



6th International Conference on Industry 4.0 and Smart Manufacturing

# Enhancing Resource Assignment Efficiency in Service Industry: A Predict-then-Optimize Approach with XGBoost

Mehran Nasseri <sup>a,b</sup>, Taha Falatouri <sup>a,b</sup>, Patrick Brandtner <sup>a,b\*</sup>, Farzaneh Darbanian <sup>a,b</sup>, Sina Mirshahi <sup>a,b,c</sup>

<sup>a</sup> University of Applied Sciences Upper Austria, Logistikum

<sup>b</sup> Josef Ressel Centre for Predictive Value Network Intelligence

<sup>c</sup> Department of Informatics and Artificial Intelligence, Faculty of Applied Informatics, Tomas Bata University in Zlin, Czech Republic

---

## Abstract

This paper addresses the critical aspect of resource planning in a service context through an integrated predictive and prescriptive approach. Utilizing real-world data from a company providing repair and maintenance services, we demonstrate the use of an XGBoost model to forecast ad-hoc service demands and subsequently optimize resource assignment using a mathematical model. Our findings show that the prediction evaluation metrics significantly improve, highlighting the superiority of complex machine learning models over baseline models such as Linear Regression. Furthermore, the integration of the prediction into the decision-making process resulted in a 26.4% lower decision error compared to the baseline model. Our research also found that the deviations in prediction and optimal objective function values are not aligned. While the average error for MAE % in prediction is 22.2%, the error for the optimal objective function is much lower, reducing to 5.3%. However, although true for our case, this might not be generalizable. Furthermore, when comparing the baseline model with these results, it is also shown that an improvement in prediction accuracy also improves decision making error. Our results indicate that a combined predict-then-optimize approach outperforms the existing methods in both predictive and prescriptive performance, demonstrating its applicability in real-world scenarios.

© 2025 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 6th International Conference on Industry 4.0 and Smart Manufacturing

*Keywords:* Resource planning, predictive analytics, optimization, XGBoost, predict-then-optimize

---

\* Corresponding author. Tel.: +43 5 0804 33450; fax: +43 5 0804 33299.

E-mail address: [patrick.brandtner@fh-steyr.at](mailto:patrick.brandtner@fh-steyr.at)

## 1. Introduction

Decision making is fundamental to a manager's role, significantly influencing the organization's strategic direction, operational effectiveness, and overall success. Managers must decide on various matters, such as purchasing procedures, supply chain operations, or human resources, with resource planning being especially critical for the optimal use of assets and personnel [1]. Resource planning represents a crucial element of every business, involving the assignment of tasks to resources to maximize task completion within specified time frames [2]. The dynamics environment of work assignment, including job cancellations, delays, illnesses, and other external factors such as traffic, necessitate a highly responsive and adaptable system [3]. Decision-makers must anticipate these dynamics to make effective and efficient resource assignments. The results of predictive analytics represent an important input to optimize the outcomes of decisions. However, prediction data might be considered doubtful due to the inherent uncertainty in its accuracy [4].

The commonly used approach in practice involves a prediction-first, followed by an optimization paradigm [5]. Supporting this paradigm, researchers aim to improve the accuracy of their predictions. This improvement is achieved through mathematical measures where the prediction error is compared to reality. Mathematical measures are highly effective in theoretical settings, where they simply compare two numbers. However, applying these metrics to real-life decision-making is more complex, as additional parameters influence predictions to enhance business profitability.

There is growing usage of artificial intelligence (AI) in different managerial tasks [6], one of those being resource allocation. Among the AI methods, deep learning models have outperformed traditional methods in various fields, including resource demand forecasting [7]. As a recent example, XGBoost has become a popular model for predicting resource demand [8].

In this paper we are going to address the critical aspect of resource planning in a service context and highlight its significance in the profit of the businesses using an integrated predictive and prescriptive approach in a real-world problem. More precisely, we show how to use XGBoost to forecast ad-hoc service demands, followed by a mathematical model to optimize resource assignment. The remainder of the paper is structured as follows: Section two provides the results of a literature review on the usage of mathematical methods in resource allocation across different domains. Section three introduces the methodological steps and the experiment, detailing the dataset specifics and prediction modelling. Subsequently, section four demonstrates the evaluation of the models and compares their performance and applicability in the context of prediction and resource allocation. Finally, section five concludes the paper and provides an outlook on future research.

