Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

USING MONTE-CARLO SIMULATION TO MEASURE THE BUSINESS-RELEVANT IMPACT OF PLANNING UNCERTAINTY ON FIELD SERVICE DELIVERY

Clemens Wolff
Michael Vössing
Niklas Kühl

Karlsruhe Institute of Technology
Karlsruhe Service Research Institute
Karlsruhe, Germany

ABSTRACT

Today, delivery of industrial maintenance services is planned under uncertainty, as, for example, the true task duration is unknown. Thus, providers rely on estimates at the stage of planning. During operations, the true task duration becomes known which results in a schedule update. This work evaluates the effects of information uncertainty on business-relevant performance indicators in industrial maintenance. We present a simulation model to reflect the evolution of information and the schedule revision during operations. Within a simulation experiment, we measure the business-relevant impact of different uncertainty levels on operations. Indeed, the findings of our research confirm the common understanding that reduced uncertainty at the stage of planning has a positive effect on service delivery. However, our research also shows that in order to achieve business-relevant impact in service delivery quality, the reduction of uncertainty at the stage of planning needs to reach a certain threshold.

1 INTRODUCTION

In industrial maintenance – a prominent example of an industrial service according to Gitzel et al. (2016) – providers plan schedules for their engineers. However, those schedules are planned under uncertainty, as, for example, the true duration of a repair task is unknown until it is actually finished. Therefore, the schedule initially planned under uncertainty is constantly updated during operations as previous estimates become certain. Thus, technician scheduling should be seen as a dynamic problem (Vössing 2017). Similar remarks have already been made to delivery services on a broader notion by Psaraftis (1995). For example, a planner may have assumed an estimated task duration of two hours, however, once an engineer is on site, he notices that the true repair takes four hours. In this case, the engineer's schedule must be updated such that it reflects the new information. Another example that requires schedule update is the arrival of a previously unknown high priority task (e.g., repair).

Researchers aim to reduce uncertainty during planning by leveraging new technologies, as, for example, predictive methods and smart machinery. In industrial maintenance, for example, predictive methods can be used on the basis of real-time machine log data. However, so far, researchers have not yet evaluated the impact of reduced uncertainty at the state of planning on service delivery. Whilst there seems to be a common agreement on the positive effect of reduced uncertainty at the stage of planning on service delivery, we are not yet able to describe the effects further. We do not know whether the impact of reduced uncertainty has a linear effect on service delivery or, for example, whether the impact increases exponentially with reduced uncertainty.

Consequently, this work aims at evaluating the impact of reduced uncertainty at the stage of initial schedule creation on service delivery in industrial maintenance. In detail, we present a simulation experiment in which the impact of reduced task duration estimate error on delivery performance is measured. Hence, this work is not concerned on how uncertainty can be reduced. Instead, it only evaluates its impact once such methods exist. However, on the basis of this work, researchers can estimate the degree of uncertainty reduction they must reach such that a true benefit in service delivery is achieved. Therefore, this work contributes to research on service delivery, a major research priority in service science (Ostrom et al. 2010; Ostrom et al. 2015). Furthermore, this work contributes on research on planning under uncertainty by addressing the gap between predictive methods and their impact on the quality of service delivery.

The remainder of this work is structured as follows: In Section 2, we introduce fundamentals and related work. In Section 3, we explain the theoretical basis for the simulation model, which itself is presented in Section 4. The simulation experiment and its results are shown and discussed in Section 5. Finally, in Section 6, we summarize this work and point out limitations and future work.

2 FUNDAMENTALS AND RELATED WORK

In this section, we introduce fundamentals of technician dispatching and elaborate on related work.

2.1 Fundamentals

In industrial maintenance, a provider dispatches his engineers to his clients' sites in order to deliver the desired service (e.g. repair or overhaul). In practice, the process of dispatching – i.e., the timely assignment of service tasks to engineers – is done by human dispatchers that follow simple dispatching strategies (Hill 1992). During dispatching, a dispatcher usually tries to achieve two contradictory goals: First, the dispatcher wants the technician to be utilized as high as possible, whilst second, also having short-term capacity to schedule upcoming unplanned tasks that have a high priority (e.g., repair) (Wolff et al. 2018). Furthermore, efficient technician dispatching results in a competitive advantage, as expensive resources (the technicians) are used efficiently.

Within technician scheduling, Li and Ierapetritou (2008) differentiate between two concepts: First, an initial schedule generation coupled with schedule revision throughout the day – also referred to as predictive-reactive scheduling (e.g., Sabuncuoglu and Bayız 2000), and second, online and offline scheduling (e.g., Sabuncuoglu 1999).

Due to the uncertainty in dispatching, dispatchers usually follow *predictive-reactive scheduling* approaches (Vössing 2017), which aligns with the first concept. This means that an initial schedule is predicted, however, during operations, providers react on unforeseen events and update their schedule (Li and Ierapetritou 2008). Therefore, the technician dispatching problem can be seen as a two stage problem: First, an initial schedule must be created, and second, the revision of the current schedule after disruptive events. After operations have finished, the final schedule is most likely to differ from the initial schedule. In the following, we briefly summarize work on the two stages of the technician dispatching problem.

