# Man-Machine Teaming: AI's Overload Management and Task Allocation

Weifei Xie[1,2,3][a], Cédric Herpson[1], Christophe Cambier[2] and Jean-Marie Dembelle[3]

[1] *Sorbonne Universite, CNRS, LIP6, F-75005 Paris, France*

[2] *Sorbonne Universite, IRD, UCAD,UGB UMI UMMISCO, F-75006 Paris, France*

[3] *Universite Gaston Berger de Saint-Louis, Laboratoire d'Analyse Numérique et Informatique, BP. 234 Saint Louis, Senegal*

*fi fi fi*

Keywords: Task Planning, Mental Workload.

Abstract: Sustained concentration can induce cognitive overload in demanding roles such as surgeons and soldiers, prompting research into workload management via surveys, performance metrics, and physiological measures. However, these conventional methods face challenges like time-intensive survey processes, unreliable metrics, and equipment dependency, complicating their application in projects like Recolte, which develops partially autonomous drone fleets for data collection.To overcome these challenges, we designed an AI-Human task distribution algorithm that assesses mental workload considering task complexity, duration, and operator competency. This algorithm also incorporates factors influencing recovery speed, including rest type, operator state, and environmental conditions. We assessed four recovery models to determine their suitability for our use case, employing AI for routine tasks to mitigate human workload. Our algorithm aims to optimize team composition by determining the optimal timing for integrating additional human or advanced AI resources to ensure mission success. Empirical findings from this study provide insights into the recovery models' impact on operational effectiveness, facilitating the analysis of success rates across different task configurations and operator settings. This method ensures continuous human training, even with the presence of AI operators for specific tasks.



## 1 INTRODUCTION

Maintaining a high level of concentration for extended periods can lead to cognitive overload, whether for a surgeon, a soldier in combat, or a driver. Cognitive overload can have a negative impact on performance, employee productivity, and may even lead to mission failure (Li et al., 2020). To prevent overload and optimize performances, it's necessary to monitor or estimate worker's mental workload and to plan recovery periods. In the Recolte project, we aim to develop a fleet of semi-autonomous drones for long-term data collection using different platforms and both artificial and human pilots. In this context, our goal in this paper is to measure the impact of the various recovery models on a mission and to study the evolution of a mission success rate according to specific tasks and operators configurations.

In the literature, methods for monitoring mental workload can be divided into three categories: subjective measurement using surveys, physiological measurement using equipment and performance measurement. It is important to note that these categories are not mutually exclusive. One disadvantage of surveys is their limited applicability in continuous missions. Physiological measurements require specialized equipment not easily usable during a mission. Performance measurement can be further divided into primary task measurement and secondary task measurement. Both have several disadvantages, including insensitivity to variations in mental load, sensitivity to differences between individuals, unreliability of the secondary task limit and the addition of extra mental load. The obstacles mentioned make it difficult to directly apply these methods for long-term continuous missions. Rather than monitoring mental workload, we thus choose to estimate it using mental workload and recovery models combined with a proposed task allocation algorithm. Our algorithm also considers the amount of training required for each mission to be successful, which is not always a concern in traditional work-rest balancing planning. Additionally, our task allocation algorithm is able to test whether it is possible to execute the mission without cognitive overload, using both human and AI pilots.

[a] https://orcid.org/0009-0008-8866-9855



Xie, W., Herpson, C., Cambier, C. and Dembelle, J.
Man-Machine Teaming: AI's Overload Management and Task Allocation.
DOI: 10.5220/0012984900004562

In *Proceedings of the 2nd International Conference on Cognitive Aircraft Systems (ICCAS 2024)*, pages 95-103
ISBN: 978-989-758-724-5

In section 2 we present the general literature of mental workloads and recovery models as well as associated scheduling algorithms. As workload can be influenced by various factors, section 3 presents four mental workload and recovery models grounded in different strand of the literature and adapted to our application context. We then present our scheduling algorithm that combines both workload and recovery criteria in section 4. In section 5 we discuss experimental results comparing the behaviours of the different models before concluding.

