.

* Corresponding author at: viale Risorgimento 2, 40136 Bologna, Italy.
E-mail address: riccardo.manzini@unibo.it (R. Manzini).

http://dx.doi.org/10.1016/j.cie.2015.06.006
0360-8352/

value suggested by the machine vendor. This assumption is critical in the presence of mechanical and mechatronic components that are mostly diffused in the modern automatic machines. However, Assumption 1 is often necessary due to the large amount of parts and components involved simultaneously and physically connected. Another assumption that frequently follows the constant failure rate is the constant frequency to execute preventive maintenance tasks (i.e., Assumption 2).

Furthermore, the provider commonly executes the preventive task on a component after a time equal to the mean time to failure (MTTF) of the task/component from the previous action and/or replacement (Assumption 3).

Unfortunately, when applied to real instances, these assumptions are not consistent, especially in presence of parts subject to "aging", e.g., "early wear out" components or "old age and rapid wear out" components (Manzini, Regattieri, Pham, & Ferrari, 2010). Furthermore, the parts and components of a production system, e.g., a packaging machine, are not "as good as new" items after repairing or a preventive action, even in case of the part replacement.

To find more concrete and realistic solutions and go beyond to the illustrated assumptions, this paper presents an original mixed integer linear programming (MILP) model for the determination of the maintenance schedule that minimizes the total cost associated to the task plan. These costs include the preventive maintenance contributions, the corrective contributions (the so-called unplanned costs), the spare parts management, and the labor accounted by the maintenance operators.

The task plan scheduling is the result of the assignment and sequencing of different preventive maintenance tasks to a set of available service orders. This set is usually known in advance and results from a deal between the supplier of maintenance service, i.e., the previously defined "service provider", and the client which requires for the maintenance of its plant. The generic service order corresponds to a time bucket located on a specific calendar date. This is the reason we adopt the terms time bucket to indicate a service order of a finite capacity.

The client purchases a calendar of preventive maintenance time buckets, and the service provider has to assign maintenance tasks to these buckets, controlling the availability of the system and reducing costs to realize a profitable service. In other words, the aim of the provider is to minimize the total cost of maintenance while guaranteeing a standard level of availability (i.e., up time) of the production system.

The remainder of this paper is organized as follows. Section 2 presents a literature review on the scheduling of the preventive maintenance. Section 3 illustrates the proposed maintenance planning model. Section 4 presents a significant case study which inspired the development of the proposed model. A sensitivity analysis is conducted to demonstrate the effectiveness of the proposed planning model. Finally, Section 5 discusses the conclusions and further research.

2. Literature review

The literature presents many contributions to preventive maintenance and scheduling issues for production systems with a special focus on operations. In particular, management science and operational research frequently discuss scheduling and optimization problems, but few studies deal with reliability and maintenance engineering (Manzini et al., 2010; Regattieri, Manzini, & Battini, 2010).

Sherwin (2000) presents a review and a discussion of the main issues in maintenance management. He also attributes significant and strategic importance to data collection to conduct effective planning and scheduling of maintenance tasks.

Many studies deal with maintenance planning applied to production and operations, e.g., models and methods to schedule preventive maintenance activities on manufacturing systems subject to failure, i.e., corrective maintenance (Hadidi, Al-Turki, & Rahim, 2012; Xiang, Cassady, Jin, & Zhang, 2014). In particular, they formulate integrated planning models to simultaneously face production and maintenance planning (Cassady & Kutanoglu, 2005; Kuo & Chang, 2007). These contributions are not based on the reliability of parts and components involved and are not suitable to strategically design a task plan tailored to a selected production system subject to failure. They do not involve the management of spare parts and the assignment of tasks in agreement with finite capacity constraints.

Duffuaa and Al-Sultan (1999) present one of the first mathematical formulation of the stochastic programming for scheduling maintenance personnel. It incorporates deterministic and stochastic contributions. Heuristic algorithms to solve the maintenance scheduling problem are proposed by Raza and Al-Turki (2007). This adoption of heuristic and meta-heuristic approaches is supported by a demonstration of the NP-hard problem complexity.

Several contributions present interval time models, i.e., reliability based static state models for the determination of the time to replace components without any discussion on capacity and time constraints, which are very important in real applications (Hui, Zheng, Liu, Zhao, & Sun, 2013). Simple and basic models are collected and illustrated by Jardine and Tsang (2006). More complex and recent contributions based on MILP are illustrated by Perez Canto (2011) and Bell and Percy (2012).

Kim and Yoo (2012) discuss the planning of maintenance actions combined with manpower by the determination of the workforce size as a relevant issue in the presence of labor-intensive actions and high labor costs.

Alardhi and Labib (2008) present a preventive maintenance scheduling model based on mixed integer programming, which is the modeling approach adopted by the authors of this paper. They include crew constraints, maintenance window constraints and time-limitation constraints, but they do not include reliability based functions.

Tam, Chan, and Price (2006) present three integer linear programming models for maintenance interval determination, minimizing cost and maximizing system availability. They adopt a Weibull distribution for failure rates, but they do not consider time capacity constraints for the execution of a task. Personnel assignment and costs are not included. Finally, spare parts contributions are not modeled.