## 2. Background

Resource assignment can be defined as determining the optimal allocation and utilization of resources, such as materials and labor, to achieve specific goals, such as minimizing costs or maximizing efficiency [9]. Traditional approaches to resource allocation exhibit several limitations, including performing allocation solely at the activity level, assigning a single resource to a single activity, and lacking the ability to dynamically capture resources or complement individual resource allocations. These traditional methods assume that parameters related to tasks and resources, such as costs, durations, and capacities, are known. However, in many real-life scenarios, these assumptions may not hold true, with some parameters being unknown and uncertain [3]. To address this uncertainty, a common practical approach involves a predict-then-optimize process. Initially, a predictive model is developed, which is then employed to improve decision-making. Second, the required decisions are made assuming that these predictions are accurate. In this approach, the quality of the decisions is mostly dependent on the quality of the prediction inputs. The limitation of this method is that costs are only considered during the optimization phase [5]. On the other hand, the predict-and-optimize approach suggests building a predictive model by directly minimizing the costs associated with the downstream decision-making task. This method primarily relies on incorporating complexity into the loss function, computation process, and generalization ability, which may not be efficient in certain scenarios [10].

To reach the optimized resource allocation different methods have been investigated in literature. In health care services, En-nahli et al. (2015) have proposed a mixed integer linear model to satisfy the availability and preferences of patients, ensure the compatibility of caregivers' skills with patients' needs, and accommodate caregivers' lunch breaks. Their model is based on several constraints: patient must visit once a day, each caregiver must start and end

their tour within their home, each patient is visited by a caregiver with the necessary qualifications and each caregiver has one lunch break during the day. The developed method improved the efficiency for large instances where exact methods are limited by computational time [11]. Heimerl and Kolisch (2010) have used heuristic approaches to assign project work to resources while minimizing costs and ensuring fitting skill levels. The implemented model showed promising computational performance [12]. In the field of preventive maintenance, another study investigated the impact of assigning employees to tasks based on their qualifications. A proactive strategy for allocating and further training multi-skilled workers was proposed with the goal to develop a work schedule that fulfills orders while considering the team's initial competence levels over the planning period. Their method proved valuable for industries where employees have diverse skills and companies must adapt them based on customer requirements [13]. Szwarc et al. (2023) focused on planning the rotational assignment of tasks to a multi-skilled workforce to ensure their competencies remain at the required level [3]. Xiao et al. (2021) developed a novel model, called Enhanced-XGB, to improve the prediction accuracy of XGBoost for complex time series. Their model dynamically adjusts tree weights in XGBoost and incorporates an offset value to fine-tune the results [8].

In the area of the postal resource allocation, another study [14] explored a recommendation model based on recurrent neural networks for human resource allocation. Their model facilitates personnel scheduling by enhancing the basic neural network and integrating cyclic, global, and local models. Sharma and Omair Shafiq (2020) utilized historical retail purchase data to forecast the likelihood of item purchases. They made an ensemble learning model that integrated Random Forests (RFs), Convolutional Neural Networks (CNNs), Extreme Gradient Boosting (XGBoost), with a voting mechanism. The model's effectiveness was measured using accuracy, precision, F1 score, sensitivity, and specificity. Their findings demonstrated that the ensemble models achieved superior performance compared to existing solutions [15].

In conclusion, while traditional resource allocation methods have proven useful in certain scenarios, the growing complexity and uncertainty in real-world tasks have highlighted the need for more dynamic and integrated approaches, such as predict-then-optimize models, which offer greater flexibility and efficiency in addressing these challenges.

### 3. Research methodology and experiments

The aim of this paper is to address the problem of assigning available resources to a number of tasks within a specific time frame. These tasks originate from customer repair needs for their heating appliances, each with a specific time duration and technical complexity that can be performed by technicians with corresponding expertise. These tasks are not deterministic and must be predicted for the upcoming time frame. To achieve this, we propose a PtO (Predict-then-Optimize) approach, as shown in Figure 1. Based on our proposed approach, we first develop a prediction model for forecasting future repair needs for heating appliances, referred to as ad-hoc orders.

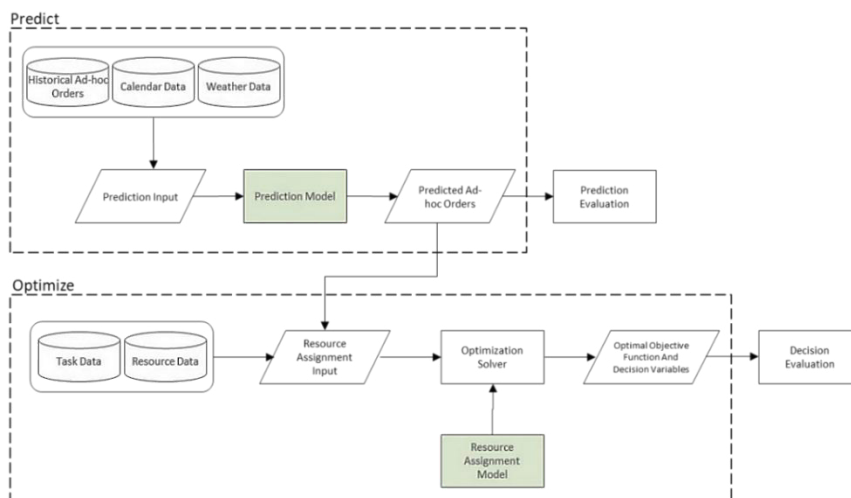


Figure 1: Our proposed PtO approach

Then, we develop a resource assignment model to use the prediction output as input to assign available resources (i.e., technicians with different expertise) to the ad-hoc orders in an optimal manner, considering business constraints.

