

## Manuscript: Graph Coloring Models for Production Line Scheduling Optimization

We would like to express our sincere appreciation to the Editor and the reviewers for their thorough evaluation of our manuscript and for their constructive and insightful comments. We have carefully revised the paper to address all observations, and all corresponding changes have been highlighted in green. We are confident that these revisions have substantially improved the rigor and clarity of the work. The following document provides a detailed, point-by-point response to each comment raised by the Associate Editor and the reviewers.

### Editor

Please consider removing Fig. 1, as it does not provide relevant information

**Reply:** We thank the reviewer for this observation. The figure has been removed from the revised manuscript as suggested.

### **Actions taken in the manuscript:**

- Figure 1 and its corresponding reference have been removed from the manuscript.

After an equation ending with a comma or semicolon, do not add indentation or capitalize the following sentence.

**Reply:** We thank the reviewer for this helpful observation. Although the specific location of the issue was not indicated, we conducted a careful review of all equations in the problem statement. In particular, we revised the explanatory sentences that follow the formulas, removing unintended indentation and correcting any improper capitalization to ensure full compliance with the formatting requirements. We trust that these targeted revisions adequately address the reviewer's concern.

### **Actions taken in the manuscript:**

- The capitalization and indentation of explanatory sentences following equations have been corrected throughout the problem statement.

Maintain consistent grammatical person and number throughout the article; avoid alternating between first-person (we do...) and impersonal constructions (it has been done)

**Reply:** We appreciate this valuable suggestion. Maintaining a consistent grammatical person indeed improves the clarity and readability of the manuscript. Accordingly, we have revised the document to ensure that all sections are written in the first-person plural.

**Actions taken in the manuscript:**

- The grammatical person has been unified across the manuscript, particularly in the simulated annealing definition subsection, the results section, and the future work discussion.

Please, avoid referring to functions/variable names from your own code. Use clear identifiable names for these parameters (i.e. TABLE 2)

**Reply:** You are right. We have revised the column names in all tables to ensure they are clearer and no longer reference implementation-specific variable names.

**Actions taken in the manuscript:**

- The variable and parameter names have been updated in Tables 2, 3, and 4 to use clear, descriptive terminology.

Multi-objective problem is not formally presented

**Reply:** Thank you for your comment. This suggestion is highly relevant and contributes significantly to improving the quality of the manuscript. We have therefore added the formal mathematical formulation of the multi-objective algorithm to the problem statement.

**Actions taken in the manuscript:**

- A dedicated subsection has been added to the problem statement presenting the formal mathematical formulation of the multi-objective scheduling problem.

In TABLE 5, the trailing digit in 232.20000000000002 is indicative of numerical cancellation or floating-point round-off error, please use two significant figures

**Reply:** Thank you for pointing this out. The value has been rounded to two decimal places as suggested.

**Actions taken in the manuscript:**

- The reported fitness values have been rounded to two decimal places for all algorithms and problem instances.
- Following the reviewers' recommendations to improve statistical clarity, the original Table 5 has been split into Tables 5, 6, and 7.

Reviewer 1

Mayor improvements

*Use of colloquial language and literal translations:* The article contains several expressions that do not conform to the expected technical-formal style of academic literature. For example,

phrases such as “Of course, these dynamic values will not be optimal for every single scheduling instance” should be avoided, as the use of “of course” introduces an unnecessary colloquial tone. Similarly, direct translations such as “As exposed previously” were detected, which is not an idiomatic construction in academic English. Instead, “As mentioned previously” or “As stated earlier” are recommended.

**Reply:** We agree with this observation. Since English is not our first language, we acknowledge the importance of carefully avoiding colloquial expressions and non-idiomatic constructions. We appreciate the reviewer’s guidance, which has helped us improve the academic tone of the manuscript.

**Actions taken in the manuscript:**

- The expression “Of course” has been replaced with “Naturally” to maintain a more formal tone.
- The phrase “As exposed previously” has been replaced with the idiomatic academic expression “As stated earlier.”

*Use of references without contextual integration:* In several passages of the article, references are listed cumulatively at the end of sentences without explaining their specific relevance. For example, in the introduction, [4], [5], and [6] are cited as supporting evidence for a claim about environmental benefits, without indicating which aspect each reference addresses. To improve the usefulness of the references, it is recommended to integrate them into the discourse by briefly justifying their inclusion (for example: “as demonstrated in [4], where graph coloring is applied in energy-constrained planning scenarios”).

**Reply:** We appreciate this suggestion. We agree that citing references without clarifying their specific contributions may reduce clarity and lead to potential misunderstandings. Accordingly, we have revised the introduction to better integrate the cited works into the narrative by explicitly indicating their relevance.