The first stage of the technician dispatching problem – the initial schedule creation – is an extension to the *Vehicle Routing Problem (VRP)*, which was initially introduced by Dantzig and Ramser (1959) as the *Truck Dispatching Problem*. Since then, many modifications have been published in order to reflect individual problem instances. The core of the VRP is described as "the problem of designing optimal delivery or collection routes from one or several depots to a number of geographically scattered cities or customers, subject to side constraints" (Laporte 1992). Until today, several reviews and taxonomies of the VRP and its extensions have been published, e.g., Eksioglu et al. (2009), Drexl (2012), and Lahyani et al. (2015). The VRP is the basis for many technician routing problems, as, for example, the work by Weigel, Don and Cao, Buyang (1999) or Kovacs et al. (2012).

Unfortunately, literature on the second stage – rescheduling within technician dispatching – is rare. Chaari et al. (2014) outline triggers and approaches for rescheduling. However, the presented rescheduling approaches for technician dispatching do not differ much from rescheduling in other domains.

2.2 Related Work

So far, research in technician dispatching has focused on the inclusion of uncertainty at the stage of initial schedule creation. A good overview of robust optimization within VRP is provided by Bertsimas and Simchi-Levi (1996). Furthermore, the work by Sungur et al. (2008) proposes a robust optimization approach for the VRP with demand uncertainty. In addition, Souyris et al. (2013) present a robust schedule optimization problem under uncertain task duration.

Even though these works provide methods on how to deal with uncertainty in technician scheduling, they lack to measure the impact of reduced uncertainty on business-relevant performance indicators. Most studies evaluate their robust solution against the naive nominal solution based on the objective value. However, the objective value does reflect business performance.

3 CONCEPT OF TECHNICIAN SCHEDULING

In this section, we explain the problem of technician dispatching as understood for the remainder of this work. Given the interest in the effects of information uncertainty on service delivery, the following assumptions are made: First, the maintenance provider follows a *predictive-reactive scheduling* approach. Second, the initial schedule for an engineer is created on a daily basis and planning happens prior to shift start. Further information on how the initial schedule is derived are given in section 3.2. Third, information uncertainty at the stage of planning is limited to uncertain task duration only. At the stage of planning, an estimated task duration is used. Once the engineer arrives at a task's site (i.e., once it is planned to start according to the currently valid schedule), the true duration becomes known. The relation between the true and the estimated task duration is further explained in section 3.1. Fourth, schedule update (i.e., rescheduling or schedule revision) is an immediate process. Thus, once a task duration becomes known, the current schedule is immediately updated resulting in a new current schedule. Updating is done according to the *right-shift* policy as explained in section 3.3.

3.1 Task Duration Estimation

Due to uncertainty, the task's true duration d^t is unknown before the task is actually started. Therefore, during planning and rescheduling before the task has started, maintenance providers rely on a task duration estimate d^e. The difference between the true service duration and the estimated service duration is referred to as the estimate error Δt. The relation between the true and estimated task duration are shown in Equation (1). The estimate error Δt follows a pre-defined distribution defined by the simulation scenario. The estimate error distributions are introduced along the simulation experiment scenarios in section 5.1.

$$d^e = d^t + \Delta t \tag{1}$$

3.2 Initial Schedule Creation

Our formalization of the initial schedule creation problem closely follows the work of Petrakis et al. (2012), however, we simplify some aspects as they are not relevant for this work. First, only schedules for one day are created. Second, engineer skills are neglected. Third, we modify the objective function to reflect an optimization objective better aligned with our findings from practice. The objective of this work is to maximize the number of tasks served, however, at minimal provider costs. Furthermore, the problem presented in this work does not use any approach of robust scheduling.

The formalized scheduling problem can be considered as a graph $G = \{V, E\}$, in which the vertex V consists of two disjunctive subsets $V = \{D, C\}$. Set $C = \{0, 1, \ldots, m-1\}$ reflects the set of engineers and – for easier understanding – their home locations. Analogue, set $D = \{0, 1, \ldots, n-1\}$ reflects the tasks and their locations. The set of edges E denotes possible travel segments for the technicians, thus, $E = \{(i, j)\}$ with $i, j \in V$. The travel times between the locations are seen as symmetric, thus $c_{ij} = c_{ji}$ for $i, j \in V$ and

$c_{ii} = 0$ for $i \in V$. One minute of travelling costs the provider f. Furthermore, the maximum regular work time – which is a hard constraint during the initial schedule creation – is denoted by t_{max}. The estimated task duration for a task i is denoted by d_i^e with $i \in D$. The shift begin of all engineers is denoted by b.

We use two decision variables to formalize the initial schedule creation within the technician scheduling problem. First, variable $a_i, i \in D$ denotes the engineer's arrival time at task i. Second, the decision variable $x_{ijk}, i, j \in V, k \in C$ is defined as follows:

$$x_{ijk} = \begin{cases} 1 & \text{if engineer k travels from node i directly to node j} \\ 0 & \text{otherwise} \end{cases}$$