The data are derived from a real-world use case involving a company that provides repair and maintenance services for heating appliances in private households located in Austria. This section provides a detailed description of the two main parts involved in our approach.

### *Predict*

In the predictive aspect of our approach, our goal is to build an effective prediction model for forecasting ad-hoc orders, as they serve as the main input for our resource assignment model. The output of this predictive model can then be integrated into the assignment model. For this purpose, we utilize various datasets in the prediction model. Our prediction dataset comprises historical daily ad-hoc orders. In total, the dataset spans more than 5 years, providing a solid basis for training and testing. Additionally, we incorporate calendar-related data and weather conditions alongside historical ad-hoc orders. Calendar-related data provide daily information such as the month, week, day of the week, day of the month, and any special events. Weather data include temperature in Celsius, wind speed in m/s, and the amount of precipitation in mm.

Based on the prediction datasets, we extract a set of relevant features to be used in the prediction model. These features were created based on prior research [16, 17], business insights, and experience. Along with the target variable (number of ad-hoc orders), these features comprise the prediction input. The features incorporated into the prediction model are presented in table 1. The prediction input dataset is divided into two parts: the training dataset, which includes approximately 80% of the data and is used for building the models, and the test dataset, which comprises the remaining 20% of the data and is held out for the final evaluation step as unseen data.

Our prediction problem involves a multistep forecast of ad-hoc orders 2 weeks ahead (10 days excluding Saturdays and Sundays) using the features presented in table 1. To achieve this, we adopt the approach proposed by [16] employing the tree-based ensemble model XGBoost, which has demonstrated strong performance in similar problems [17]. This model utilizes an ensemble technique that combines the predictions of a large number of decision trees. In this study, we implemented XGBoost using the Python xgboost library [18].

Table 1 Extracted features for the input of the prediction model

Feature	Type	Source Dataset
Number of repair requests at $t$ ( $t=1,2,\dots,7$ ) days before the current date.	Numeric	Historical Orders
Month of the year	Categorical-Nominal	Calendar Data
Week of the Year	Categorical-Nominal	Calendar Data
Day of the week	Categorical-Nominal	Calendar Data
Indication of a special day or event.	Categorical-Nominal	Calendar Data
Indication of working day or non-working day.	Categorical-Nominal	Calendar Data
Indication of working day or non-working day of the previous day	Categorical-Nominal	Calendar Data
Indication of working day or non-working day of the next day	Categorical-Nominal	Calendar Data
Temperature at $t$ ( $t=1,2,\dots,7$ ) days before the current date.	Numeric	Weather Data
Wind speed at $t$ ( $t=1,2,\dots,7$ ) days before the current date.	Numeric	Weather Data
Amount of precipitation at $t$ ( $t=1,2,\dots,7$ ) days before the current date.	Numeric	Weather Data

### *Optimize*

In the decision-making phase of our approach, we aim to model our problem by developing a resource assignment model. Initially, we proposed it in a general form and then adapted it to a simplified version tailored to our specific use case. This ensures that both the theoretical part of our work, developed with a general model, and the practical side, with fewer constraints and variables, remain applicable to our real-world use case.

Our generalized resource assignment problem is how to assign tasks to resources in order to maximize the number of tasks done in specific time intervals, while ensuring that each resource does not exceed the allowed working hours and that each task is completed by only one resource in a single day.

This problem can be formulated as mathematical optimization model. The primary objective of mathematical optimization is to formulate a mathematical model comprising a set of mathematical relationships for a decision problem and search for its optimal solution. This process typically involves four key elements: Data constants known as parameters, decision variables, constraints, and the objective function [19]. It includes defining decision variables and an objective function, which is to be optimized subject to certain constraints with constant parameters.

We consider the following assumptions, which guide the development of our resource assignment model to ensure it meets both the theoretical and practical requirements of our real-world use case:

- A1: Completing each task requires a specific time and yields a specific amount of profit.
- A2: Assigning at least one task to each resource imposes a fixed cost for using that resource for the day.
- A3: Each resource has a maximum allowable working hours per day, referred to as resource capacity, and tasks assigned to resources each day must not exceed this available capacity.
- A4: Each task can only be completed on specific days using specific resources, due to spare part requirements and the technical complexity that requires resources with corresponding expertise.