**Actions taken in the manuscript:**

- The references have been more explicitly integrated into the phrasing of the introduction.
- The contributions of references [4], [5], and [6] have been clarified within the text. The revised paragraph now reads as follows:

“As reported in [4], [5], and [6], graph coloring problem–based algorithms have been successfully applied to scheduling tasks. Accordingly, we will cover a graph-theoretic optimization method based on the graph coloring problem to reduce idle times, balance workloads, and minimize energy consumption, ultimately enhancing both operational and environmental performance.”

*Mixed Organization of Content in Section IV:* The inclusion of subsections H. Algorithm Summary and I. Simulated Annealing Algorithm in Section IV makes it too long and mixes conceptual, technical, and experimental elements. This is particularly noticeable because the

three algorithms are described again in Section V. It is suggested that the methodological description (Section IV) be clearly separated from the implementation and experimental setup (for example, in a specific section such as “Experimental Setup”) to avoid redundancies and improve readability.

**Reply:** Thank you for this observation. We acknowledge that Section IV may appear extensive; however, we consider the inclusion of the simulated annealing algorithm to be methodologically justified, as it serves as the baseline against which the proposed genetic algorithms are evaluated. In Section IV, we focus on describing how each algorithm operates from a methodological perspective, whereas Section V is dedicated exclusively to the experimental configuration and parameter settings. Therefore, no redundancy exists between these sections. Nevertheless, since simulated annealing is not one of our proposed contributions, we have reduced its prominence in the methodology and retained only the information strictly necessary to understand the experimental procedure.

**Actions taken in the manuscript:**

- We explicitly clarified that the simulated annealing algorithm is used solely as a baseline by adding the following sentence at the beginning of the corresponding subsection:

“To evaluate the performance of the proposed genetic algorithms, we employ a simulated annealing (SA) algorithm as a baseline. SA is a well-established metaheuristic commonly used for combinatorial optimization problems and provides a suitable reference for assessing solution quality and convergence behavior.”

- We removed non-essential descriptive content related to simulated annealing and retained only the key elements required for understanding the experimental comparison.
- Specifically, we removed the following paragraph:

“At each iteration, a new candidate solution is accepted if it improves the objective function; otherwise, it may still be accepted with a probability that decreases as the algorithm's temperature parameter is reduced. This mechanism allows the method to escape local optima during the early stages and gradually converge toward high-quality solutions as the temperature approaches zero.”

*Weak connection between text and figures in section V:* The figures showing the evolution of the algorithms (Figures 5, 6, 7) are barely mentioned in the text, without any interpretation. This limits their usefulness. For example, Figure 8, which represents feasible solutions obtained by the multi-objective GA, should be explained as a visualization of the Pareto front in terms of conflicts vs. energy performance.

**Reply:** We agree that the figures in Section V would benefit from a more explicit interpretation in the manuscript. Without adequate discussion, the results may appear to rely primarily on the numerical tables. In response to this comment, we have expanded the accompanying text to provide a clearer explanation and interpretation of all figures, explicitly linking them to the observed algorithmic behavior and experimental outcomes.

**Actions taken in the manuscript:**

- We have added the following paragraph to provide a detailed explanation of Figures 4, 5, and 6:

“Even though problem instances FT10 and LA01 are more complex, Figures 5 and 6 show a fast convergence behavior very similar to that observed for the FT06 instance in Figure 4. In all cases, conflict-free solutions are reached as the mean fitness approaches zero-conflict values, which can be attributed to the applied elitism strategy. While this elitism promotes rapid convergence, it also entails that once a feasible solution is found, further fitness improvements become progressively slower and may nearly stagnate. However, as will be shown in Tables 5, 6, and 7, the solutions obtained are sufficiently efficient when compared to the other algorithms.”

- We have added the following paragraphs to further explain Figures 7, 8, and 9:

“Although Figures [7] and [9] display only a single point, multiple feasible solutions were obtained in both cases. However, all these solutions share identical objective values, which results in a single visible point on the Pareto front. This behavior is likely a consequence of the elitism and selection mechanisms employed by the algorithm, which tend to promote rapid convergence toward a small set of dominant solutions, thereby limiting population diversity.

Figure [8] provides a broader view of the Pareto front for the FT10 instance, where all solutions are infeasible. The figure illustrates how the algorithm explores different trade-offs among infeasible solutions, with some individuals prioritizing reductions in peak load and others focusing on minimizing the number of conflicts. Nevertheless, the Pareto front remains relatively narrow, indicating limited diversity among the solutions identified for this instance.”