# 2 STATE OF THE ART

Xie and Salvendy (2000) suggests that the evaluation of mental workload should consider several aspects, including instantaneous workload, peak workload, average workload, accumulated workload, and overall workload. Instantaneous workload reflects workload fluctuations and is typically measured through physiological metrics. Peak workload, which represents the highest level of instantaneous workload, is critical for identifying potential mental overload. Accumulated workload is the total workload experienced during a task, calculated from instantaneous workload measurements. Average workload is the mean of all instantaneous workload values. Overall workload can be assessed using a mapping function based on instantaneous workload or another function that considers both average and accumulated workload. These mapping functions consider the complexity of the task and individual differences, providing a comprehensive view of the mental demands placed on professionals in high-stakes environments.

$$W_{peak} = MAX(W_i(t)) \tag{1}$$

$$W_{accumulated}(t) = \int_0^t W_i(u)du \tag{2}$$

$$W_{average} = W_{accumulated}(t)/t \tag{3}$$

(4)

$$W_{overall} = f_1(W_i(t)) = f_2(W_{average}(t), W_{accumulated}(t)) \tag{5}$$

In the context of multitasking, it is assumed that the cumulative workload associated with simultaneous tasks is additive in nature (Xie and Salvendy, 2000), as represented by the equation:

$$W_{\text{overall-multi}} = \sum_{i=0}^{n} W_{\text{overall-i}} \tag{6}$$

This formulation, however, does not account for the potential mitigating effects of interspersed breaks between tasks. Extensive research has demonstrated that breaks can significantly facilitate recovery from high workload levels. Consequently, a revised model incorporating the impact of breaks is proposed:

$$W_{\text{overall}} = \sum_{i=0}^{n} W_{\text{overall-i}} + \sum_{j=0}^{m} R_{\text{break-j}} \tag{7}$$

In this revised model, $R_{\text{break-j}}$ represents the reduction in mental workload attributed to the $j$th break. The efficacy of these breaks in reducing mental workload is influenced by various factors, including the nature of the relaxation activity undertaken during breaks, the surrounding environment, and the duration of the break, among others.

For our project, "Recolte," the practical challenges of deploying physiological measurement equipment for each pilot to directly assess workload necessitate an alternative approach. As such, we aim to approximate the workload through an estimation based on the workload associated with individual tasks.

## 2.1 Mental Workload Accumulation

In the study conducted by Xie and Salvendy (2000), it was articulated that the mental workload is predominantly influenced by the task at hand and the individual involved. Further elaboration by Hancock et al. (2021) elucidated that the mental workload is modulated by the chosen work procedure, which is influenced by four pivotal factors: the nature and complexity of the task, the degree of autonomy and the scope for decision-making allowed by the situational context, the operator's level of expertise, and the operator's current mental state. In the cognitive domain, processing can be categorized into three distinct levels (Drenth, 1998):

1. Skill-Based Processing: Automatic and highly practiced actions requiring minimal thought. They are fast and efficient but susceptible to attention-related errors.
2. Rule-Based Processing: Applies learned rules to somewhat familiar situations. They are more thoughtful than skill-based; useful for semi-routine situations.
3. Knowledge-Based Processing Engages in novel situations requiring deep thought and problem-solving. They are highly flexible and powerful but slow and demands significant mental resources.

We can thus conclude that with different expertise the same task can conduct to different workload.

### 2.2 Mental Workload Recovery

Numerous studies have demonstrated that incorporating breaks into the work schedule can significantly enhance both employee health and performance. This practice can lead to increased productivity with reduced investment, ultimately benefiting organizational efficiency. Albulescu et al. (2022) discusses how work demands deplete psychological resources, which can be replenished during periods of rest such as sleep, weekends, and vacations. The author also notes that recovery can occur during shorter periods of downtime, including lunch breaks, work breaks, and even brief micro-breaks. Caldwell et al. (2009) highlights napping as a highly effective non-pharmacological method for enhancing alertness in flight pilots. The research also points out the time required for falling asleep and waking up within designated nap durations, citing a NASA study which determined that an actual sleep time of 26 minutes out of a 40-minute nap period is typical. This observation indicates that napping may not be practical for micro-breaks because of the associated time constraints, thus rendering it unsuitable for short breaks. Additionally, research by Brazaitis and Satas (2023) demonstrates that a 10-minute break following 50 minutes of work does not sufficiently mitigate cognitive fatigue, suggesting that such intervals are inadequate for preventing cognitive decline during longer tasks. Albulescu et al. (2022) demonstrated that shorter, more frequent breaks can effectively reduce fatigue levels without adversely affecting productivity.