We use the following notations for the mathematical formulation of our model:

Sets and indexes:

- $i \in I$  Set of tasks
- $r \in R$  Set of available resources
- $t \in T$  Set of time periods in days

Parameters:

- $M$  A large constant
- $d_{irt}$  Time duration for completing task  $i$  by resource  $r$  on day  $t$  (in minutes)
- $c_{rt}$  Available capacity of resource  $r$  on day  $t$  (in minutes)
- $f_{rt}$  Fixed cost of using resource  $r$  on day  $t$
- $a_{irt}$  A parameter that indicates whether assigning task  $i$  on day  $t$  to resource  $r$  is allowed, with a value of 1 if allowed, and  $M$  otherwise
- $p_{irt}$  Profit from assigning task  $i$  on day  $t$  to resource  $r$ , without considering the fixed cost of using resources

Decisions variables:

- $x_{irt}$  Binary variable which is 1 if task  $i$  on day  $t$  is assigned to resource  $r$ , and 0 otherwise
- $z_{rt}$  Binary variable which is 1 if at least one task on day  $t$  is assigned to resource  $r$

Considering the above notation, our problem can be formulated using the following integer programming (IP) model:

$$\text{Max } Z = \sum_{i \in I} \sum_{r \in R} \sum_{t \in T} p_{irt} x_{irt} - \sum_{r \in R} \sum_{t \in T} f_{rt} z_{rt} \quad (1-1)$$

s. t.

$$\sum_{i \in I} d_{irt} x_{irt} \leq c_{rt} z_{rt} \quad \forall r \in R, t \in T \quad (1-2)$$

$$\sum_{r \in R} a_{irt} x_{irt} \leq 1 \quad \forall i \in I, t \in T \quad (1-3)$$

$$x_{irt} \in \{0,1\} \quad \forall i \in I, r \in R, t \in T \quad (1-4)$$

$$z_{rt} \in \{0,1\} \quad \forall r \in R, t \in T \quad (1-5)$$

The objective function (1-1) maximizes overall profits. The first part of the objective function represents the total profit from completing tasks, and the second part represents the fixed cost of using resources. Constraint (1-2) ensures that the tasks assigned to resources each day do not exceed the available capacity for each resource on that day. Additionally, if at least one task on day  $t$  is assigned to resource  $r$ ,  $z_{rt}$  is set to one, thereby activating the fixed cost in the objective function. Constraint (1-3) ensures that each task is assigned to at most one resource and on one day. The  $a_{irt}$  parameter ensures the allowance of assigning a task to a resource on a specific day by applying 1 when it is allowed and a large constant  $M$  when it is not allowed. Constraints (1-4) and (1-5) define the domains of variables, which are binary. Our proposed model can be considered as an extended version of the generalized assignment problem [20], which is known to be NP-hard. Considering this, to adapt our proposed model for practical application in our real-world scenario, we make the following adjustment to the initial assumptions, reducing the complexity of our problem without compromising practicality:

- Tasks can be classified into a limited number of similar types, each having the same duration and profit regardless of resources and days.
- Each task can be completed by any resource on any day.
- Each resource has the same fixed cost.

With these adjustments, we redefined our notation and formulation with the corresponding values of sets and parameters based on predicted ad-hoc orders, task data, and resource data from our real-world use case as follows:

Sets and indexes:

- $i \in I$  Set of task types,  $I = \{1, 2, 3\}$
- $r \in R$  Set of available resources,  $R = \{1, 2, 3, \dots, 15\}$
- $t \in T$  Set of time periods in days,  $T = \{1, 2, \dots, 10\}$

Parameters:

- $d_i$  Time duration for completing task type  $i$  (in minutes),  $d_1 = 140, d_2 = 110, d_3 = 90$
- $c_{rt}$  Available capacity of resource  $r$  on day  $t$  (in minutes), We consider a standard workweek of 5 days, allocating 510 minutes for regular working days and 360 minutes for Fridays.  $c_{r1} = c_{r2} = c_{r3} = c_{r4} = c_{r6} = c_{r7} = c_{r8} = c_{r9} = 510, c_{r5} = c_{r10} = 360 \forall r \in R = \{1, 2, 3, \dots, 15\}$
- $y_{it}$  Maximum Number of task type  $i$  on day  $t$ , it represents the ad-hoc orders of each task type that is based on the Predicted Ad-hoc Orders.
- $f$  Fixed cost of using resource  $r$  on day  $t$ ,  $f = 30$
- $p_i$  Profit from assigning task type  $i$ , without considering the fixed cost of using resources,  $p_1 = 14, p_2 = 11, p_3 = 9$