Given the above information, the technician dispatching problem can be defined as shown in 2:

$$\begin{aligned}
\underset{a,x}{\text{minimize}} \quad & \sum_{i\in V}\sum_{j\in V}\sum_{k\in C} c_{ij}x_{ijk}f + (n - \frac{1}{2}\sum_{i\in D}\sum_{j\in D}\sum_{k\in C} x_{ijk})p && && \text{(2a)}\\
\text{subject to} \quad & \sum_{i\in V\setminus\{j\}} (x_{ijk} - x_{jik}) = 0 && \forall j \in V, k \in C, && \text{(2b)}\\
& \sum_{j\in C} x_{ijk} \le 1 && \forall i \in D, k \in C, && \text{(2c)}\\
& \sum_{j\in C} x_{ijk} = 0 && \forall i \in D, k \in C \setminus \{i\}, && \text{(2d)}\\
& \sum_{i\in V}\sum_{k\in C} x_{ijk} \le 1 && \forall j \in C, && \text{(2e)}\\
& a_j \ge (b + c_{ij})x_{ijk} && \forall i \in D, j, k \in C, && \text{(2f)}\\
& a_j \ge (a_i + d_i^e + c_{ij})x_{ijk} && \forall i, j, k \in C, && \text{(2g)}\\
& (a_i + d_i^e + c_{ij})x_{ijk} \le b + t_{max} && \forall i, k \in C, j \in D, && \text{(2h)}\\
& x_{ijk} \in \{0, 1\} && \forall i, j \in V, k \in C && \text{(2i)}
\end{aligned}$$

Flow is conserved with constraint (2b), ensuring that all engineers that arrive at a task also leave it again. Constraint (2c) ensuring that each engineer leaves his home depot only once, whereas constraint 2d assigns each engineer to one home location. Constraint (2e) limit the amount of visits from engineers to one for each task. Constraints (2f) and (2g) ensure correct engineer arrival times at the tasks locations. Furthermore, constraint (2h) ceils the maximum working time of each engineer. Finally, constraint (2i) forces x to be binary, thus imposing a binary flow on the graph G. According to Miller et al. (1960) and Kulkarni and Bhave (1985), sub-tour elimination is implicitly done by constraints (2f), (2g), and (2h). The objective function minimizes total provider costs by including a penalty p for unscheduled tasks and travel time costs f per travelling minute. As overtime is not permitted within the initial schedule creation, each engineer works a maximum of his regular work time – for which he receives a fixed wage. Thus, regular work wages are schedule-independent and not included in the objective function. As such, the formalized problem is not a Mixed-Integer Program. However, by the introduction of auxiliary variables and large constants, non-linear constraints can be linearized (Petrakis et al. 2012).

3.3 Rescheduling

During operations, the current schedule for the given technician is updated following the *right-shift* policy every time a change in a task's duration becomes known. In this work, the true task duration becomes known when the engineer starts delivery of the task, i.e., when the task's current start time has been reached.

In this context, the current schedule is either the day's initial schedule for the day or an already modified schedule.

The *right-shift* policy – in its base form – naively updates the schedule by propagating duration changes for any element by propagating it through the schedule until its end is reached. Thus, in the naive approach, if the task duration of task i is changed by Δt_i, the updated start times $a_i^{updated}$ of subsequent tasks can be calculated as shown in Equation (3). We chose the *right-shift* policy as a suitable rescheduling approach, as it is a simple policy that minimizes both, deviation from the initial schedule and disruptions to the planning function of the schedule (Leon et al. 1994).

$$a_i^{updated} = a_i + \Delta t_i \tag{3}$$

Given the domain of industrial maintenance, two more constraints need to be taken into account: First, an engineer cannot access a machine at any time, as industrial machines are involved in a production schedule and must be shut down prior to maintenance actions (Paz and Leigh 1994). In this work, we assume that the client plans engineer works according to the initial schedule and shuts the machine down 30 minutes prior the communicated start time. Thus, the new start times of tasks can be calculated according to Equation (4).

$$a_i^{updated} = \begin{cases} a_i + max(\Delta t_w, -30) & \Delta t_w < 0 \text{ (i.e., true task duration < estimated task duration)} \\ a_i + \Delta t_w & \Delta t_w > 0 \text{ (i.e., true task duration > estimated task duration)} \end{cases} \tag{4}$$

Second, an engineer has a maximum overtime after which he must return home. Consequently, if the application of the *right-shift* policy results in a violation of the maximum work time including overtime, tasks are removed from the end of the schedule until the engineer is able to return home without exceeding his overtime. Removed tasks would – in practice – be added to the backlog and be scheduled the following day. In case the overtime is exceeded when the last task's duration is extended, the engineer remains at the location and finishes the task. In this work, the maximum overtime t_o is set to $t_o = 2$ hours, which aligns with legal obligations in Germany. Figure 1 shows the applications of the *right-shift* policy for three examples. Using the described rescheduling police, technician schedules are independent of each other, as changes on one technician's schedule do not influence another technician's schedule.

3.4 Performance Indicators for Service Delivery

The impact of task duration estimate errors on the actual service delivery is evaluated on the basis of performance indicators. This work builds on performance measures for industrial maintenance as introduced by Meier et al. (2013). In detail, the following relevant performance indicators are used in this work:

- *Tasks performed:* The number of tasks performed by all engineers within the given day.
- *Engineer utilization:* The utilization (= on site and travel) of the engineers with regard to t_{max}
- *Travel Time:* The percentage of work time an engineer is travelling.
- *Tasks started on-time:* The percentage of tasks started on time. Every task started within $-15/+30$ minutes of the initially planned start time.
- *Cancelled Tasks:* Percentage of tasks that had to be cancelled as overtime was exceeded.
- *Overtime:* The sum of overtime (in hours) of all engineers. This indicator is directly cost-relevant for maintenance providers, as higher wages are paid during overtime.