Tucker (2003) suggests that the type of relaxation activity undertaken during a break can influence the extent of recovery achieved. For instance, engaging in activities such as taking a short nap (less than 15 minutes) or drinking coffee may be more effective at combating sleepiness compared to simply taking a break. Hoover et al. (2022) conducted an experimental comparison of two types of break activities: physical (stationary biking) and relaxation (progressive muscle relaxation). Their findings indicate that relaxation activities are more effective in promoting psychological detachment and relaxation, while physical activities are superior in replenishing energy levels.

### 2.3 Work-Rest Balance Scheduling

Another critical approach to preventing overload involves strategic task scheduling. It is essential to develop a plan that avoids overload while simultaneously not compromising work efficiency. In Jaber and Neumann (2010) introduced a mixed-integer linear programming (MILP) model designed to address task scheduling issues, incorporating the impact of human fatigue into the modeling process. The goal of this MILP model is to enhance productivity while reducing the physical strain on workers. However, solving such a MILP problem remains a challenging task. Zhang et al. (2022) studied task scheduling in a human-robot collaborative assembly cell with the goal of balancing job cycle time and human fatigue. They developed an MILP model for this issue, using job cycle time as the objective function and imposing a constraint on maximum human fatigue. To address the complexity of solving this MILP model, they suggested the use of a genetic algorithm. But to our knowledge there is currently no task scheduling algorithm taking into account both task difficulty and the AI ability for the scheduling. Studies such as Zhang et al. (2022) and Jaber and Neumann (2010) propose that specific types of work consistently impact humans in the same manner; however, their approach does not align with the Project Recolte. For example, in this project, the complexity of aerial tasks is influenced by factors like wind conditions, the presence of obstacles, and the number of turns required. Furthermore, whereas other studies may prescribe fixed tasks for robots, the Recolte project focuses on adapting tasks based on their difficulty rather than defining them by type.

To meet this need, we propose an algorithm that takes into account the type of rest activity, operator capabilities and task complexity. Our algorithm also must also take into account the fact that the intervention of AI shoud not negatively affect human learning.

## 3 PROPOSED MENTAL WORKLOAD ACCUMULATION AND RECOVERY MODELS

Drenth (1998) highlighted that mental workload accumulation is affected by the environment, the task, and the operator. In the context of Recolte, where the operator's environment remains relatively constant during the mission, the primary factors influencing mental workload accumulation are the task and the operator. Consequently, we categorize the flying task into three levels of difficulty—easy, medium, and difficult—based on criteria such as the number of turns relative to the task area, the presence of a guiding routine, and whether the map is known or unknown. In addressing the operator dimension of Recolte's mission to make drones widely accessible, we recognize the significant variability in operator expertise. Accordingly, we classify operators into three distinct categories based on their skill level: novice, intermediate, and advanced. This stratification allows us to appropriately match task difficulty with operator capability, facilitating both efficient and effective drone operation.

## 3.1 Mental Workload Accumulation Model

We propose a mathematical model to quantify the mental workload induced by task $i$ on operator $j$ as follows:

$$mw_i = duration_i * coef_{ij} \quad (8)$$

- $mw_i$ represents the mental workload generated by task $i$ for operator $j$.
- $duration_i$ denotes the estimated duration of task $i$
- $coef_{ij}$ is a predefined coefficient that reflects the difficulty of task $i$ adjusted for the skill level of operator $j$. This coefficient ensures that the level of the operator is appropriately matched to the task difficulty, such as assigning only difficult tasks to advanced operators.