- We have added the following paragraph to further explain Figures 10, 11 and 12:

“Contrary to the behavior observed in Figure [10], Figures [11] and [12] exhibit a clearer convergence pattern once the exploration phase concludes, with a consistent downward trend in the fitness values across iterations. This behavior is consistent with the expected dynamics of the simulated annealing algorithm and indicates that, for these more complex instances, the balance between exploration and exploitation is more effective. In these cases, the initial exploration allows the algorithm to escape poor-quality regions of the search space, while the subsequent exploitation phase progressively refines the solutions toward higher-quality configurations.”

*Lack of justification for the choice of certain parameters:* The choice of certain factors in the algorithm's parameters (e.g.,  $\text{population} = \max(15 \cdot |O|, 200)$ ) is not justified. Although it is mentioned that they may require adjustment, it would have been helpful to clarify whether they were chosen through trial and error, based on previous results in the literature, or for empirical reasons.

**Reply:** We appreciate the reviewer's valuable comments. Presenting different parameter choices without adequate explanation could indeed be confusing for the reader. Therefore, we have added a clarification explaining the rationale behind the parameter choice for each algorithm.

**Actions taken in the manuscript:**

- Added an explanatory paragraph to the results section describing the rationale behind the parameter selection process:  
"All parameters in each algorithm were selected using a trial-and-error approach, with two to three executions performed to ensure that changes in parameter values had a significant impact on the algorithm's results and that the observed improvements were not due to the algorithm's inherent randomness."
- Changed Table 1's Number of Generations from "200" to " $\approx 500$  (Dynamic)".
- Added the following paragraph to Section V-A:  
"As it can be observed, the "Number of Generations" parameter differs from the single-objective algorithm ( $\max(\text{nops} \cdot 10, 500)$  instead of  $\max(\text{nops} \cdot 5, 200)$ ) This is due to the differences between the two algorithms, requiring a higher number of generations."

*Lack of statistical repetition in experiments with stochastic algorithms:* Since both GA and SA are stochastic algorithms, reporting a single run is insufficient. It would be more rigorous to run multiple runs with different seeds and report metrics such as mean, standard deviation, and best/worst results. This would allow for an evaluation of the algorithm's robustness and stability.

**Reply:** We completely agree that relying on a single fixed seed is insufficient for a stochastic metaheuristic. We have addressed this critical point by re-engineering our experimental validation to provide statistically robust conclusions.

Specifically:

1. Multiple Runs: We re-executed all three algorithms using multiple independent runs with different random seeds.
2. Statistical Reporting: The results tables have been completely overhauled to report the Mean and Standard Deviation ( $\pm$  SD) for all key metrics (Cost/Objective Value and Conflicts).
3. Clarity: To address the lack of clarity in the previous results table (as noted by other reviewers), we have replaced the single wide table with three separate, instance-specific tables. This structure improves readability and clearly distinguishes between the reported metrics.

We are confident that these changes significantly enhance the rigor and credibility of our experimental findings.

**Actions taken in the manuscript:**

- Re-executed all experiments using multiple independent random seeds (N=5 or N=10) for statistical validation.
- Replaced the original "Results" table with Tables 5, 6, and 7, which now report the Mean  $\pm$  Standard Deviation for Cost/Objective Value and Conflicts across all runs.
- Updated the manuscript's Results section to ensure it is made clear that multiple runs of our algorithm were performed

*Limited literature review on sustainability:* Although the article's focus is explicitly on energy efficiency, the literature review (section II) does not include relevant references on green scheduling or sustainable manufacturing.

**Reply:** We agree that, given the explicit focus of our work on energy efficiency, it is important to incorporate relevant references on green scheduling and sustainable manufacturing in the literature review. To address this concern, we have expanded Section II by adding recent survey papers that specifically cover energy-efficient scheduling and sustainability-oriented approaches.

**Actions taken in the manuscript:**

- We have added three survey papers to the literature review. The following paragraph has been incorporated into the Related Work section:

"Going a step further, recent surveys such as [13], [14] and [15] provide comprehensive reviews of energy-efficient scheduling problems, identifying evolutionary approaches—particularly genetic algorithms—as prominent and effective solution techniques."

Minor improvements

*Section title change: "Related Works"  $\rightarrow$  "Related Work":* The title of section II appears as "Related Works", when the accepted usage in most conferences and journals is "Related Work", in the singular.

**Reply:** Thank you for pointing this out. The section title has been corrected accordingly.

**Actions taken in the manuscript:**

- The section name has been changed from "Related Works" to "Related Work".