4 SIMULATION MODEL

In this section, we further elaborate on the simulation model. First, we give a brief overview of the simulation environment. Second, we define its major components. Third, we talk about model validation, before fourth, talking about parametric values used.

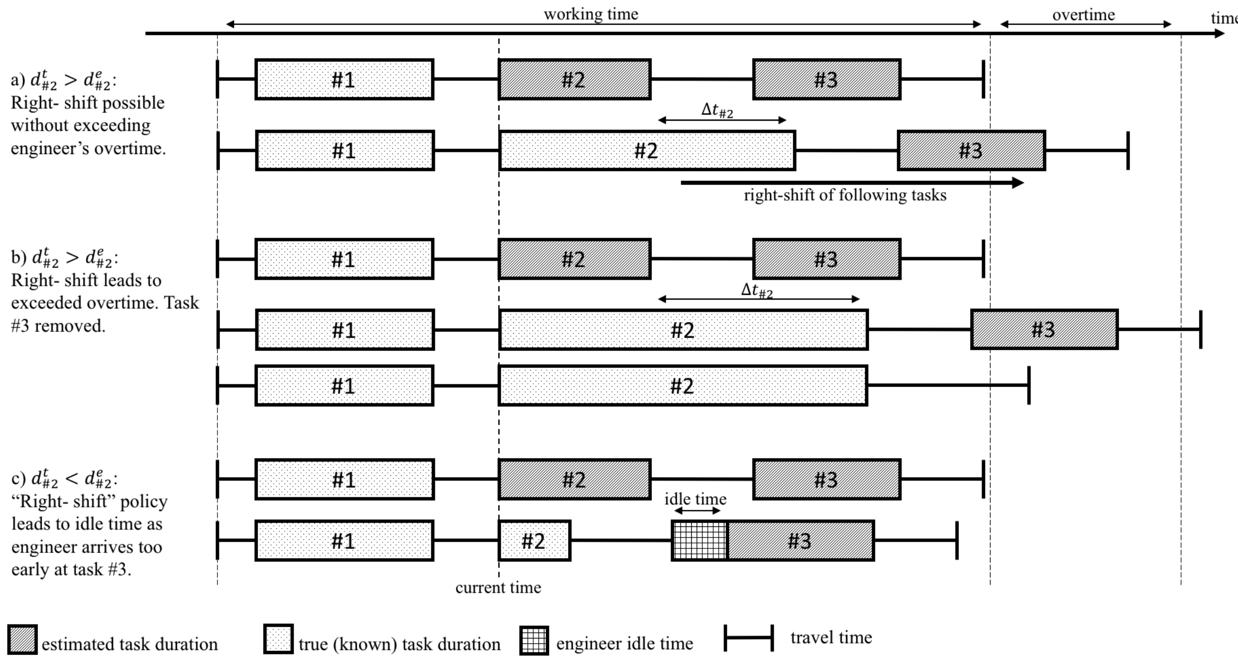

Figure 1: Examples for right-shift scheduling.

4.1 Simulation Environment Overview

Within the simulation model, we are able to simulate the entire process from the initial schedule creation through the evolution of information over time. With the evolution of information, the simulation model updates the current schedule to reflect the currently known information. Once the simulation has finished, the current schedule becomes the final schedule. Finally, the final schedule can be compared against the initial schedule.

The components of the simulation environment are displayed in Figure 2. As an input, the simulation model receives a list of tasks to be scheduled and a set of engineers that are available to deliver the requested services. Given this information, the initial schedule is created using the initial schedule solver, which is further explained in Section 4.2. Once the initial schedule is created, the environment simulates the modification of task duration at the time a task is scheduled to start. Thus, at the planned beginning of a task, its duration is modified to reflect the true task duration, opposed to the previously used estimated task duration. The new service duration then triggers the schedule to be updated. Finally, the final schedule is compared to the initial schedule in order to determine performance indicators as described in Section 3.4.

4.2 Simulation Environment Components

In the following, we walk through the individual components of the simulation environment and explain how they are implemented. First, we explain the initial schedule solver, before second, elaborating on the task duration and schedule update components. The last two components are realized using a discrete-event simulation approach. The components of the simulation environment are further explained below.

- *Initial Schedule Solver:* The initial schedule is using two heuristics – as described by Petrakis et al. (2012) – to solve the initial schedule problem. The solver first creates a valid schedule using the *Least Insertion Cost* heuristic, which is further improved by applying a *Variable Neighborhood Search*. For solving, we applied the neighborhoods as described by Petrakis et al. (2012) except for the *UR1_Swap_2*, *UR1_Swap_21*, *UR1_Swap_12*, *UR2_Relocate_1*, and *UR2_Swap_1* neighborhoods with 2000 iterations.