Decisions variables:

- $x_{irt}$  Number of task type  $i$  assigned to resource  $r$  on day  $t$
- $z_{rt}$  Binary variable which is 1 if at least one task on day  $t$  is assigned to resource  $r$

Mathematical model:

$$Max Z = \sum_{i \in I} \sum_{r \in R} \sum_{t \in T} p_i x_{irt} - \sum_{r \in R} \sum_{t \in T} f z_{rt} \quad (2-1)$$

s. t.

$$\sum_{i \in I} d_i x_{irt} \leq c_{rt} z_{rt} \quad \forall r \in R, t \in T \quad (2-2)$$

$$\sum_{r \in R} x_{irt} \leq y_{it} \quad \forall i \in I, t \in T \quad (2-3)$$

$$x_{irt} \geq 0 \text{ and integer} \quad \forall i \in I, r \in R, t \in T \quad (2-4)$$

$$z_{rt} \in \{0,1\} \quad \forall r \in R, t \in T \quad (2-5)$$

The objective function (2-1) represents the overall profit, and constraint (2-2) ensures that the tasks assigned to resources each day do not exceed the available capacity for each resource on that day. Constraint (2-3) ensures that the number of assigned tasks for each task type does not exceed its maximum available number. Constraints (2-4) and (2-5) define the domains of variables, specifically changing  $x_{irt}$  from binary to integer. To solve the problem, we used the Python library Pyomo [21] with the solver CBC [22]. Our problem is an Integer Programming problem consisting of 180 constraints and 600 variables, of which 450 are integer variables and 150 are binary variables.

#### 4. Results

In this section, we discuss the results of our approach. We evaluated both the prediction model, referred to as prediction evaluation, and the output of the resource assignment model, referred to as decision evaluation, as shown in figure 1. We tested our method over a period of nine months, running the model at the beginning of each week and advancing one week for the next iteration. This aligns with the company's current approach, where they replan their resources weekly for the following two weeks. For each iteration, we first ran the prediction model to obtain the predicted ad-hoc orders and then fed the prediction output into the resource assignment model to calculate the optimal values. We evaluated both the prediction accuracy and the optimal objective function values based on actual and predicted data for unseen data. The results are shown in table 2.

Table 2: The result of prediction and decision evaluation

Week	Actual		XGBOOST				LR				
	Optimal Objective Value	Optimal Objective Value	% Diff	% MAE	% RMSE	$R^2$	Optimal Objective Value	% Diff	% MAE	% RMSE	$R^2$
1	2,202	2,239	1.7%	21%	27%	88%	2,033	8%	28%	37%	78%
2	2,114	2,272	7.5%	21%	28%	88%	2,179	3%	28%	37%	79%
3	1,883	2,241	19.0%	30%	43%	73%	2,205	17%	31%	43%	72%
4	1,689	2,152	27.4%	34%	48%	66%	2,078	23%	33%	44%	71%
5	1,849	2,165	17.1%	26%	35%	84%	1,869	1%	19%	27%	91%
6	1,806	1,920	6.3%	22%	29%	90%	1,591	12%	30%	58%	62%
7	1,901	1,958	3.0%	24%	31%	88%	1,895	0%	32%	50%	70%
8	2,286	2,219	2.9%	22%	31%	84%	1,738	24%	29%	45%	67%
9	2,298	2,349	2.2%	19%	29%	85%	2,222	3%	25%	35%	78%
10	2,327	2,455	5.5%	16%	21%	92%	2,272	2%	26%	32%	82%
11	2,399	2,421	0.9%	19%	29%	83%	2,249	6%	22%	30%	81%
12	2,377	2,506	5.4%	20%	30%	87%	2,367	0%	27%	38%	79%
13	2,368	2,530	6.8%	23%	31%	88%	2,345	1%	31%	48%	71%
14	2,450	2,559	4.4%	23%	31%	87%	2,353	4%	30%	43%	75%
15	2,613	2,530	3.2%	24%	36%	83%	2,302	12%	42%	60%	53%
16	2,337	2,260	3.3%	25%	40%	82%	2,236	4%	35%	51%	71%
17	2,010	1,970	2.0%	24%	35%	87%	1,997	1%	40%	77%	37%
18	2,331	2,278	2.3%	16%	20%	95%	1,967	16%	34%	52%	66%
19	2,596	2,572	0.9%	18%	23%	93%	2,218	15%	33%	46%	71%
20	2,501	2,639	5.5%	18%	24%	93%	2,611	4%	18%	24%	93%
21	2,551	2,616	2.5%	17%	24%	93%	2,428	5%	23%	33%	88%
22	2,560	2,498	2.4%	25%	34%	85%	2,356	8%	34%	43%	76%
23	2,559	2,507	2.0%	16%	22%	93%	2,454	4%	19%	27%	90%
24	2,548	2,624	3.0%	21%	32%	86%	2,615	3%	32%	44%	74%
25	1,804	1,978	9.6%	41%	58%	67%	1,920	6%	56%	90%	20%
26	1,441	1,613	11.9%	35%	47%	85%	1,559	8%	55%	91%	43%
27	2,167	2,095	3.3%	24%	31%	90%	1,577	27%	42%	71%	48%