*Terminological and stylistic inconsistencies in technical terms:* The article uses expressions such as "multi objective" and "multi-objective," or "job shop scheduling" and "Job-Shop

Scheduling.” It is recommended to maintain a clear and consistent convention: use “multi-objective” with a hyphen and define “Job-Shop Scheduling (JSS)” from the beginning before using the acronym.

**Reply:** We concur with this observation. Terminological inconsistencies can negatively affect the clarity, precision, and overall quality of the manuscript. Accordingly, we have revised the text to ensure the consistent use of technical terminology throughout the paper, including standardized forms and definitions..

**Actions taken in the manuscript:**

- We have replaced all occurrences of “multi objective” with the standardized hyphenated form “multi-objective.”
- We have unified the terminology related to job shop scheduling by using the form “Job-Shop Scheduling”.

*Unjustified choice of the number of elites:* The GA's elitism scheme explanation states that the top 5 individuals ( $N_{elite}$ ) are retained per generation, but this choice is not justified. Since the population size is dynamic (instance-dependent), a brief proportional justification would be expected (e.g., “5 individuals represents approximately 0.3% of the population for median instances”).

**Reply:** We acknowledge that selecting a fixed elitism parameter without explicit justification may be unclear, particularly when the population size varies across problem instances. To address this concern, we have added a detailed explanation in the methodology section clarifying the rationale behind the chosen elite size.

**Actions taken in the manuscript:**

- We have added the following paragraph to the *Replacement and Elitism* subsection of the methodology:

“We set the elite size to 5 for both genetic algorithms. This value was selected to ensure that a small yet representative subset of high-quality solutions is preserved across generations, regardless of the dynamic population size. Using a fixed and relatively small elite size prevents excessive selection pressure as the population grows, while still guaranteeing that the best individuals are not lost. Moreover, since tournament selection already introduces a degree of elitism, limiting the number of explicitly preserved individuals helps maintain population diversity and reduces the risk of premature convergence. Empirically, this configuration provided a good balance between solution preservation and exploration across problem instances of varying sizes.”

*Lack of information on computational performance:* The article does not report the execution times or computational efficiency of the algorithms, which limits its practical applicability. It would be very useful to include, at least as a guideline, how long each method took per instance, to assess the feasibility of its use in real-world or scalable environments.



**Reply:** Thank you for this suggestion. Indeed, execution time is a relevant factor when addressing scheduling problems, as it directly affects the practical viability and scalability of the proposed algorithms. We have therefore taken this concern into account and introduced the corresponding modifications to the manuscript.

#### **Actions taken in the manuscript:**

- Execution time statistics have been added to Tables [5], [6] and [7].
- A discussion of the computational performance and execution times has been incorporated into the Results section.

“Regarding execution time, as expected, simulated annealing is the fastest algorithm, since the genetic algorithms must evaluate the fitness function for a large number of individuals at each generation. Nevertheless, the single-objective genetic algorithm exhibits competitive execution times for small- and medium-sized problem instances. For larger instances, however, its execution time increases significantly; for example, as shown in Table~[6], it requires  $223.4 \pm 2.18$  seconds (nearly four minutes) to complete. Depending on production requirements, such execution times may be impractical. However, as previously discussed in the analysis of Figures~[4], [5], and [6], the algorithm could be terminated earlier—once a feasible solution is found—at the cost of a slightly worse objective value but with a substantial reduction in execution time, still yielding more robust solutions than simulated annealing.

The multi-objective genetic algorithm presents the highest execution times across all problem instances. When combined with its inability to consistently find conflict-free solutions—particularly for larger instances—this significantly limits its applicability in real-world scenarios. Nevertheless, in contexts where execution time is less critical and solution quality for small problem instances is the primary concern (e.g., offline scheduling or next-day production planning), the multi-objective approach may still represent a viable alternative.”

- Additional remarks on execution time and practical applicability have been included in the Conclusions section.
  - “Among the algorithms evaluated, the single-objective genetic algorithm demonstrates the strongest overall performance. It consistently produces feasible solutions across all tested instances and achieves better fitness values than the multi-objective approach for medium- and large-sized problems. *In addition, it exhibits a competitive execution time, making it suitable for practical and time-constrained scheduling scenarios.*”
  - “For problem instances with fewer operations and/or machines, the multi-objective genetic algorithm remains a viable alternative. Its ability to explore trade-offs between objectives may provide advantages in scenarios where solution diversity is desirable; *however, its higher execution time must be carefully considered when assessing its applicability in time-sensitive environments.*”

*Unclear Legend in Table 5:* The table summarizing the experimental results has columns labeled “Results” and “Conflicts” without explicit explanation. To avoid misunderstandings, it would be helpful to include something like this in the legend: “Results: objective function value achieved; Conflicts: number of conflicts remaining in the best solution.”