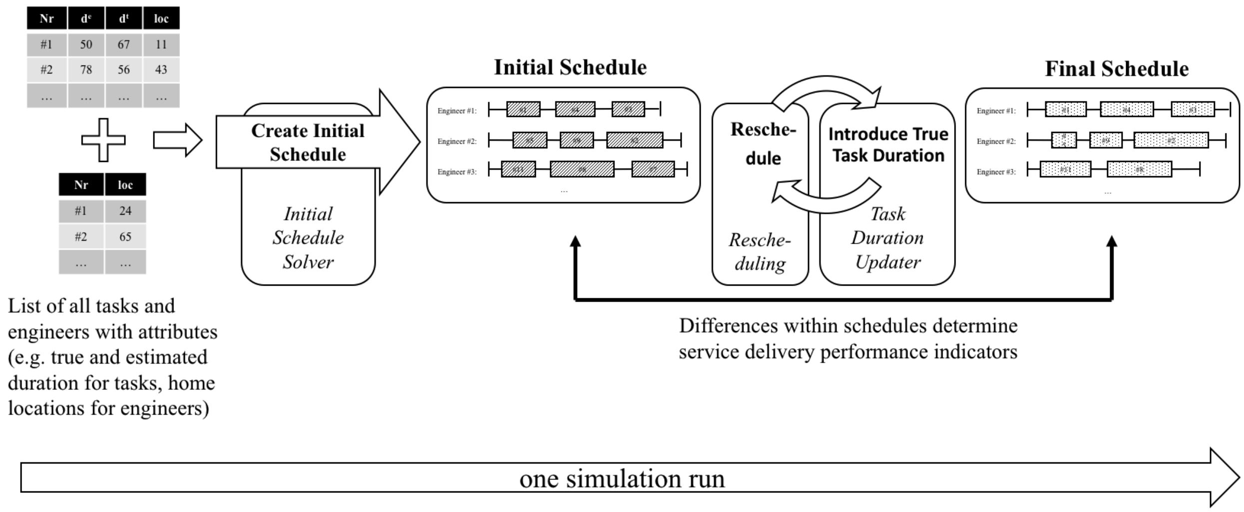

Figure 2: Simulation environment (actions and their simulation environment components).

- *Task Duration Updater:* This module reflects the introduction of new information. In detail, whenever a task starts, it unveils its true duration, as opposed to the previously used estimated duration. Therefore, this module provokes rescheduling.
- *Rescheduling:* This module is responsible for updating the current schedule within the simulation environment. It uses the *right-shift* policy as described in Section 3.3.

4.3 Model Validation

The simulation model was validated according to Sargent (2009). First, conceptual validation was done by presenting the initial underlying concepts to experts from two medium-sized German OEM providing maintenance services to their customers. The initial concepts were modified to include one remark the practitioners made. Second, the simulation model was validated by having two researchers following the schedule updates during $n = 150$ simulation runs closely, a validation method closely aligned with the validation through *Animation* (Sargent 2009).

4.4 Simulation Parameters

The simulation experiment was conducted using the parameters given in Table 1. Given the task duration and travel time distribution, an engineer should – on average–be able to conduct three tasks within his regular work time, resulting in a theoretical engineer utilization of ${}^{7.5h}/_{8h} = 93.75\%$ (three tasks and four journeys per day). The engineer set size is chosen to be 15, and for each engineer one additional task to the–on average–feasible three tasks is added, resulting in a total of 60 tasks. On execution of the simulation environment, a set of 2,000 locations is created which remains constant during the experiment. However, at each simulation run, tasks are randomly re-assigned to locations. Overall, the travel time in minutes between two locations follows a normal(45, 10) distribution. Furthermore, the deviation of the estimated task duration from the true task duration is provided by a deviation distribution. Given the context of this work, the deviation distribution is used as simulation experiment factor and its different values are introduced in the following section.

5 SIMULATION EXPERIMENT

As initial schedules are created on a daily basis and operational schedule updates are only reflected within the schedule for one engineer on one specific day, the schedules on different days can be seen as

Table 1: Parameter values used within the simulation experiment.

Parameter	Value	Parameter	Value (in min)
Variable Neighborhood Search Iterations	2000	Work Time t_{max}	480
Penalty for Unplanned Task p	5000	True Task Duration d^t	normal(90, 15)
Costs Travel Time f	5	Travel Time Matrix c	normal(45, 10)
Number Tasks n	60	max overtime t_o	120
Number Technicians m	15		

independent of each other. Therefore, different instances of the *initial schedule creation* problem can be treated independently of each other. Furthermore, we evaluate the impact of stochastic variables (estimate error) on service delivery performance. According to Raychaudhuri (2008), this is a typical case for the application of a *Monte Carlo simulation*, in which many individually independent instances of a stochastic problem set are generated, simulated, and finally, jointly analyzed. Due to the repetition, researchers evaluate many different combinations of stochastic instances and, therefore, increase the likelihood of gaining valid insights (Law 2015).

5.1 Experiment Factors

In Section 3.1, we introduced the notion of the true and estimated task duration. This work evaluates the effect of estimate errors on service delivery quality. Within the experiment, we vary the task duration estimate error distribution. Therefore, the estimate error distribution is the factor of this simulation experiment. The different estimate error distributions are provided in Table 2 and each error distribution reflects one simulation experiment scenario. Hence, with growing scenario number, the uncertainty at the stage of planning is reduced, until finally – in scenario 9 – no estimate error (i.e., no uncertainty) is assumed. All error distributions follow a normal distribution with mean zero, indicating no systematic estimate error.

Table 2: Experiment scenarios – distribution of task duration error (in minutes).