## 3.2 Mental Workload Recovery Model

As introduced in section 2, recovery from mental workload can be influenced by various factors including the operator's current state, the type of relaxation activity engaged in, and the duration of the recovery period. Based on this understanding, we developed four models, each grounded in different strands of literature:

1. Linear Model: This model posits a direct, proportional relationship between the intensity and duration of the workload and the required recovery time. It is based on the principle that recovery time increases linearly with increases in the cognitive demands of the task. This model can be especially useful for tasks with predictable and consistent cognitive loads.(Asadayoobi et al., 2023) The model is defined by the equation $R = \varepsilon * duration$ where $\varepsilon$ denotes the recovery speed, $duration$ is the length of the break , $R$ measures the recovery of mental workload.
2. Linear Model With Delays: Inspired by Caldwell et al. (2009), our model recognizes that the body requires time to initiate the recovery process, such as the time needed to fall asleep when using napping as a recovery activity.The model is defined by the equation $R = \varepsilon * (duration - delay)$ if $duration > delay$ or $R = 0$ if $duration <= delay$ ($delay$ is a predefined constant).
3. Exponential Model: An exponential model influenced by pre-break mental workload.(Jaber and Neumann, 2010) The model is defined by the equation $R = F * exp^{\mu * duration}$, $F$ is pre-break mental workload level.$\mu$ is constant.
4. Quadratic Model: Albulescu et al. (2022)suggested that micro-breaks might be more efficient than extended breaks. In response, we propose a quadratic model $at^2 + bt + c$ to explore this relationship. Here, $a, b, c$ are coefficients that shape the recovery function. We set $a$ negative to model scenarios where shorter breaks yield better recovery, implying a decrease in recovery benefit as break duration increases. Conversely, to examine conditions where longer breaks could be more advantageous, we make $a$ positive, indicating increased recovery efficiency with longer break durations.

# 4 PROPOSED SCHEDULING ALGORITHM

Having developed models for both mental workload accumulation and recovery, it is crucial to consider Recolte's objective of enabling novice operators to effectively execute missions. To achieve this, integrating artificial intelligence (AI) into the drone systems to provide necessary assistance becomes essential.

For safety concerns, our algorithm should be capable of assessing the feasibility of missions given the human operator's expertise and the support of AI. Specifically, the algorithm needs to evaluate:

1. Mission Feasibility: Assess if the human operator can complete the mission with AI support without suffering cognitive overload.
2. Risk of Overload: Identify potential mental overload risks during the mission.
3. Recovery Model Selection: Recommend an appropriate recovery model based on observed stress levels and task nature, choosing suitable relaxation activities that align with the mission timeline and operator needs.
4. Dynamic Adjustments: Enable the system to dynamically adjust risk assessments and recovery plans as mission parameters change, such as variations in task complexity or assignment.

The algorithm should also have the capability to train human operators, for example, by enabling them to execute tasks proactively while minimizing the risk of overload. From there we propose Algorithm (1)

: *Distribution* [1] below and introduce the associated notations :

- $W^s_{ij}$: the mental workload at the start of task $i$ of the operator $j$
- $W^e_{ij}$: the mental workload at the end of task $i$ of the operator $j$
- $Wn_{ij}$: the mental workload will be needed for complete the rest task attribute from task $i$ for the operator $j$. If there is no task assign after task $i$, $Wn_{ij} = W_{ij}$. If there are tasks assigned to operator $j$ after task $i$, identify the task $c$ that is the closest subsequent task to task $i$, then $Wn_{ij} = W_{ij} + Wn_{cj}$
- $n_d$ number of preset difficulty levels
- $d_i$ difficulty of task $i$(easy, medium, and difficult for 1,2,3)
- $l_j$ level of operator $j$ (novice, intermediate, and advanced for 1,2,3)
- $W_{ij}$ mental workload generate by task $i$ for the operator $j$
- $Wmax_j$ the maximum mental workload an operator can tolerate without negatively influencing their performance

In this algorithm, we initially assign tasks to operators such that the difficulty of each task corresponds with the operator's level of expertise—easy tasks to novice operators, medium tasks to intermediate operators, and difficult tasks to advanced operators, ensuring not to exceed their maximum workload capacities. For tasks that remain unassigned, we proceed to allocate these to higher-level operators. However, such assignments must comply with the specified equation $Wmax_a - W^s_{ta} > Wn_{ta}$(where the $a$ present operator, $t$present the task) to ensure that the addition of any new task does not cause the operator to surpass their tolerable workload threshold. This approach guarantees that all tasks are appropriately assigned while maintaining operational efficiency and preventing operator overload.