28	2,583	2,372	8.2%	20%	26%	90%	2,029	21%	37%	61%	48%
29	2,493	2,484	0.4%	18%	24%	93%	2,525	1%	17%	27%	91%
30	2,435	2,479	1.8%	17%	24%	93%	2,458	1%	18%	26%	92%
31	2,324	2,393	3.0%	20%	28%	91%	2,342	1%	22%	30%	90%
32	2,264	2,387	5.4%	19%	30%	90%	2,409	6%	17%	26%	92%
33	2,221	2,299	3.5%	20%	27%	91%	2,232	0%	19%	24%	93%
34	2,185	2,349	7.5%	25%	32%	87%	2,314	6%	25%	30%	89%
35	2,302	2,302	0.0%	23%	31%	85%	2,165	6%	22%	27%	88%
36	2,376	2,323	2.2%	20%	29%	85%	2,258	5%	21%	29%	85%
37	2,418	2,373	1.9%	14%	20%	93%	2,283	6%	20%	28%	87%
38	2,274	2,357	3.6%	20%	28%	89%	2,379	5%	27%	35%	83%
39	1,970	2,098	6.5%	23%	31%	88%	2,089	6%	36%	61%	56%
<b>Average</b>	<b>2,252</b>	<b>2,317</b>	<b>5.3%</b>	<b>22.2%</b>	<b>30.8%</b>	<b>87.0%</b>	<b>2,183</b>	<b>7.2%</b>	<b>29.1%</b>	<b>43.1%</b>	<b>73.9%</b>
<b>Median</b>	<b>2,324</b>	<b>2,349</b>	<b>3.3%</b>	<b>21.3%</b>	<b>29.6%</b>	<b>88.1%</b>	<b>2,236</b>	<b>5.0%</b>	<b>28.0%</b>	<b>38.1%</b>	<b>77.8%</b>

To evaluate the prediction accuracy, we used the metrics shown in table 3. It's worth noting that metrics like MAPE are not suitable for our analysis because the ad-hoc orders can include zero values in our data. To ensure consistency in scale and to facilitate comparison between different models, we divided the MAE and RMSE by the average actual value, using their percentages as evaluation metrics.

Table 3: Prediction evaluation metrics

Mean Absolute Error % (MAE %)	Root Mean Square Error (RMSE%)	Coefficient of Determination ( $R^2$ )
$MAE \% = \frac{1}{n} \sum_{t=1}^n \frac{ \hat{y}_t - y_t }{\bar{y}}$	$RMSE \% = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2}}{\bar{y}}$	$R^2 = 1 - \frac{\sum_{t=1}^n (\hat{y}_t - y_t)^2}{\sum_{t=1}^n (y_t - \bar{y})^2}$

\*Where  $y_t$  and  $\hat{y}_t$  are the actual and predicted ad-hoc order for the time interval (day) t, respectively.

For decision evaluation, we compared the percentage difference between the predicted and actual optimal values. To better understand the performance of our model, we also included the results of the baseline Linear Regression model (represented as "LR" in the table) model, which is the current approach the organization is using.

The prediction evaluation metrics showed a significant improvement compared to this baseline model. On average, MAE % improved from 29.1% to 22.2% (a 23.7% improvement), RMSE % improved from 43.1% to 30.8% (a 28.5% improvement), and  $R^2$  improved from 73.9% to 87.0% (a 17.7% improvement). This demonstrates the power of complex machine learning models over traditional statistical models in the given use case.

However, the evaluation of the prediction alone does not show if there is also an improvement in resource planning and decision-making. Its importance becomes evident when it is used in the decision-making process. Since the deviation from the objective function is crucial for the company's budget planning, we compared optimal objective function values for predicted and actual inputs. As shown in table 2, the average deviation for the objective function is 5.3% for our prediction approach, compared to 7.2% for the baseline model. This significant improvement of 26.4% demonstrates that our model was able to enhance both predictive and prescriptive performance.

Another finding of our research is that the deviations in prediction and optimal functional values are not aligned. While the average error for MAE % in prediction is 22.2%, the error for the optimal objective function is much lower, reducing to 5.3%. However, although true for our case, this might not be generalizable [5]. Furthermore, when comparing the baseline model (prediction MAE % of 29.1% and optimal objective function error of 7.2%) with these results, it is shown that an improvement in prediction accuracy also improves decision making error.