Scenario	Error Distribution	Scenario	Error Distribution	Scenario	Error Distribution
1	normal(0, 120)	4	normal(0, 75)	7	normal(0, 30)
2	normal(0, 105)	5	normal(0, 60)	8	normal(0, 15)
3	normal(0, 90)	6	normal(0, 45)	9	normal(0, 0)

5.2 RESULTS

Each scenario is executed $k = 640$ times and the responses are measured. The average performance indicator values for each scenario are shown in Table 3. In the following, we briefly walk through the individual performance indicators and interpret their measures for the different scenarios.

Given the results as shown in table 3, our findings support the general understanding that reduced uncertainty at the stage of planning impacts service delivery positively. There is no performance indicator that shows a negative trend between a reduced estimate error and the performance indicator.

Furthermore, we note that the number of tasks performed increases with a reduced duration estimate error between scenarios one to six. This can be explained as a high estimate error resulting in rather large schedule updates may result in sudden, large idle times at the end of a shift for a technician. As no new tasks are planned throughout the day, idle time prior to shift end results in lost capacity.

In addition, we see that each, the utilization, the percentage of travel time, and the percentage of cancelled tasks, varies within a range of around 3% for each case. As explained, cancelled tasks need to be

Table 3: Experiment results (mean values per scenario).

Scenario	1	2	3	4	5	6	7	8	9
Tasks Performed	45.6	46.6	47.8	49.2	50.7	52.4	53.5	52.6	53.4
Utilization	0.9247	0.9306	0.9373	0.9443	0.9508	0.9537	0.9477	0.9209	0.9235
Travel Time	0.2527	0.2577	0.2620	0.2674	0.2730	0.2780	0.2807	0.2712	0.2772
Cancelled	0.0196	0.0221	0.0225	0.0223	0.0196	0.0133	0.0042	0	0
On-Time	0.5354	0.5264	0.5280	0.5285	0.5326	0.5601	0.6253	0.8206	1
Overtime (in h)	13.1	13.8	14.7	15.3	15.9	15.4	11.8	5.3	0

rescheduled at following days, therefore, being added to the *backlog* of tasks that is used as input for the technician dispatching problem of following days. It is likely that cancelled tasks result in penalties to be paid to the customer by the provider. The utilization ranges between 92.09% and 95.37%, the percentage of travel time between 25.27% and 28.07%, and the percentage of cancelled tasks between 0 and 2.25%. Furthermore, we observe no clear monotonic trend between a reduced task duration estimate error and those performance indicators. Therefore, we conclude that the engineer utilization, the percentage of travel time, and the percentage of cancelled tasks are rather independent of the task duration estimate error.

However, in addition, we want to highlight the following points: First, there are no cancelled tasks in any simulation run once the task duration estimate error was below normal(0, 15), i.e., scenario 8 and 9. Second, in Section 4.4, we calculated a theoretical utilization of 0.9375. In scenario 9 – which reflects no duration estimate error – our observations report a mean engineer utilization of 0.9234, which is very close to the theoretical utilization.

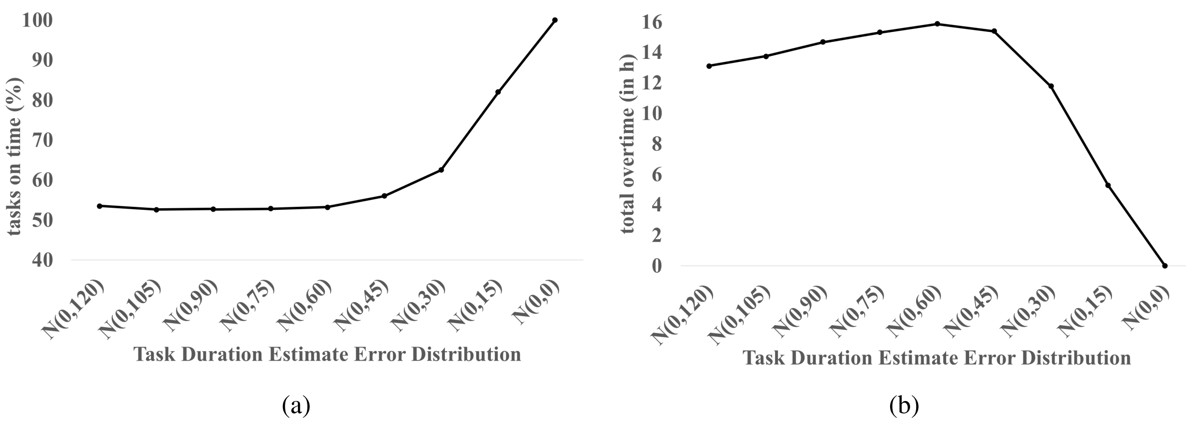

Figure 3: Mean performance indicator values; (a) on-time; (b) overtime.

When looking at *on-time* delivery – as depicted in figure 3a – we note a sudden increase at an estimate error distribution of normal(0, 30) or below (i.e., scenarios 7 to 9). From a delivery perspective, this is an important finding as it indicates that a service provider is required to meet a certain prediction accuracy in order to see a drastic increase in delivery performance. We want to point out that – in this work – the *on-time* performance indicator is defined by the percentage of tasks started within a -15/+30 minute time window around the estimated start of a task, which corresponds to the required minimum estimate error distribution to see a drastic increase in performance. Turning this thought around, this implies that service providers could – if they know their estimate errors – engineer performance indicators according to their estimate errors in order to optimize for the *on-time* performance indicator. However, those are initial thoughts and further research is required to further support this idea.