Algorithm 1: Distribution.

**Input:** $agentPool\ taskPool$, $recovryModel$
**Output:** $tabDistribution, riskOverload$
divide both the agent pool and the task pool based on their respective levels and difficulty such that task level i is in $taskPool[i]$
$riskOverload = False$
**for** i in range$n_d$ **do**
  **for** task t in $taskPool[i]$ **do**
    **for** agent a in $agentPool[i]$ **do**
      calculate $W^s_{ta}$
      if $W^s_{ta} + W_{ta} < Wmax_j$ assign the task to agent a, delete the task in $taskPool[i]$, update $W^e_{ta}$ break
    **end for**
  **end for**
**end for**
**for** i in range$n_d$ **do**
  update workload necessary list for agent in $agentPool[i]$
**end for**
**for** i in range$n_d$ **do**
  **for** task t in $taskPool[i]$ **do**
    **for** j in range$i+1, n_d$ **do**
      **for** agent a in $agentPool[j]$ **do**
        calculate $W^s_{ta}$
        calculate $Wn_{ta}$
        **if then**$Wmax_a - W^s_{ta} > Wn_{ta}$
          assigning the task to agent a, delete the task in $taskPool[i]$, update $W^e_{ta}$ break
        **end if**
      **end for**
    **end for**
  **end for**
  **for** i in range$n_d$ **do**
    update workload necessary list for agent in $agentPool[i]$
  **end for**
**end for**
**for** i in range$n_d$ **do**
  if exist task in $taskPool[i]$ then $riskOverload = True$
**end for**
**return** $tabDistribution, riskOverload$

# 5 EXPERIENCE

## 5.1 Common Parameter

In this section we explore our model through a case study inspired by the Project *Recolte*. *Recolte* aims to develop a fleet of partially autonomous solar-powered drones designed for prolonged data collection involving a combination of various platforms and human pilots. Specifically, we analyze the performance of a drone undertaking 100 consecutive flying tasks, each lasting approximately 15 minutes without interruption. The difficulty of each task is randomly assigned in each trial according to a predefined probability distribution.

For the operator component, we have examined

[1] Source code : https://github.com/xiewf2019/distribution

various configurations to address the following questions with application of our scheduling algorithm:

- What are the benefits of integrating an AI, even at a novice level, compared to relying solely on a human operator?
- Which recovery model is most suitable for this scenario?
- If an AI and a human operator with equivalent skill levels are both involved, will the human operator receive sufficient training?

In the calculations that follow, the units utilized within the accumulation and recovery model are expressed as percentages. For instance, if the operator accumulates a value of 50, this indicates that they have reached 50% of their maximum capacity. For the mental workload accumulation coefficients, inspired by the work in Drenth (1998), we have considered scenarios where the operator's skill level is lower than the task difficulty, rendering the operator unable to perform the task. In USA Air Carrier Operations Survey, they mentioned that after 2 hour operation needs a break, so we consider that for a novice pilot we need break after 1 hour operation. The coefficients are determined based on the indices $i$ and $j$, as specified below.

Table 1: Mental workload accumulation coefficients based on human level $j$and task difficulty$i$.

| i \ j | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 1.5 | 1.125 | 0.75 |
| 2 | - | 2 | 1.5 |
| 3 | - | - | 3 |

With this coefficient we can assumed that after one hour operation the novice operator will be consider as overload. Building on the research presented in Hoover et al. (2022), and recognizing that drone flight involves greater stress than the e-tasks referenced in Hoover et al. (2022),according their result they consider that 15 minute break is enough for 30 minute e-task work, as drone flight is more stress than e-task so we consider that we need 45 minute to recover for 30 minute flight we have accordingly established the numerical values for the recovery models:

- $\mu$ for exponential model: 0.096 , we take the value of work Jaber and Neumann (2010) (model 7 in experience)
- $a, b, c$ for quadratic model: $1/15, 0, 0$ (model 1 in experience) , $1/30, 0.5, 0$ (model 2 in experience), $-1/30, 1.5, 0$ (model 4 in experience), $-1/15, 2, 0$ (model 5 in experience)
- *delay* for delay model: 3 (model 6 in experience)
- $\varepsilon$ for linear model: 1 (model 3 in experience)

For the aspect of mental workload capacity, unlike other systems, we assume that everyone reaching their limit will be overloaded. We acknowledge that maintaining optimal capacity can be challenging, and for safety reasons, we consider a person to be overloaded if they reach 80% of their capacity. So we set all the operator has 80 as mental workload capacity.