Moghaddam and Usher (2011) present two non-linear models. The first minimizes the global cost; the other maximizes the system reliability. They adopt increasing failure rates, but they do not consider the time capacity constraints and the time duration of tasks.

Ebrahimipour, Najjarbashi, and Sheikhalishahi (2013) present non-linear models for parallel machines focusing on the difference between maintenance (not as good as new) and replacement (as good as new) activities. They use a Weibull distribution of failure rates, and tasks take a uniformly distributed amount of time. These models are not suitable to scheduling multiple tasks for a complex machine in agreement with the time capacity constraints.

Different modeling approaches to preventive maintenance scheduling are illustrated by Zhang and Nakamura (2005) and Xu, Xueshan, Wang, and Sun (2012), the first illustrating a simulation model and the latter heuristic algorithms, which do not support the decision making techniques adopted by the authors of this paper.

Other recent contributions on maintenance scheduling are proposed by Levi, Magnanti, Muckstadt, Segev, and Zarybnisky (2014), Bajestani and Banjevic (2014), Tantardini, Portioli-Staudacher, and Macchi (2014), Tarakci, Ponnaiyan, and Kulkarni (2014), and Gustavsson, Patriksson, Strömberg, Wojciechowski, and Önnheim (2014). The latter propose an ILP model to schedule the preventive maintenance of the plant components over finite discretized time batch, given a common set-up cost and component costs dependent on the lengths of the maintenance intervals.

Literature does not yet present effective models for cost-based, reliability-based and time resource constraints maintenance planning and scheduling.

The model proposed in this paper addresses the scheduling problem, including both costs (labor cost, spare parts cost, and failure cost) and reliability issues, i.e., preventive and corrective tasks according to finite capacity constraints. The model meets complex instances without any limitations due to the number of tasks and the unit time periods. The sensitivity analysis conducted in this paper demonstrates the effectiveness of the proposed model in complex instances, renouncing the optimum and accepting near-optimal feasible solutions as necessary in real and complex applications.

The original contribution of the proposed model is due to both the mathematical formulation and its applicability to large and real instances as demonstrated by the case study illustrated below.

3. Maintenance planning model

This model deals with the scheduling of maintenance actions in a planning period of time made of pre-defined set of time buckets. The previously defined client purchases a calendar of preventive maintenance time buckets and the provider schedules the maintenance action in according with this calendar.

The problem object of this study is the planning and scheduling of maintenance tasks in the available time buckets corresponding to a set of planned stop periods and preventive maintenance actions. The aim is to minimize the total expected and probabilistic cost made of preventive maintenance, corrective maintenance, spare parts and personnel workload. The cost is "probabilistic" because it quantifies the so-called additional failure cost due to unplanned failures and repair actions.

Along the machine down-time, the maintenance service operators are involved in repair actions, which are the tasks of the scheduling problem. The proposed model assigns the available bucket times to alternative types of operators, e.g., user's operators or service provider's operators, in agreement with the capacity constraints. In particular, the capacity of the generic bucket is finite in time, and the generic maintenance action has a deterministic duration partially consuming this capacity.

The basic assumption of the proposed model is that the generic task can be executed after a variable time from the last execution and/or the starting time of the machine. These values are opportunistic in agreement with the cost minimization and capacity constraints. The scheduler has to decide which task and when to conduct it, in agreement with the time and capacity constraints.

This is a resource-constraints scheduling problem because of the finite capacity constraints assignment problem (Pinedo, 2005). It is similar to a bin-packing problem with different size bins, the so-called variable size bin-packing problem (Bang-Jensen & Larsen, 2012). The generic bin corresponds to the service order, and the related bucket time is subject to finite capacity on the number and typology of operators involved.

The proposed model is mixed-integer and linear, even though the failure probability function of the generic item is introduced. To have a linear model and a realistic unplanned cost function, we assumed the trend of the failure probability function as a path-wise linear function illustrated in Fig. 1 and associated to a generic task i. This trend assumes the as-good-as-new assumption by the execution of the preventive maintenance action. Given a generic task i, the failure function is the result of three parameters:

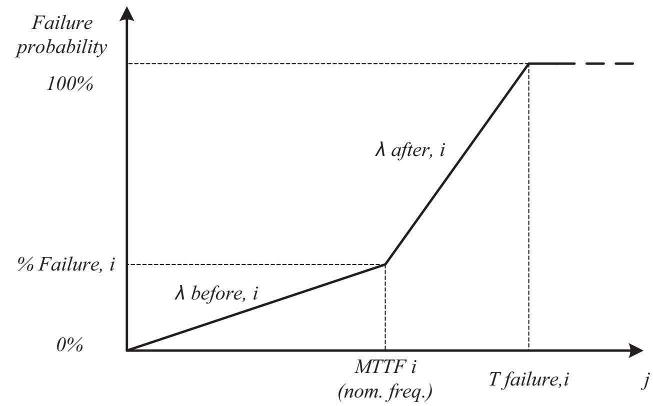

Fig. 1. Path-wise linear failure probability function for the evaluation of the unplanned costs of task i.