**Reply:** To address the lack of clarity in the previous results table (as noted by other reviewers), we have replaced the single wide table with three separate, instance-specific tables. This structure improves readability and clearly distinguishes between the reported metrics

**Actions taken in the manuscript:**

- Table [5] has been divided into tables [5], [6] and [7].

*Typographical inconsistency in the “D. Assumptions” section:* Assumption 1 appears in plain text, while Assumptions 2–6 are indented. This typographical inconsistency is visually jarring. It is recommended to standardize the formatting (for example, by removing the indentation and using \noindent in LaTeX) to present all assumptions consistently.

**Reply:** Thank you for your observation. To ensure consistent formatting across the section, we have reorganized the assumptions into a numbered list so that all items appear aligned at the same typographical level.

**Actions taken in the manuscript:**

- All assumptions have been included in a numbered list.

*Outdated References:* Although some recent sources are included, there are also several significantly older references (e.g., from 1989, 1992, and 2016). While these may be valid for theoretical foundations, it is recommended to supplement them with more current literature.

**Reply:** Thank you for your observation. We acknowledge the presence of several older references in our manuscript. However, these works correspond to foundational contributions that formally define the classical algorithms and theoretical principles on which our approach is built. Since these references are used strictly to establish the underlying methodological framework (not to support recent developments or empirical improvements) we believe they remain appropriate and necessary. For this reason, we have opted to retain them. Nevertheless, we have reviewed the manuscript to ensure that all discussions of contemporary advances are supported by up-to-date literature where relevant.

## Reviewer 2

En general vuestro trabajo está bien planteado, no obstante, antes de publicarse hay ciertos aspectos de entidad (major revision) que deben refinarse:

## 1. Methodology (GA y SA)

Hay inconsistencias entre las distintas tablas en el número de generaciones (por ejemplo, Tabla 1 frente a Tablas 2 y 3). Sería conveniente unificarlas y dejar claro qué configuración se ha usado realmente en los experimentos.

**Reply:** We appreciate your valuable comments. The differing number of generations in each experiment would be confusing for the reader. We have fixed Table 1's inaccurate values and added an explanation detailing why the number of generations is different between the single-objective and the multi-objective experiments, as this distinction is important.

### Actions taken in the manuscript:

- Changed Table 1's Number of Generations from "200" to " $\approx 500$  (Dynamic)".
- Added the following paragraph to Section V-A:  
"As it can be observed, the "Number of Generations" parameter differs from the single-objective algorithm ( $\max(\text{nops} \cdot 10, 500)$  instead of  $\max(\text{nops} \cdot 5, 200)$ ) This is due to the differences between the two algorithms, requiring a higher number of generations."

## 2. Experimental Results

Las métricas no son muy claras en la tabla 5 "Results" ¿Qué se está comparando exactamente? No se analizan componentes relevantes como el makespan sería conveniente visualizar cómo se reduce. No hay validación estadística, solo se usa una semilla fija para reproducibilidad, pero no se repiten los experimentos para múltiples semillas, tampoco se analizan intervalos de confianza. No se menciona posibles causas de por qué no hay soluciones libres de conflictos para el FT10.

**Reply:** We completely agree that relying on a single fixed seed is insufficient for a stochastic metaheuristic. We have addressed this critical point by re-engineering our experimental validation to provide statistically robust conclusions.

Specifically:

4. Multiple Runs: We re-executed all three algorithms using multiple independent runs with different random seeds.
5. Statistical Reporting: The results tables have been completely overhauled to report the Mean and Standard Deviation ( $\pm$  SD) for all key metrics (Cost/Objective Value and Conflicts).
6. Clarity: To address the lack of clarity in the previous results table (as noted by other reviewers), we have replaced the single wide table with three separate, instance-specific tables. This structure improves readability and clearly distinguishes between the reported metrics.

We are confident that these changes significantly enhance the rigor and credibility of our experimental findings.

**Actions taken in the manuscript:**

- Re-executed all experiments using multiple independent random seeds (N=5 or N=10) for statistical validation.
- Replaced the original "Results" table with Tables 5, 6, and 7, which now report the Mean  $\pm$  Standard Deviation for Cost/Objective Value and Conflicts across all runs.
- Updated the manuscript's Results section to ensure it is made clear that multiple runs of our algorithm were performed

3. La figura (1) debería especificarse de dónde se ha obtenido

**Reply:** As suggested by the Associate Editor, we removed the figure from the manuscript, as it did not provide relevant information on the topic, making it no longer necessary to add a source for the figure.