The simulation creates a scenario where operators completes 100 consecutive tasks. In this model, each task is performed by an operator whose capacity must not be less than the requirements of the task. If the assigned operator for a task is at risk of overload, i.e., the sum of the cumulative mental workload and the expected workload exceeds a predetermined threshold and no other operator is available to replace him or no available operator is capable of completing the task, the task is considered to be a failure due to the potential risk of overload. To avoid the effects of randomness and to assess the impact of task difficulty distributions on the results, 100 tasks with specified difficulty distributions were generated. The algorithm was then applied to each set of tasks to assess the risk of overload. Each distribution was repeated 100 times to ensure that it was not affected by random variations.

## 5.2 Experience and Result

To answer aux questions, in the initial simulation, we compared three configurations:

- An advanced-skilled individual paired with a novice.
- An advanced-skilled individual paired with a novice AI.
- An advanced-skilled individual, a beginner individual, and a beginner AI.

The first two configurations assess the impact of replacing a human with an AI, while the third configuration explores the effects of introducing an AI into the scenario. In the referenced figure (1), the y-axis

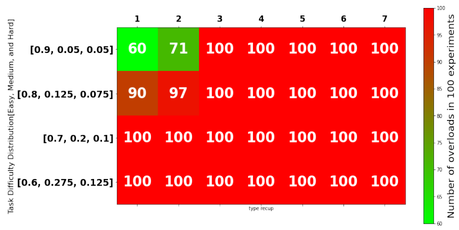


Figure 1: The result with 100 simulation, one human operator level advanced, one human level novice.

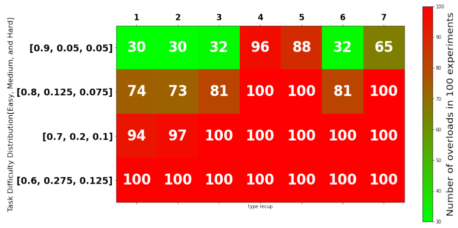


Figure 2: The result with 100 simulation, one human operator level advanced, one AI level novice.

represents the probability distribution across different levels of task difficulty, categorized as novice, intermediate, and advanced. For instance, the distribution [0.9, 0.05, 0.05] signifies a 90% probability of the task being classified as novice level, a 5% probability for intermediate, and a 5% probability for advanced level. The x-axis delineates various types of recovery models. The numbers displayed on the grid indicate the frequency, out of 100 simulations, at which the operator is unable to manage the tasks, signifying a risk of overload. For example, a value of 54 means that in 100 simulations, there were 54 instances where the risk of overload was present.

To address the first question, a comparison of Figures 1 and 2 clearly indicates that integrating AI enhances the likelihood of success in simulations. Specifically, in a scenario where the task difficulty distribution is [0.9,0.05,0.05] and Model 1 serves as the recovery model, the AI-integrated configuration results in 30 failures, compared to 60 failures in the non-AI configuration. This trend is consistent across various cases, with AI-assisted configurations outperforming those without AI. The benefits of AI integration are particularly evident when the probability of encountering difficult tasks is low, significantly reducing the rate of task failures.

In Figure 1 and Figure 2, we observe that Models 1 and 2 exhibit fewer instances of overload risk. Specifically, Figure 1 shows that in 100 simulations, Recovery Models 1 and 2 register 60 and 71 cases of overload risk, respectively, while all other models consistently reach 100 cases of overload. Similarly, Figure 2 confirms that the risk of overload for Models 1 and 2 is lower compared to other recovery models. This indicates a potential advantage in the effectiveness of Models 1 and 2 in managing workload to prevent overload. we conclude that quadratic model with a positive quadratic coefficient is mostly efficient in our simulation case. And quadratic model with a negative $a$ are not suitable for this situation.