(1) $T_{failure}$, the time from the last execution of task i to the unit of time corresponding to "failure certainty", i.e., the failure probability is assumed equal to 1. Obviously this is a model assumption.
(2) the *MTTF*, named as the nominal frequency f_i.
(3) $\%_{failure,i}$, the percentage of failure occurrence at the time equal to the *MTTF*.

These parameters uniquely define the constant failure rates $\lambda_{before,i}$ and $\lambda_{after,i}$, which refer to the period "before" and the period "after" the nominal frequency (see "nom. freq." in Fig. 1).

Because many mechatronic devices compose complex machines, the manufacturer does not necessarily know the failure behavior of such subsystems. Although it is not responsible for reliability assessment and predictions, the manufacturer can quickly quantify the aforementioned parameters ($T_{failure}$, f_i and $\%_{failure,i}$) through on-field monitoring and interviews with maintenance operators. We assume that the generic service order stops the machine during the time bucket. Then, the number of service orders is established in advanced by a contract. The provider of the maintenance service tries to best plan and schedule the preventive actions to guarantee the standard and nominal productivity of the systems, i.e., reduce the total cost of maintenance, including the repair costs of unplanned actions.

Other important assumptions at the basis of the proposed model:

- the set of tasks to be scheduled is known;
- the number of time buckets is pre-defined;
- the duration of a specific task is constant;
- the unit costs (e.g. spare parts, personnel, additional failure cost) are known and deterministic.

The proposed original mixed integer linear programming (MILP) model for the assignment of preventive maintenance actions to the available and capacity constraints time buckets is following.

The objective function is:

$$\min z = \sum_{i,j,k} a_{ijk}(Cf_i + d_i C_{lab,k}) + Cr_i(\lambda_{before\,i}(f_i a_{ijk} - q_{ijk}) + \lambda_{after\,i} r_{ijk}) \quad (1)$$

subject to

$$\sum_i (d_i/tech_{jk}a_{ijk} + Wu_i h_{ij}) \leqslant Cp_{jk} \quad \forall j,k \tag{2}$$

$$\sum_k y_{ik} = 1 \quad \forall i \tag{3}$$

$$\sum_j a_{ijk} \leqslant y_{ik}M \quad \forall i,k \tag{4}$$

$$h_{ij} = \sum_k a_{ijk} \quad \forall i,j \text{ where } Cp_{jk} > 0 \text{ and } Wu_i > 0 \tag{5}$$

$$\sum_{s\in\{j-T_{failure,i}+1,j\}} a_{isk} \geqslant 1\cdot y_{ik} \quad \forall i,j,k \text{ and } j \geqslant T_{failure,i} \tag{6}$$

$$\sum_j a_{ijk} \geqslant 1\cdot y_{ik} \quad \forall i,j,k \text{ and such that } f_i \leqslant n \tag{7}$$

$$b_{ijk} \leqslant b_{i,j-1,k} + 1 \quad \forall i,j,k \tag{8}$$

$$b_{ijk} \geqslant b_{i,j-1,k} + 1 - M\cdot a_{ijk} \quad \forall i,j,k \tag{9}$$

$$b_{ijk} \leqslant M(1-a_{ijk}) \quad \forall i,j,k \tag{10}$$

$$b_{i0k} = 0 \quad \forall i,k \tag{11}$$

$$w_{ijk} \leqslant b_{i,j-1,k} + 1 - f_i \quad \forall i,j,k \tag{12}$$

$$w_{ijk} \geqslant b_{i,j-1,k} + 1 - f_i - M(1-a_{ijk}) \quad \forall i,j,k \tag{13}$$

$$w_{ijk} \leqslant Ma_{ijk} \quad \forall i,j,k \tag{14}$$

$$r_{ijk} - q_{ijk} - p_{ijk} = w_{ijk} \quad \forall i,j,k \tag{15}$$

$$p_{ijk} \leqslant M(1-a_{ijk}) \quad \forall i,j,k \tag{16}$$

$$r_{ijk} \leqslant M\ del_{ijk} \quad \forall i,j,k \tag{17}$$

$$q_{ijk} \leqslant M\ ear_{ijk} \quad \forall i,j,k \tag{18}$$

$$ear_{ijk} + del_{ijk} = a_{ijk} \quad \forall i,j,k \tag{19}$$

M is a constant subject to:

$$M \gg n \text{ and } \gg \max(f_i) \tag{20}$$

where

$i = 1,\ldots,m$	Tasks
$j = 1,\ldots,n$	Time buckets
$k = 1,\ldots,K$	Operators' typologies
d_i	Duration of task i
f_i	Nominal frequency of task i. It is valued in number of time buckets.
$C_{lab,k}$	Labor cost of operator k per time unit, for example (€/h)
Cp_{jk}	Capacity of service order j with operator k. This capacity is the result of the number and the typologies of the operators involved.
Cf_i	Fixed cost of task i. It includes the spare parts contribution.
Cr_i	Additional failure cost of task i. This is the additional unplanned and probabilistic cost (in case a failure occurs)
$T_{failure,i}$	The time to failure certainty for the task i according with the previously defined Path-wise linear failure probability function (see Fig. 1)
$\lambda_{before,i}$	Constant failure rate before the nominal frequency
$\lambda_{after,i}$	Constant failure rate after the nominal frequency
$tech_{jk}$	Number of operators assigned to the service order j in case the task is executed by the typology of operators k
Wu_i	Warm-up time for the task i. This is the time window after the execution of the task. It is a warm-up time to have the system ready to produce at nominal conditions. This time reduces the available time to conduct preventive maintenance actions during the generic time bucket

The decisional variables are:

$$a_{ijk} = \begin{cases} 1 & \text{if task } i \text{ is executed in time bucket } j \text{ by operators of typology } k \\ 0 & \text{otherwise} \end{cases}$$

$$y_{ik} = \begin{cases} 1 & \text{if task } i \text{ is assigned to operators } k \text{ of typology } k \\ 0 & \text{otherwise} \end{cases}$$

$b_{ijk} \geqslant 0$	the distance from the last execution of task i, adopting the operators of typology k
w_{ijk}	shift time from the unit time corresponding to the nominal frequency execution of task i, given the operators of typology k
$r_{ijk} \geqslant 0$	delay time of task i from the nominal frequency, given the operators of typology k
$q_{ijk} \geqslant 0$	early time of task i from the nominal frequency, given the operators of typology k
$p_{ijk} \geqslant 0$	auxiliary variable, given the operators of typology k

$$ear_{ijk} = \begin{cases} 1 & \text{if task } i \text{ is executed at the nominal frequency} \\ & \text{or before it} \\ 0 & \text{otherwise} \end{cases}$$

$$del_{ijk} = \begin{cases} 1 & \text{if task } i \text{ is exceuted with a delay considering the} \\ & \text{nominal frequency} \\ 0 & \text{otherwise} \end{cases}$$

$$h_{ij} = \begin{cases} 1 & \text{if task } i \text{ is executed in service order } j \\ 0 & \text{otherwise} \end{cases}$$

Eq. (2) are the temporal capacity constraints given the generic service order j. In Eq. (3) the task i is assigned uniquely to one operator k. Thanks to Eq. (4) that task i is definitely devoted to the operator k, through the whole taskplan of duration T units of time. This assumption is common in real applications to facilitate the execution of tasks.

Eq. (5) introduces the previously defined variable h_{ij} to manage the warm-up time not associated with a specific typology of operators k. Eq. (6) sets the range, after the nominal frequency, in which the task can be scheduled. The basic assumption is that the task is scheduled not later than $T_{failure,i}$ corresponding to the event "certainty of failure".

Eq. (7) forces the execution of tasks with nominal frequency f_i shorter than the total number of service order n selected for the taskplan.

Eqs. 8–11 define the time spent from the last execution of a task by the introduction of the variable b_{ijk}. In particular, Eq. (11) initializes the variable b_{ijk} to 0, which is reasonable for parts and machines that are new at time 0.

Eqs. 12–14 define the time shift w_{ijk} between each execution and the nominal frequency of task i.

Eqs. 15–19 quantify the early time q_{ijk} or the delay time r_{ijk} as previously defined. In particular, Eq. (19) establishes that the two binary variables ear_{ijk} and del_{ijk} cannot be equal to 1 at the same time and are both forced to 0 if a_{ijk} is 0.

Table 1 exemplifies the role of the main variables of the model and the constraints from (8) to (18) involving those variables. A planning period of 13 units of time is selected and illustrated, from $j = 3$ to 15. This is the so-called "time line". Assume that the first execution of a task i is placed in j equal to 4, which corresponds to the nominal frequency of execution for the task i.

From Table 1:

- The task is executed at j equal to 4, 10, and 13. However, the nominal frequency, after each execution, would be in j equal to 4 (i.e., 4 periods from the beginning), 8 (i.e., 4 periods from the unit time 4), and 14 (i.e., 4 periods from the unit time 10).
- The variable b counts the time that passes from the last execution. b is 0 when the task is executed (i.e., a assumes the value 1).
- The variable w defines the shift between each execution and the nominal frequency based on the last execution. Therefore, if the task is not executed ($a = 0$), w is null. In the other cases, if the task is delayed, w is positive. If the task is anticipated, w is negative. If the task is executed at its nominal frequency, w is 0.
- The variable r calculates the delay (if w is positive).
- The variable q calculates the anticipation (if w is negative).

4. Case study

This section presents a real case study of scheduling preventive maintenance actions for complex packaging machines. We decided to illustrate this case in a what-if environment in order to demonstrate the effectiveness of the proposed model when subject to different system configurations and instance dimensions.