Furthermore, we see a drastic decrease in technician overtime with reduced estimate error, as depicted in Figure 3b. Our findings indicate that the drastic decrease starts once an estimate error distribution of normal(0, 45) or below (i.e., scenarios 6 to 9) is met. For the provider, this is very important as additional overtime results in increased engineer wages.

From a monetary perspective, overtime directly influences the provider's profit negatively, whereas the number of tasks performed ensures revenue. Given the assumption of this work that a provider does not schedule tasks initially during overtime, a provider tries to perform the highest number of tasks within the fewest overtime. Our findings observe that the number of performed tasks reaches its maximum in scenario seven. Furthermore, performance values for on-time delivery and overtime improve drastically within scenario seven. Consequently, a provider should not only try to reduce the duration estimate error, but reach an estimate error distribution of at least normal(0, 30) (i.e., scenarios 7 to 9) or below. Even with a further decrease in the estimated error, additional delivery quality can be achieved.

However, decreased estimate errors are likely to result in higher investments, as prediction methods are likely to be more complex. Therefore, a provider must trade off the benefits of additional delivery quality gained by a smaller estimate error with the additional costs of creating a model capable of predicting with the required estimate error.

6 CONCLUSION

Based on a simulation experiment, this work measures the impact of reduced uncertainty (in the form of task duration estimate errors) on service delivery performance in a field service delivery environment. In detail, uncertainty at the stage of schedule planning was introduced by adding a service task duration estimate error. During the simulation experiment, various distributions for the duration estimate error are assumed. Delivery quality is measured by the number of tasks performed, engineer utilization, travel time, cancelled tasks, on-time delivery, and engineer overtime.

The key findings of this work can be summarized in the following three points: First, the number of tasks performed, engineer utilization, travel time, and cancelled tasks seem to be rather independent of the duration estimate error. Second, *on-time* delivery improves greatly once the estimate error (in minutes) follows a distribution below normal(0, 30). Third, once the estimate error (in minutes) is below a normal(0, 45) distribution, engineers' overtime drastically decreases.

This work marks first steps into the evaluation of estimate errors on service delivery. Therefore, naturally, this work has limitations and lays the basis for further work: First, this work relies on assumed distributions for travel time, task duration, and estimate errors. Therefore, further work needs to be done on justifying those assumptions. Second, the optimization model presented in Section 3.2 needs to incorporate further constraints to better reflect practice. For example, skills were neglected. Additionally, customers may have constrained service time windows that need to be taken into account. In addition, this work can further profit from a differentiation between repair and overhaul tasks, as each type introduces individual requirements. Third, the initial schedule is based on naive scheduling. In practice, a provider assuming uncertainty may choose to create the initial schedule following robust optimization techniques. Finally, we call for similar research on other sources of uncertainty, as, for example, the arrival of urgent service demand.

Though this work explores a new direction of research, it already indicates first managerial implications. Given the positive impact of a reduced estimate error on delivery performance, providers of industrial maintenance should try to estimate uncertain information as good as possible. However, this work also shows that the impact is fairly low until a certain threshold has been met. Thus, maintenance providers should really drive towards finding this threshold in order to leverage the full potential of new predictive technologies.

ACKNOWLEDGMENTS

The research and development project that forms the basis for this report is funded under project No. 01MD16015B within the scope of the Smart Services World technology programme run by the Federal Ministry for Economic Affairs and Energy and is managed by the project management agency at the German Aerospace Center (DLR-PT). The author is responsible for the contents of this publication.

REFERENCES

Bertsimas, D. J., and D. Simchi-Levi. 1996. "A New Generation of Vehicle Routing Research: Robust Algorithms, Addressing Uncertainty". *Operations Research* 44(2):286–304.

Chaari, T., S. Chaabane, N. Aissani, and D. Trentesaux. 2014. "Scheduling under Uncertainty: Survey and Research Directions". In *Proceedings of the 2014 International Conference on Advanced Logistics and Transport (ICALT), 1st–3rd May, Tunis, Tunisia*, 229–234.

Dantzig, G. B., and J. H. Ramser. 1959. "The Truck Dispatching Problem". *Management Science* 6(1):80–91.

Drexl, M. 2012. "Rich Vehicle Routing in Theory and Practice". *Logistics Research* 5(1):47–63.

Eksioglu, B., A. V. Vural, and A. Reisman. 2009. "The Vehicle Routing Problem: A Taxonomic Review". *Computers & Industrial Engineering* 57(4):1472–1483.

Gitzel, R., B. Schmitz, H. Fromm, A. Isaksson, and T. Setzer. 2016. "Industrial Services as a Research Discipline". *Enterprise Modelling and Information Systems Architectures* 11(1):4–1–22.

Hill, A. V. 1992. "An Experimental Comparison of Dispatching Rules for Field Service Support". *Decision Sciences* 23(1):235–249.

Kovacs, A. A., S. N. Parragh, K. F. Doerner, and R. F. Hartl. 2012. "Adaptive Large Neighborhood Search for Service Technician Routing and Scheduling Problems". *Journal of Scheduling* 15(5):579–600.

Kulkarni, R., and P. Bhave. 1985. "Integer Programming Formulations of Vehicle Routing Problems". *European Journal of Operational Research* 20(1):58–67.

Lahyani, R., M. Khemakhem, and F. Semet. 2015. "Rich Vehicle Routing Problems: From a Taxonomy to a Definition". *European Journal of Operational Research* 241(1):1–14.