Our preliminary findings suggest that the inclusion of AI novices could potentially reduce the risk of overload, particularly when the probability of encountering difficult tasks is relatively low.

To answer the question if both an AI and a human operator of equivalent skill levels are involved, will the human receive adequate training, we conducted a comparative analysis between Figure 1 and Figure 2. This analysis focused on the proportion of tasks performed by novice operators in two scenarios: one involving both human novices and advanced operators, and the other incorporating human advanced operators, and AI novices.

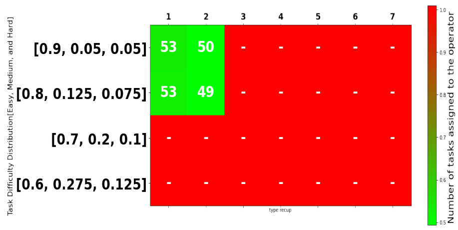


Figure 3: The portion of task assigned to human novice for the case human novice + advanced.

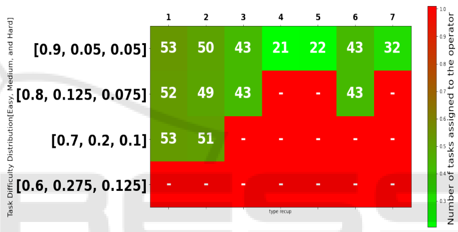


Figure 4: The portion of task assigned to human novice for the case human novice + advanced + AI novice.

In figures 3 and figure 4, both the y-axis and x-axis are identical with figure 1. The numbers displayed on the grid represent the proportion of tasks assigned to a human novice in simulations where there is no risk of overload. For instance, a value of 54 indicates that, out of 100 tasks, 54 are assigned to the human novice. In comparing Figure 3 with Figure 4, it becomes evident that integrating AI not only maintains the need for training novice operators but also enhances their capacity to handle more complex missions. This observation is further supported when analyzing Figure 1 and Figure 2, which demonstrate that the presence of AI enables less experienced operators to undertake more challenging tasks. This adaptation indicates a substantial improvement in both operational capacity and training efficiency, as AI assists in managing and mitigating the potential challenges faced by novice operators in high-demand scenarios.

### 5.3 Summary of Findings for Each Question

1. Regarding the first question, the results indicate that integrating an AI system proves beneficial in reducing cognitive overload.
2. Concerning the second question, our analysis reveals that a quadratic model with positive coefficients is most appropriate for our case.
3. The results from the third question indicate that integrating AI into the workflow does not change the quantity of tasks assigned to trainees. This demonstrates that the inclusion of AI has no adverse effects on the volume of operational training tasks.

### 5.4 Discussion

Our experimental findings affirm the critical role of AI integration in mitigating cognitive overload. Nevertheless, addressing overload requires a multidimensional approach that extends beyond AI facilitation to include flexibility, learning capabilities, and unforeseen event management. Crucially, maintaining comprehensive training for human pilots emphasizes the imperative of human oversight.

We have identified a quadratic model with positive coefficients as optimal for our experimental context. Yet, real-world applications necessitate consideration of environmental impacts, feasibility, and benefits associated with downtime activities. Our algorithm not only supports AI assistance but also strategically manages the workload on trainees—essential in training-centric scenarios. To bolster security and maintain operational readiness, an AI-first strategy is recommended, ensuring human operators are prepared to intervene when AI limitations surface.

## 6 CONCLUSION

This paper studied the tasks scheduling of a Human-AI collaborative drone-piloting mission. A novel task scheduling algorithm integrating micro-breaks during the mission was proposed to schedule necessary recovery time for human workers during their working hours to deal with the human mental workload overload while while fulfilling the designated mission. Experiments show that our method allows to integrate AI in order to reduce the risk of overload while preserving the training of human pilots. Regarding the recovery model, results require to be tested with observation from real-life situations.

## REFERENCES


Albulescu, P., Macsinga, I., Rusu, A., Sulea, C., Bodnaru, A., and Tulbure, B. T. (2022). ”Give me a break!” A systematic review and meta-analysis on the efficacy of micro-breaks for increasing well-being and performance. *PLOS ONE*, 17(8):e0272460.