This case deals with the maintenance scheduling of maintenance service provider of a leading company that manufactures automatic packaging systems and machines located in different countries worldwide. The provider has to minimize the global cost of maintenance due to the preventive and corrective actions. More precisely, this cost is due to spare parts use, which is necessary in both preventive and corrective actions (see Cf_i and Cr_i) and personnel use. In this case study, the cost of personnel is assumed to be equal to zero because the model fits the aim of the provider who defines the specifications of a maintenance plan for a future planning period. The cost of local personnel use is not charged to the provider. Finally, the cost of the provider's personnel is equal to zero because it is defined in advance in this case study; in the

Table 1
Variables a, b, w, r and q. An example.

Equations of the model		Time buckets j												
	j	3	4	5	6	7	8	9	10	11	12	13	14	15
	a_{1j}	0	1	0	0	0	0	0	1	0	0	1	0	0
	O = on-time execution; D = delay execution; E = early execution		O						D			E		
From (8) to (11)	$b_{ij} = \begin{cases} if & a_{ij} = 0,\ then\ b_{ij-1} + 1 \\ else & b_{ij} = 0 \end{cases}$	3	0	1	2	3	4	5	0	1	2	0	1	2
From (12) to (14)	$w_{ij} = \begin{cases} if & a_{ij} = 1,\ then\ w_{ij} = b_{ij-1} + 1 - f_i \\ else & w_{ij} = 0 \end{cases}$	0	0	0	0	0	0	0	2	0	0	−1	0	0
From (15) to (18)	$r_{ij} = \begin{cases} if & w_{ij} > 0,\ then\ r_{ij} = w_{ij} \\ else & r_{ij} = 0 \end{cases}$	0	0	0	0	0	0	0	2	0	0	0	0	0
	$q_{ij} = \begin{cases} if & w_{ij} < 0,\ then\ q_{ij} = -w_{ij} \\ else & q_{ij} = 0 \end{cases}$	0	0	0	0	0	0	0	0	0	0	1	0	0

Table 2
Static analysis. Statistics before scheduling.

Instance	Index	Instances (i.e., Tasks × Service Orders)			
		Instance 1 20 × 20	Instance 2 20 × 40	Instance 3 40 × 20	Instance 4 40 × 40
A	Total number of task executed – no constraints	92	188	184	376
B	Avg. number of tasks per service order	4.6	4.7	9.2	9.4
C	Time ratio, i.e., (the total requested time)/(total available time – no capacity constraints)	76.1%	77.2%	76.1%	77.2%
D	Total cost of spare parts for the tasks executed without capacity constraints (K€)	154.7	314.4	309.4	628.8
E	Total unplanned cost due to failure risk before the nominal frequency and in agreement with the failure rate λ_{before} (K€)	2.285	5.72	4.57	11.44
F	Objective function without capacity constraints, i.e., the total cost (total cost of spare parts + total unplanned cost) (K€)	156.98	320.12	313.97	640.2

Table 3
Instances 1–4. Comparison between the results generated by the solver and the static KPI.

KPIs comparison	Instance 1 (20 × 20)		Instance 3 (40 × 20)		Instance 2 (20 × 40)		Instance 4 (40 × 40)	
	New KPI	Static KPI	New KPI	Static KPI	New KPI	Static KPI	New KPI	Static KPI
Time (sec)	24		600		600		600	
O.F. (K€)	146.6	156.9	293.1	313.9	302.6	320.1	598.7	640.2
Gap%	0%		3.7%		7%		8.3%	
Planned cost (K€)	143.3	154.7	286.6	309.4	291.3	314.4	574.6	628.8
% of planned cost over the total cost	97.7%	98.5%	97.8%	98.5%	96.2%	98.2%	95.9%	98.2%
Number of executed tasks	84	92	168	184	174	188	346	376

generic service order, the involvement of the operators from the provider is fixed and decided in advance in the contract between the provider and the user.

The proposed model is solved by the Gurobi solver via the AMPL (A Mathematical programming Language, © 2013 AMPL Optimization Inc.) software using an Intel® Quad Core, 2.4 GHz processors, with 8 GB RAM. We illustrate the results obtained by the application to four different subsets of tasks and time buckets belonging to the case study:

- Instance 1, made of 20 tasks and 20 time buckets.
- Instance 2, made of 20 tasks and 40 time buckets.
- Instance 3, made of 40 tasks and 20 time buckets.
- Instance 4, made of 40 tasks and 40 time buckets.

The duration of the generic task is deterministic and is from 15 to 120 min. The nominal frequency f_i can be from 1 to 20, i.e., there are tasks executed just once in the planning horizon and tasks executed at every time bucket, that is, 20 times for Instance 1. The number of typologies of operators is equal to two to distinguish the "service provider personnel" from the "user personnel" but without any costs.

Table 2 presents a few statistics and key performance indicators – KPI, for the four selected instances. They measure the level of complexity and feasibility of the proposed model when applied before scheduling, i.e., before the execution of the MILP solver. This is the reason we call this "static analysis". Index A quantifies the number of scheduled tasks with respect to the nominal frequency and in the absence of capacity constraints. Similarly, index B quantifies the average number of tasks per time bucket, i.e., service order. Index C measures the ratio of total requested time as a result of the execution of tasks in agreement with the nominal frequency and the total available time in the absence of capacity constraints. The higher this ratio is, the more difficult it is to find admissible solutions to the capacity constraints problem. In particular, this KPI is more than 76% for the four selected instances.