Laporte, G. 1992. "The Traveling Salesman Problem: An Overview of Exact and Approximate Algorithms". *European Journal of Operational Research* 59(2):231–247.

Law, A. M. 2015. *Simulation Modeling and Analysis*. 5th ed. New York, NY: McGraw-Hill Education.

Leon, V. J., S. D. Wu, and R. H. Storer. 1994. "Robustness Measures and Robust Scheduling for Job Shops". *IIE Transactions* 26(5):32–43.

Li, Z., and M. Ierapetritou. 2008. "Process Scheduling Under Uncertainty: Review and Challenges". *Computers & Chemical Engineering* 32(4-5):715–727.

Meier, H., H. Lagemann, F. Morlock, and C. Rathmann. 2013. "Key Performance Indicators for Assessing the Planning and Delivery of Industrial Services". *Procedia CIRP* 11:99–104.

Miller, C. E., A. W. Tucker, and R. A. Zemlin. 1960. "Integer Programming Formulation of Traveling Salesman Problems". *Journal of the ACM* 7(4):326–329.

Ostrom, A. L., M. J. Bitner, S. W. Brown, K. A. Burkhard, M. Goul, V. Smith-Daniels, H. Demirkan, and E. Rabinovich. 2010. "Moving Forward and Making a Difference: Research Priorities for the Science of Service". *Journal of Service Research* 13(1):4–36.

Ostrom, A. L., A. Parasuraman, D. E. Bowen, L. Patrício, and C. A. Voss. 2015. "Service Research Priorities in a Rapidly Changing Context". *Journal of Service Research* 18(2):127–159.

Paz, N. M., and W. Leigh. 1994. "Maintenance Scheduling: Issues, Results and Research Needs". *International Journal of Operations & Production Management* 14(8):47–69.

Petrakis, I., C. Hass, and M. Bichler. 2012. "On the Impact of Real-time Information on Field Service Scheduling". *Decision Support Systems* 53(2):282–293.

Psaraftis, H. N. 1995. "Dynamic Vehicle Routing: Status and Prospects". *Annals of Operations Research* 61(1):143–164.

Raychaudhuri, S. 2008. "Introduction to Monte Carlo Simulation". In *Proceedings of the 2008 Winter Simulation Conference*, edited by S. J. Mason et al., 91–100. Piscataway, New Jersey: IEEE.

Sabuncuoglu, I. 1999. "Rescheduling Frequency in an FMS with Uncertain Processing Times and Unreliable Machines". *Journal of Manufacturing Systems* 18(4):268–282.

Sabuncuoglu, I., and M. Bayız. 2000. "Analysis of Reactive Scheduling Problems in a Job Shop Environment". *European Journal of Operational Research* 126(3):567–586.

Sargent, R. G. 2009. "Verification and Validation of Simulation Models". In *Proceedings of the 2009 Winter Simulation Conference*, edited by M. D. Rossetti et al., 162–176. Piscataway, New Jersey: IEEE.

Souyris, S., C. E. Cortés, F. Ordóñez, and A. Weintraub. 2013. "A Robust Optimization Approach to Dispatching Technicians Under Stochastic Service Times". *Optimization Letters* 7(7):1549–1568.

Sungur, I., F. Ordóñez, and M. Dessouky. 2008. "A Robust Optimization Approach for the Capacitated Vehicle Routing Problem with Demand Uncertainty". *IIE Transactions* 40(5):509–523.

Vössing, M. 2017. "Towards Managing Complexity and Uncertainty in Field Service Technician Planning". In *2017 IEEE 19th Conference on Business Informatics (CBI)*, 24th–26th July, Thessaloniki, Greece, 312–319.

Weigel, D., and B. Cao. 1999. "Applying GIS and OR Techniques to Solve Sears Technician-Dispatching and Home-Delivery Problems". *Interfaces* 29(1):112–130.

Wolff, C., M. Vössing, B. Schmitz, and H. Fromm. 2018. "Towards a Technician Marketplace Using Capacity-Based Pricing". In *Proceedings of the 51th Hawaii International Conference on System Sciences*, January 3rd–6th, Waikoloa Village, Hawaii, 1553–1562.

AUTHOR BIOGRAPHY

CLEMENS WOLFF is a research associate at Karlsruhe Service Research Institue (KSRI) at the Karlsruhe Institute of Technology (KIT), Germany. Currently, Clemens focuses on system-oriented service delivery, a delivery approach aiming at minimizing not only provider, but also customer costs. His email address is clemens.wolff@kit.edu, and his web address is ksri.link/wolff.

MICHAEL VÖSSING is a research associate at Karlsruhe Service Research Institue (KSRI) at the Karlsruhe Institute of Technology (KIT), Germany. He explores how maintenance providers can leverage cyber-physical systems and machine learning for proactive decision making – e.g., to enable availability-based maintenance contracts. His email address is michael.voessing@kit.edu, and his web address is ksri.link/voessing.

NIKLAS KÜHL is a post-doctoral research associate at Karlsruhe Service Research Institue (KSRI) at the Karlsruhe Institute of Technology (KIT), Germany. Niklas' research focus is on analytics, machine learning and artificial intelligence in service systems. His email address is niklas.kuehl@kit.edu, and his web address is ksri.link/kuehl.