Asadayoobi, N., Taghipour, S., and Jaber, M. (2023). Optimising stochastic task allocation and scheduling plans for mission workers subject to learning-forgetting, fatigue-recovery, and stress-recovery effects. *Expert Systems with Applications*, 229:120524.

Brazaitis, M. and Satas, A. (2023). Regular short-duration breaks do not prevent mental fatigue and decline in cognitive efficiency in healthy young men during an office-like simulated mental working day: An EEG study. *International Journal of Psychophysiology*, 188:33–46.

Caldwell, J. A., Mallis, M. M., Caldwell, J. L., Paul, M. A., Miller, J. C., and Neri, D. F. (2009). Fatigue Countermeasures in Aviation. *Aviation, Space, and Environmental Medicine*, 80(1):29–59.

Drenth, P. J. D., editor (1998). *Organizational psychology*. Number 4 in Handbook of work and organizational psychology / ed. by Pieter J. D. Drenth. Psychology Press, Hove.

Hancock, G., Longo, L., Young, M., and Hancock, P. (2021). MENTAL WORKLOAD. In Salvendy, G. and Karwowski, W., editors, *HANDBOOK OF HUMAN FACTORS AND ERGONOMICS*, pages 203–226. Wiley, 1 edition.

Hoover, C. S., Ragsdale, J. M., and Ayres, T. B. (2022). An experimental test of resource recovery from physical and relaxation work break activities. *Stress and Health*, 38(3):477–489.

Jaber, M. Y. and Neumann, W. P. (2010). Modelling worker fatigue and recovery in dual-resource constrained systems. *Computers & Industrial Engineering*, 59(1):75–84.

Li, K., Xu, S., and Fu, H. (2020). Work-break scheduling with real-time fatigue effect and recovery. *International Journal of Production Research*, 58(3):689–702.

Tucker, P. (2003). The impact of rest breaks upon accident risk, fatigue and performance: A review. *Work & Stress*, 17(2):123–137.

Xie, B. and Salvendy, G. (2000). Prediction of Mental Workload in Single and Multiple Tasks Environments. *International Journal of Cognitive Ergonomics*, 4(3):213–242.

Zhang, M., Li, C., Shang, Y., Huang, H., Zhu, W., and Liu, Y. (2022). A task scheduling model integrating micro-breaks for optimisation of job-cycle time in human-robot collaborative assembly cells. *International Journal of Production Research*, 60(15):4766–4777.

# APPENDIX

## Experience and Result 2

We also have compared other configurations:

- human advanced, human intermediate
- human advanced, AI intermediate

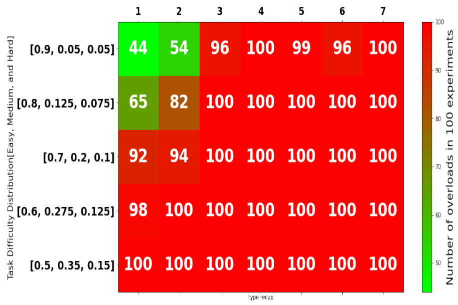


Figure 5: The result with 100 simulation, one human operator level advanced, one human operator level intermediate.

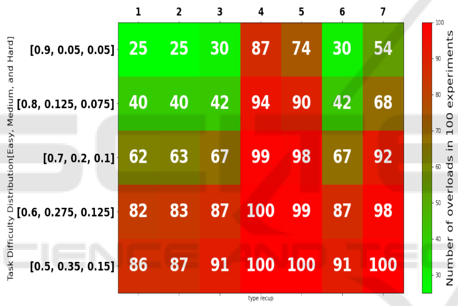


Figure 6: The result with 100 simulation, one human operator level advanced, one AI operator level intermediate.



Comparing Figure 2 and Figure 5, it is apparent that a novice AI proves more beneficial than an intermediate human. Additionally, when assessing Figure 2 and Figure 6, the benefits of high level AI in terms of risk avoidance become more pronounced as the probability of encountering difficult tasks increases. While the improvement may not be significant at lower probabilities, the value of high level AI in managing higher-difficulty tasks becomes increasingly crucial as the likelihood of such challenges grows. We don't test scenarios involving the insertion of an advanced AI because it's clear that such an AI could handle all tasks efficiently, eliminating the risk of overload.