The total cost of spare parts (index D) refers to the cost generated by the execution of preventive maintenance actions in the absence of unplanned (corrective) and reliability-based actions.

Index E quantifies the unplanned cost due to Eq. (1) and the existence of failure rate λ_{before} for each task i. Finally, index F quantifies the total objective function (planned and unplanned contributions).

Table 4
Distribution of scheduled tasks.

Instance	Delay tasks	On-time tasks	Early scheduled tasks	Total scheduled tasks
1 (20 × 20)	3	69	12	84
2 (20 × 40)	12	148	14	174
3 (40 × 20)	15	145	25	185
4 (40 × 40)	25	289	32	346

Table 5.1
Instance 4. Input data for the 2 pairs of tasks.

Task	Nominal frequency (no. of service orders)	$T_{failure}$ (no. of service orders)	*Spares cost* (preventive maintenance unit cost)	Additional failure cost (unit cost)
a	16	24	1.8	0.2
b	16	24	0.3	3.3
c	8	12	0.4	1
d	8	12	2	0.2

Table 5.2
Instance 4, schedule (Legend: O = on-time execution; D = delay execution; E = early execution).

Task \ j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
a																				D																				
b																O																O								
c								O								O						E								O										
d												D												D								O								

4.1. Results in a what-if environment

Given the NP-hard complexity of the decision problem, the computation time to achieve the optimal solution depends on the considered instances. Table 3 reports the percentage gap between the obtained feasible solution and the best lower bound computed by the solver.

For Instance 1, the illustrated solution is the optimal (Gap is 0%). It is generated in 24 s. The objective function, O.F. (K€) in Table 3, performs better than the one quantified before scheduling by the previously defined static analysis (see Table 2). The number of tasks executed is 84, which differs from the 92 estimated by the static analysis. Similar results are obtained for different solving times when applied to the other selected and more complex instances.

Table 4 presents the distribution of the scheduled tasks in "delay tasks" (1), "on-time tasks" (2), i.e., tasks scheduled in the time bucket corresponding to the nominal frequency, and "early scheduled tasks" (3), i.e., scheduled before the nominal frequency. This demonstrates that the adoption of the nominal frequency of scheduling does not reveal the optimal solution, even if a large amount of tasks are executed on time with respect to the nominal frequency (e.g., 289 tasks of 349 scheduled for the Instance 4). Given a task, sometimes it is scheduled on time and sometimes it is scheduled before or after the nominal frequency, owing to the effects of cost data, failure rates and capacity constraints.

Focusing on the case of Instance 4, Table 5 reports a selection of tasks for a better comprehension of the proposed model and the generated solutions. Table 5.1 presents the input data for two pairs of tasks selected from the group of 40. In particular, considering the pair of tasks (a, b), the cost of preventive (i.e., planned) maintenance for task a is lower than that cost for b and vice versa, given the additional failure (i.e., unplanned) cost. Table 5.2 illustrates the executions for the selected pairs of tasks in the planned task plan. The timeline reported in Table 5.2 is [1, 40] (i.e., $T = 40$ time buckets).

Table 6 presents the summary results for the selected pairs of tasks. In particular, given the pair of tasks (a, b), b is executed more frequently because of its higher "unplanned" unit cost. The execution of task a is expensive each time it is executed, but the cost for the delay due to the failure rate is not relevant. This is the reason it is executed in time bucket 20 with a delay.

Similarly, considering the pair couple of tasks (c, d), task c has a higher failure cost, so delays are avoided. Task d presents low failure cost but significant preventive costs, mainly due to the cost of the expensive spare parts. As a result, c is executed more frequently than d.

Table 7 shows the values of the objective function for Instance 4 (i.e., 40 × 40) and different available solving times, demonstrating that sub-optimal solutions are evidently effective for real industry applications. The renouncing of the optimum solution is necessary in the presence of large and complex instances. A feasible but not optimal solution drives the decision-maker toward the opportunity to plan the preventive actions, the use of personnel and the use of spare parts, quantifying the global expected cost. The assessment of the expected cost drives the discussion on the targets and costs of a service of maintenance between the supplier and the user of a production system.

Table 6
Results generated by the solver for the pairs of tasks (a and b) and (c and d).

Task	Number of executions in the planning period T	Avg. distance between executions	Cost paid along task plan	Planned cost	Failure cost
a	1	20	1.93	1.8	0.13
b	2	16	0.6	0.6	0
c	4	7.5	1.79	1.6	0.19
d	3	11	6.4	6	0.4

Table 7
Objective function for different available solving times. Instance 4 (40 × 40).