The managerial insights derived from the results highlight the significant advantages of integrating predictive analytics with optimization in resource planning. By using the predict-then-optimize approach with the XGBoost model, managers can make more informed decisions, resulting in a 26.4% reduction in decision errors compared to traditional methods. This indicates that improving prediction accuracy directly enhances operational efficiency and resource allocation, allowing businesses to optimize labor costs and maximize task completion. The results suggest that companies, particularly in service industries, can benefit from adopting machine learning models for forecasting demand and dynamically adjusting resource allocation, ultimately leading to increased profitability and better service delivery. Furthermore, managers should be aware that prediction errors do not necessarily translate into poor decisions,

as optimization models can compensate for some level of inaccuracy. This provides confidence in applying advanced analytics to decision-making processes even in uncertain environments, as shown in previous research [23, 24].

## 5. Discussion and Conclusion

In this research, we addressed a real-world problem by leveraging both predictive and prescriptive approaches. First, we formulated our use case into a mathematical model that could be solved efficiently while maintaining practicality. Then, we used predictive methods to forecast the uncertain parameter of our model, specifically ad-hoc orders. We evaluated the results of both predictive and decision-making performance and compared them to the current approaches used by the company. Our results indicate that our approach outperforms the existing methods in both predictive and prescriptive performance, demonstrating its applicability in real-world scenarios. From a practitioner perspective, the developed predict-then-optimize approach has been developed and tested with real-world data from a company providing repair services, showing its scalability and effectiveness in dynamic environments where demand is uncertain and fluctuating. Practical benefits include the ability to optimize labor costs and time allocation for technicians, helping businesses operate more profitably while maintaining service quality. From a theoretical perspective, the research contributes to the body of knowledge by showing that better prediction accuracy directly impacts the quality of decision-making in resource allocation, evidenced by the 26.4% reduction in decision error when compared to baseline models.

However, our research has some limitations. One limitation is that the prediction model aims to forecast ad-hoc orders without considering the decision-making objectives. While our results show an acceptable level of performance, this approach might not always be optimal. Another limitation is that our findings may not be directly applicable to other domains or datasets, as the performance and characteristics of models can vary significantly based on the specific data used. Additionally, there are many prediction techniques and algorithms that could produce different outcomes compared to our approach, so this study might not provide a comprehensive understanding of the optimal model for our specific use case.

Future research could explore several promising areas to enhance the effectiveness of resource planning and decision-making. Firstly, applying different prediction models and comparing their performance could identify the model that yields the best results. Secondly, investigating the robustness of the relationship between prediction output and optimization output using techniques like robust optimization could help determine the prediction error ranges within which the optimization model remains effective. Additionally, extending our approach to other resource assignment problems with varying objective functions or constraints could test its generalizability. Finally, developing a method to integrate predictive and optimization models by incorporating resource planning into the predictive model through a loss function could simultaneously optimize both prediction and decision-making processes.

## Acknowledgements

This research was funded by the Christian Doppler Research Association as part of Josef Ressel Centre for Predictive Value Network Intelligence (JRC PREVAIL).

## References

- [1] Maritan CA, Lee GK. (2017) "Resource Allocation and Strategy". *Journal of Management* 2017; 43(8): 2411–20 [https://doi.org/10.1177/0149206317729738]
- [2] Müller R, Vette M, Mailahn O. (2016) „Process-oriented Task Assignment for Assembly Processes with Human-robot Interaction”. *Procedia Cirp* 2016; 44: 210–5 [https://doi.org/10.1016/j.procir.2016.02.080]
- [3] Szwarz E, Bocewicz G, Golińska-Dawson P, Banaszak Z. (2023) "Proactive Operations Management: Staff Allocation with Competence Maintenance Constraints". *Sustainability* 2023; 15(3): 1949 [https://doi.org/10.3390/su15031949]
- [4] Ingle C, Bakliwal D, Jain J, Singh P, Kale P, Chhajed V. (2021) "Demand Forecasting : Literature Review On Various Methodologies. In: Demand Forecasting". *Literature Review On Various Methodologies; 2021. IEEE*; 1–7.
- [5] Elmachtoub AN, Grigas P. Smart (2022). "Predict, then Optimize". *Management Science* 2022; 68(1): 9–26 [https://doi.org/10.1287/mnsc.2020.3922]