**Actions taken in the manuscript:**

- Removed Figure 1 from the manuscript.

Algunos detalles menores restantes:

1. **Título**

El título es genérico podría reflejar el enfoque energy aware o mencionar el empleo de algoritmos genéticos como estrategia de resolución.

**Reply:** Thank you for your suggestion, to better reflect the methodological contribution of our manuscript, we have updated the title to: "Energy-Aware Production Line Scheduling Optimization using Graph Coloring and Genetic Algorithms".

**Actions taken in the manuscript:**

- Changed title from "Graph Coloring Models for Production Line Scheduling Optimization" to "Energy-Aware Production Line Scheduling Optimization using Graph Coloring and Genetic Algorithms".

2. **Abstract**

Cada vez que aparezca un nuevo acrónimo se debe explicar su significado, esto es especialmente relevante en el abstract. Adicionalmente se podría mejorar la frase: "demonstrate superior performance compared to the simulated annealing" especificando la magnitud de la mejora. Tampoco está clara la pregunta de investigación que se quiere resolver.

**Reply:** Thank you for pointing this out. Thanks to your comment, we found an acronym which was not defined in the manuscript. As for the quoted phrase, we have specified the magnitude of the improvement in order to ensure the difference is clear.

**Actions taken in the manuscript:**

- Defined the meaning of JSPLIB the first time it is mentioned in the manuscript.
- Modified the following sentence in the Abstract:  
“Both genetic algorithms demonstrate superior performance compared to the simulated annealing baseline, with a 89.91% average difference in overall fitness of the solutions between the approaches.”

**3. Introducción**

Motivación y contexto es un poco difuso mezclando términos variados como (Industria 5.0, sostenibilidad, scheduling). Por otro lado, el research gap frente a otros trabajos de graph coloring en job-shop debería explicitarse un poco más. Falta una frase tipo “To the best of our knowledge, no previous work has...”

**Reply:** Thank you for your comment. We do agree that the research gap is not stated clearly in the paper, thus, we have added a clarifying statement to the manuscript that reinforces the novelty of our contributions.

**Actions taken in the manuscript:**

- Added the following sentence to our Introduction:  
“To the best of our knowledge, no previous work has combined a pure graph coloring formulation with multi-objective genetic algorithms specifically for peak-load minimization in this specific job-shop context.”

**4. Related Work**

Modificad el título de la sección por “Related work” en singular. El repaso de graph coloring es correcto, pero falta explicar mejor en qué se diferencia vuestra propuesta de trabajos previos cercanos (p.ej. mixed graph coloring en job-shop). Tal y como está ahora no queda claro que falta en la literatura ¿No hay nadie que haya usado un GA en este contexto?

**Reply:** We appreciate these helpful observations. The title of the “Related Works” section was also pointed out by other reviewers, and has thus already been addressed in the manuscript. As for the research gap, it has also been addressed at the previous point.

**Actions taken in the manuscript:**

- Changed Section II’s title from “Related Works” to “Related Work”
- Added the following sentence to our Introduction:  
“To the best of our knowledge, no previous work has combined a pure graph coloring formulation with multi-objective genetic algorithms specifically for peak-load minimization in this specific job-shop context.”

**5. Problem Formulation**

Una tabla de notación con todos los símbolos ( $C_{max}$ ,  $L_{peak}$ ,  $P_{mach}$ ,  $P_{prec}$ , etc.) mejoraría la lectura.

**Reply:** Thank you for the suggestion. While we do agree that a table might aid the reader in understanding what each symbol means, each new symbol is explained after its first appearance in the paper, thus, making this table redundant. We have, however, decided to make the understanding of some of these symbols easier by presenting their meanings in a list-like format.

**Actions taken in the manuscript:**

Modified the following explanations in Section III:

- Changed “where  $L_{\text{peak}}(x)$  is the maximum number of simultaneous machines (peak load),  $C_{\text{max}}$  is the makespan, and  $P_{\text{mach}}$  and  $P_{\text{prec}}$  represent the count of machine and precedence conflicts, respectively.” to  
“where:
  - $L_{\text{peak}}(x)$  is the maximum number of simultaneous machines (peak load),
  - $C_{\text{max}}$  is the makespan,
  - $P_{\text{mach}}$  and  $P_{\text{prec}}$  represent the count of machine and precedence conflicts, respectively.”
- Changed “where  $L_{\text{peak}}(x)$  is the peak load,  $C_{\text{max}}(x)$  is the makespan, and  $P_{\text{mach}}(x)$  and  $P_{\text{prec}}(x)$  represent the number of machine and precedence conflicts, respectively.” to  
“where:
  - $L_{\text{peak}}(x)$  is the peak load,
  - $C_{\text{max}}(x)$  is the makespan,
  - $P_{\text{mach}}(x)$  and  $P_{\text{prec}}(x)$  represent the number of machine and precedence conflicts, respectively.”