Solving time (sec)	Taskplan objective function (K€)
15	668.1
30	616.9
60	603.2
300	598.7
900	598.7
3600	598.7

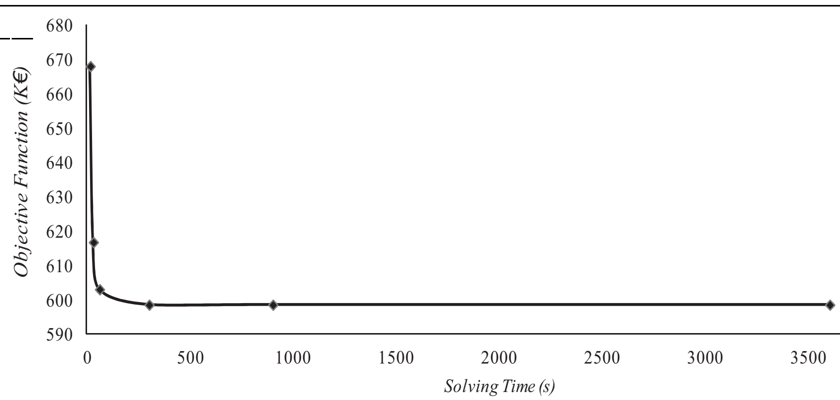

Table 8
Sensitivity analysis. Results from different solving times. Instance 1.

Static KPI Instance 1										
Index C (%)		131	114	97.9	85.6	**76.1**	45.7	38.1	32.6	11.4
Objective function (K€)	156.98	Infeasible	152.3	146.7	146.6	**146.6**	146.6	142.4	142.2	142.2
Solving time to find the optimal solution (sec)			2220	75	50	**24**	40	180	70	13
Gap (%)			0	0	0	**0**	0	0	0	0
Planned cost (K€)	15.98		139.7	143.3	143.3	**143.3**	143.3	139.9	139.8	139.8
% of planned cost above the total cost (%)	100		91.7	97.7	97.7	**97.7**	97.7	98.2	98.3	98.3
Number of executed tasks	92		81	84	84	**84**	84	82	81	81

The analysis conducted in this section is not a validation but a discussion of the results obtained by the application of the proposed model to a set of realistic case studies searching for (and renouncing to) optimal solutions.

4.2. Sensitivity analysis for different solving times

Table 8 presents the results of the proposed scheduling model applied to the tasks and time buckets of the Instance 1, for different values of the aforementioned index C (i.e., the ratio between the requested theoretical time and the available time) defined in the static analysis. The original value of C in Instance 1 was 76.1% (see the bold numbers in Table 8). This analysis demonstrates that the more constrained the instance (e.g., $C = 114\%$), the higher the number of tasks to be delayed.

5. Conclusions and further research

This paper presents an original cost-based and reliability-based MILP model for scheduling preventive tasks on complex machines subject to failure. In particular, a set of tasks are available, and for each task, there is a nominal frequency of execution according to the MTTF of the parts and components involved. The aim is to define the best schedule that minimizes the global cost of maintenance due to the planned, i.e., preventive, and corrective maintenance cost, in agreement with the capacity constraints. A failure probability is introduced in the model with an important role in the objective function to quantify the additional cost of maintenance due to the events of failure. The proposed model differs from the existing contributions illustrated by the literature because it is cost-based and reliability-based with respect to the capacity constraints, as necessary in real applications.

This model can support the production planner and the machine vendor in designing the task plan of a complex production system subject to failure for mid- to long-term periods and in optimizing the plan to minimize the maintenance expected costs, maximize benefits and control failure events. In particular, the proposed economic evaluation addresses the maintenance service provider as well as the clients in finding the best schedule that satisfies the goals of both stakeholders. This is an ex-ante and strategic maintenance service planning.

Furthermore, the proposed model can be applied by a company to plan the short-term (e.g., weekly) preventive actions and the use of resources, e.g., spare parts, equipment and personnel, necessary to conduct maintenance actions. This is further ex-ante but operational planning.

Finally, this model can support the ex-post analysis of the historical preventive and corrective maintenance plans to plan and conduct reliability evaluations on the most critical parts and components. This is ex-post planning.

The model has been applied to a complex case study that demonstrates how effective it is in reducing the total cost and improving the quality of service, even in the presence of a large number of tasks and time buckets.

Further developments will focus on the impacts of adopting multi-criteria and multi-objective optimization in task plan management and scheduling. Reliability-based and cost-based criteria can be assessed compared with other criteria, for example, the balancing of the expected saturation of the time buckets (i.e., service orders).

Further research is also expected on the development and application of new reliability-based scheduling models and methods, including the assessment of logistic issues, e.g., the management of spare parts, personnel (crew teams), and equipment, in a multi-location environment made up of different locations worldwide and in agreement with the customer service level and cost reduction.

References

Alardhi, M., & Labib, A. W. (2008). Preventive maintenance scheduling of multi-cogeneration plants using integer programming. *Journal of the Operational Research Society, 59*(4), 503–509.

Bajestani, M. A., & Banjevic, D. (2014). Integrated maintenance planning and production scheduling with Markovian deteriorating machine conditions. *International Journal of Production Research, 52*(24), 7377–7400.