- [6] Darbanian F, Brandtner P, Falatouri T, Nasseri M. (2024) “Data Analytics in Supply Chain Management: A State-of-the-Art Literature Review”. *OSCM: An Int. Journal* 2024; 1–31 [https://doi.org/10.31387/oscm0560411]
- [7] Lim B, Zohren S. (2020) “Time-series forecasting with deep learning: a survey”. *Philos Trans A Math Phys Eng Sci* 2021; 379(2194): 20200209 [https://doi.org/10.1098/rsta.2020.0209][PMID: 33583273]
- [8] Xiao C, Huang J, Wu W, Yin Y, Chang H. (2021) “Enhanced-XGB: An Online Service Resource Demand Forecasting Method for Colocation Data Centers”. In: *Enhanced-XGB: An Online Service Resource Demand Forecasting Method for Colocation Data Centers*; 2021. IEEE; 436–44.
- [9] Taha HA. (2007) “Operations research: An introduction”. 8th ed. Upper Saddle River N.J.: Pearson/Prentice Hall 2007.
- [10] Vanderschuere T, Verdonck T, Baesens B, Verbeke W. (2022) „Predict-then-optimize or predict-and-optimize? An empirical evaluation of cost-sensitive learning strategies.” *Information sciences* 2022; 594: 400–15 [https://doi.org/10.1016/j.ins.2022.02.021]
- [11] En-nahli L, Allaoui H, Nouaouri I. (2015) “A Multi-objective Modelling to Human Resource Assignment and Routing Problem for Home Health Care Services”. *IFAC-PapersOnLine* 2015; 48(3): 698–703 [https://doi.org/10.1016/j.ifacol.2015.06.164]
- [12] Heimerl C, Kolisch R. (2010) “Work assignment to and qualification of multi-skilled human resources under knowledge depreciation and company skill level targets”. *International journal of production research* 2010; 48(13): 3759–81 [https://doi.org/10.1080/00207540902852785]
- [13] Bocewicz G, Golińska-Dawson P, Szwarc E, Banaszak Z. (2023) “Preventive maintenance scheduling of a multi-skilled human resource-constrained project’s portfolio”. *Engineering Applications of Artificial Intelligence* 2023; 119: 105725 [https://doi.org/10.1016/j.engappai.2022.105725]
- [14] Shi W, Li Q. (2021) “Human Resources Balanced Allocation Method Based on Deep Learning Algorithm”. *Scientific Programming* 2021; 2021: 1–9 [https://doi.org/10.1155/2021/4681959]
- [15] Sharma A, Omair Shafiq M. (2020) “Predicting purchase probability of retail items using an ensemble learning approach and historical data”. In: *Predicting purchase probability of retail items using an ensemble learning approach and historical data*; 2020. IEEE; 723–8.
- [16] Falatouri T, Brandtner P, Nasseri M, Darbanian F. (2023) “Maintenance Forecasting Model for Geographically Distributed Home Appliances Using Spatial-Temporal Networks”. *Procedia Computer Science* 2023; 219: 495–503 [https://doi.org/10.1016/j.procs.2023.01.317]
- [17] Nasseri M, Falatouri T, Brandtner P, Darbanian F. (2023) “Applying Machine Learning in Retail Demand Prediction—A Comparison of Tree-Based Ensembles and Long Short-Term Memory-Based Deep Learning”. *Applied Sciences* 2023; 13(19): 11112 [https://doi.org/10.3390/app131911112]
- [18] Chen T, Guestrin C. (2016) “XGBoost”. In: Krishnapuram B, Shah M, Smola A, Aggarwal C, Shen D, Rastogi R, editors. *XGBoost*; 2016. New York, NY, USA: ACM; 785–94.
- [19] Pardalos PM, Hearn DW, Kallrath J. (2004) “Modeling Languages in Mathematical Optimization”. Boston, MA: Springer US 2004.
- [20] Pentico DW. (2007) “Assignment problems: A golden anniversary survey”. *European Journal of Operational Research* 2007; 176(2): 774–93 [https://doi.org/10.1016/j.ejor.2005.09.014]
- [21] Hart WE, Laird CD, Watson J-P, et al. (2017) “Pyomo — Optimization Modeling in Python”. Cham: Springer International Publishing 2017.
- [22] Forrest J, Lougee-Heimer R. (2005) “CBC User Guide”. In: Greenberg HJ, Smith JC, editors. *Emerging Theory, Methods, and Applications*. INFORMS 2005; 257–77.
- [23] Udokwu, C., Brandtner, P., Darbanian, F., & Falatouri, T. (2022). Proposals for addressing research gaps at the intersection of data analytics and supply chain management. *Journal of Advances in Information Technology*, 13(4).
- [24] Brandtner, P. (2024). Digital Transformation of Supply Chain Management-Challenges and Strategies for Successfully Implementing Data Analytics in Practice. In *Proceedings of the 2024 8th International Conference on E-Commerce, E-Business, and E-Government* (pp. 36-42).