**6. Conclusiones y Referencias/Código**

Hay limitaciones importantes que no se han señalado como el uso de instancias sintéticas y no datos en fábricas reales y sobre el modelo de energía simplificado cómo máquinas idénticas, sin costes de paradas, etc. Los futuros trabajos son algo genéricos.

**Reply:** Thank you for your valuable comment. It is true that some of these limitations should have been addressed in the conclusion, thus, we expanded the Conclusion to explicitly acknowledge them, specifically the use of synthetic instances (JSPLIB) rather than real-world factory data. On the other hand, the assumption of a simplified model (uniform machine energy consumption, full-time machine availability, etc) is already mentioned in the Conclusion.

**Actions taken in the manuscript:**

- Modified the wording in Section VI of the manuscript, changing “the scalability to larger industrial instances” to “the scalability to larger, real industrial instances, rather than synthetic instances”.

Reviewer 3

**Major error**

1. **Lack of statistical validation**

Because a fixed random seed was used to obtain the results, no statistical validation was performed. The algorithm should be repeated with different seeds to compute confidence intervals, averages, and related statistical measures.

**Reply:** We completely agree that relying on a single fixed seed is insufficient for a stochastic metaheuristic. We have addressed this critical point by re-engineering our experimental validation to provide statistically robust conclusions.

Specifically:

1. Multiple Runs: We re-executed all three algorithms using multiple independent runs with different random seeds.
2. Statistical Reporting: The results tables have been completely overhauled to report the Mean and Standard Deviation ( $\pm$  SD) for all key metrics (Cost/Objective Value and Conflicts).
3. Clarity: To address the lack of clarity in the previous results table (as noted by other reviewers), we have replaced the single wide table with three separate, instance-specific tables. This structure improves readability and clearly distinguishes between the reported metrics.

We are confident that these changes significantly enhance the rigor and credibility of our experimental findings.

**Actions taken in the manuscript:**

- Re-executed all experiments using multiple independent random seeds (N=5 or N=10) for statistical validation.
- Replaced the original "Results" table with Tables 5, 6, and 7, which now report the Mean  $\pm$  Standard Deviation for Cost/Objective Value and Conflicts across all runs.
- Updated the manuscript's Results section to ensure it is made clear that multiple runs of our algorithm were performed

**Minor error**

*Introduction*

1. **Page 1 (Figure 1): Missing source.**

The figure does not include a source. If the image is original, this is not an issue; however, if it is not, a proper citation must be provided. Images without an associated licence cannot be used.

**Reply:** As suggested by the Associate Editor, we removed the figure from the manuscript, as it did not provide relevant information on the topic, making it no longer necessary to add a source for the figure.

**Actions taken in the manuscript:**

- Removed Figure 1 from the manuscript.

*Statement of the problem*

2. **Page 4 (D. ASSUMPTIONS): Unjustified assumptions.**

Several assumptions are stated without any justification. It is recommended to provide a rationale for each assumption.

**Reply:** Thank you so much for the recommendation. Without the context of the future work, these assumptions may appear nonsensical. An explanation for the assumptions has been added to the manuscript.

**Actions taken in the manuscript:**

- The following paragraph has been added to Section III-D:  
“We make these assumptions that simplify the scheduling instances and operations so that we can focus on the scheduling logic and proposed algorithm on this paper. Future work will take these elements into account to further refine the algorithm and make it even more applicable in more realistic scenarios.”

*Proposed Evolutionary Approach*

3. **Page 4 (Figure 3): Figure not referenced in the text.** The figure is not cited within the text. All figures and tables must always be referenced explicitly in the document.

**Reply:** Thank you for pointing this oversight out. We also found this mistake is also made with Figure 2, thus, we will make sure to correctly reference them both in the manuscript.

**Actions taken in the manuscript:**

Added the following sentences to Section IV-A and IV-D respectively:

- “... which can be observed in Figure 2”
- “A visual example of the crossover and mutation processes can be observed in Figure 3”

*Results*

4. **Page 6: Missing Google Colab specifications.** Although it is mentioned that the implementation was executed in Google Colab, the hardware specifications, such as available RAM, should also be provided.