Bang-Jensen, J., & Larsen, R. (2012). Efficient algorithms for real-life instances of the variable size bin packing problem. *Computers and Operations Research, 39*(11), 2848–2857.

Bell, S. A., & Percy, D. F. (2012). Modelling uncertainty in preventive maintenance scheduling. *Quality and Reliability Engineering International, 28*(6), 604–615.

Cassady, C. R., & Kutanoglu, E. (2005). Integrating preventive maintenance planning and production scheduling for a single machine. *IEEE Transactions on Reliability, 54*(2), 304–309.

Duffuaa, S. O., & Al-Sultan, K. S. (1999). A stochastic programming model for scheduling maintenance personnel. *Applied Mathematical Modelling, 25*, 385–397.

Ebrahimipour, V., Najjarbashi, A., & Sheikhalishahi, M. (2013). Multi-objective modeling for preventive maintenance scheduling in a multiple production line. *Journal of Intelligent Manufacturing*.

Gustavsson, E., Patriksson, M., Strömberg, A.-B., Wojciechowski, A., & Önnheim, M. (2014). Preventive maintenance scheduling of multi-component systems with interval costs. *Computers & Industrial Engineering, 76*, 390–400.

Hadidi, L. A., Al-Turki, U. M., & Rahim, A. (2012). Integrated models in production planning and scheduling, maintenance and quality: A review. *International Journal of Industrial and Systems Engineering, 10*(1), 21–50.

Hui, Y., Zheng, H., Liu, J., Zhao, Y., & Sun, Z. (2013). A scheduling model based on genetic algorithms for parallel and batched maintenance tasks. In *2013 International conference on quality, reliability, risk, maintenance, and safety engineering (QR2MSE)* (pp. 744–747).

Jardine, A. K. S., & Tsang, A. H. C. (2006). *Maintenance, replacement, and reliability. Theory and applications*. Taylor & Francis.

Kim, S., & Yoo, S. (2012). The combined manpower planning and preventive maintenance strategies by aggregate planning for the service division of vending machines. *Journal of the Operational Research Society, 63*(8), 1034–1050.

Kuo, Y., & Chang, Z. A. (2007). Integrated production scheduling and preventive maintenance planning for a single machine under a cumulative damage failure process. *Naval Research Logistics (NRL), 54*(6), 602–614.

Levi, R., Magnanti, T., Muckstadt, J., Segev, D., & Zarybnisky, E. (2014). Maintenance scheduling for modular systems: Modeling and algorithms. *Naval Research Logistics, 61*(6), 472–488.

Manzini, R., Regattieri, A., Pham, H., & Ferrari, E. (2010). *Maintenance for industrial systems. Springer series in reliability engineering*. London: Springer-Verlag.

Moghaddam, S. K., & Usher, J. S. (2011). Preventive maintenance and replacement scheduling for repairable and maintainable systems using dynamic programming. *Computers & Industrial Engineering, 60*(4), 654–665.

Perez Canto, S. (2011). Using 0/1 mixed integer linear programming to solve a reliability-centered problem of power plant preventive maintenance scheduling. *Optimization and Engineering, 12*(3), 333–347.

Pinedo, M. L. (2005). *Planning and scheduling in manufacturing and sevices*. Springer.

Raza, S. A., & Al-Turki, U. M. (2007). A comparative study of heuristic algorithms to solve maintenance scheduling problem. *Journal of Quality in Maintenance Engineering, 13*(4), 398–410.

Regattieri, A., Manzini, R., & Battini, D. (2010). Estimating reliability characteristics in presence of censored data: a case study in a light commercial vehicle manufacturing system. *Reliability Engineering & System Safety, 95*, 1093–1102.

Sherwin, D. (2000). A review of overall models for maintenance management. *Journal of Quality in Maintenance Engineering, 6*(3), 138–164.

Tam, A. S. B., Chan, W. M., & Price, J. W. H. (2006). Optimal maintenance intervals for a multi-component system. *Production Planning & Control: The Management of Operations, 17*(8), 769–779.

Tantardini, M., Portioli-Staudacher, A., & Macchi, M. (2014). A model for considering the impact of rescheduling planned maintenance activities in a maintenance service contract. *Production Planning and Control, 25*(3), 241–259.

Tarakci, H., Ponnaiyan, S., & Kulkarni, S. (2014). Maintenance-outsourcing contracts for a system with backup machines. *International Journal of Production Research, 52*(11), 3259–3272.

Xiang, Y., Cassady, C. R., Jin, T., & Zhang, C. W. (2014). Joint production and maintenance planning with machine deterioration and random yield. *International Journal of Production Research, 52*(6), 1644–1657.

Xu, B., Xueshan, H., Wang, Y., & Sun, D. (2012). Risk-based optimization model for system maintenance scheduling. In *IEEE international conference on power system technology (POWERCON), 2012*.

Zhang, T., & Nakamura, M. (2005). Reliability-based optimal maintenance scheduling by considering maintenance effect to reduce cost. *Quality and Reliability Engineering International, 21*, 203–220.