**Reply:** We appreciate the suggestion. While it is easier to ensure reproducibility using a cloud tool such as Colab, we should have made sure to specify the runtime settings we used. We will make sure to add them to the manuscript.

**Actions taken in the manuscript:**

The runtime settings of the Google Colab instance used for each experiment has been added to the Results of the manuscript:



- Runtime Type: Python 3
- Hardware Accelerator: CPU
  - System Ram: 12.7 GB
  - Disk: 107.7 GB
- Runtime Version: Latest (recommended)

**5. Tables 2, 3, and 4: Missing justification for hyperparameters.**

While listing the hyperparameters enhances reproducibility, it is important to justify the choice of these parameters.

**Reply:** Thank you so much for the comment. As pointed out, without an explanation, our choice of parameters may seem arbitrary. We will make sure to explain them further in our manuscript.

**Actions taken in the manuscript:**

- The following paragraph was added to Section V-B:  
“The selected configuration was determined based on our experimentation aimed at identifying parameter values that provide robust performance within a reasonable execution time. As stated before, they could be fine tuned to a greater extent with further experimentation, but from all our testing these values seem to offer a good balance between results and execution time.”

*Global Comments*

**6. Figure and table numbering format**

It is recommended not to use parentheses when referring to figures or tables. Instead of Figure (1), use Figure 1.

**Reply:** We really appreciate the correction. All incorrect instances of incorrectly referred figures and tables will be addressed in the manuscript.

**Actions taken in the manuscript:**

- References to all figures and tables in the manuscript have been changed from “Figure (X)” to “Figure X” and from “Table (X)” to “Table X”

*Reviewer 4*

Your manuscript presents a valuable contribution to the field of production line scheduling optimization by modelling the problem through graph coloring formulations and comparing single-objective GA, multi-objective GA, and simulated annealing on standard JSPLIB instances. The problem is relevant and the mathematical modelling is solid. However, several aspects of the experimental validation and results analysis need further clarification and strengthening before the work can be considered for publication.

## **Specific Comments and Suggestions for Improvement:**

### **Major comments:**

#### **1. Lack of statistical validation of GA results.**

GA performance is reported based on a single run per problem instance, without showing variability across independent runs. For a stochastic metaheuristic, this is not sufficient to support robust conclusions.

**Reply:** We completely agree that relying on a single fixed seed is insufficient for a stochastic metaheuristic. We have addressed this critical point by re-engineering our experimental validation to provide statistically robust conclusions.

Specifically:

1. Multiple Runs: We re-executed all three algorithms using multiple independent runs with different random seeds.
2. Statistical Reporting: The results tables have been completely overhauled to report the Mean and Standard Deviation ( $\pm$  SD) for all key metrics (Cost/Objective Value and Conflicts).
3. Clarity: To address the lack of clarity in the previous results table (as noted by other reviewers), we have replaced the single wide table with three separate, instance-specific tables. This structure improves readability and clearly distinguishes between the reported metrics.

We are confident that these changes significantly enhance the rigor and credibility of our experimental findings.

### **Actions taken in the manuscript:**

- Re-executed all experiments using multiple independent random seeds (N=5 or N=10) for statistical validation.
- Replaced the original "Results" table with Tables 5, 6, and 7, which now report the Mean  $\pm$  Standard Deviation for Cost/Objective Value and Conflicts across all runs.
- Updated the manuscript's Results section to ensure it is made clear that multiple runs of our algorithm were performed

### **Minor comments:**

#### **2. Use of third-party images without attribution or licensing information.**

One figure appears to be directly taken from external sources (Figure 1 from Wikimedia Commons) but is presented as if it were original, with no source or license information in the captions. This is problematic from a copyright standpoint.

**Reply:** As suggested by the Associate Editor, we removed the figure from the manuscript, as it did not provide relevant information on the topic, making it no longer necessary to add a source for the figure.

**Actions taken in the manuscript:**

- Removed Figure 1 from the manuscript.

**3. Inconsistent table formatting and parameter naming:**

Table 3 uses parameter names in all caps with underscores (e.g. POPULATION\_SIZE, NUM\_GENERATIONS), while Table 5 and the rest of the manuscript use standard labels with spaces and normal capitalization (e.g. "Population size", "Makespan").

**Reply:** Thank you so much for pointing out this oversight. We will make sure to fix this inconsistency in the manuscript.

**Actions taken in the manuscript:**

- All table headers and parameter labels have been standardized to Sentence case (e.g., 'Population size') for consistency.