Araştırma Makalesi/Research Article
LYDIA BUSSINESS REVIEW
Yıl/Year: 2025, Cilt/Volume: 01, Issue/Sayı: 01

Hybrid Genetic Algorithm Performance for PFSP: Leveraging NEH Heuristics in Initial Population

Münire YALÇİN[a]

[a]Independent researcher, ORCID: 0009-0003-9673-1974, munireoymanyalcin1@gmail.com

ABSTRACT

The Permutation Flowshop Scheduling Problem (PFSP), with the objective of minimizing the maximum completion time, remains a critical and computationally challenging NP-hard problem in manufacturing. This study investigates the strategic impact of hybridizing a standard Genetic Algorithm (GA) by injecting high-quality solutions generated by the Nawaz-Enscore-Ham (NEH) constructive heuristic into the initial population. Two distinct models, GA-Random (purely randomized baseline) and GA-Hybrid (5% NEH initialization), were rigorously compared using the complete Taillard benchmark dataset. The results demonstrate that the GA-Hybrid model significantly outperforms the baseline GA-Random model across all problem sizes and exhibits competitive performance against the best-known solutions reported in the literature. The findings underscore that for NP-hard scheduling problems, superior performance is achieved through strategic hybridization that combines the global exploration of GAs with the powerful local exploitation provided by problem-specific heuristics during the initialization phase.

KEYWORDS: Permutation Flowshop Scheduling Problem (PFSP), Genetic Algorithm (GA), Nawaz-Enscore-Ham (NEH) Heuristic, Makespan Minimization, Hybrid Metaheuristics.

ARTICLE HISTORY

Recevied: 08.12.2025

Accepted: 24.12.2025

JEL CLASSIFICATION

M11

1. Introduction

Scheduling problems lie at the heart of manufacturing and service industries, representing a critical decision-making process where limited resources must be allocated to tasks to optimize specific objectives. Among these, the Permutation Flowshop Scheduling Problem (PFSP) is one of the most extensively studied and industrially relevant classes of scheduling problems. In a typical flowshop environment, a set of n jobs must be processed on m machines in the same technological order. The challenge lies in determining the optimal sequence of jobs that minimizes a specific performance criterion, most commonly the total completion time, known as the makespan (C_"max").

Despite its simple definition, the PFSP is computationally complex. For problems with more than two machines ($m>2$), it is classified as NP-hard, meaning that the computational time required to find an exact optimal solution grows exponentially with the problem size ($n!$). Consequently, traditional exact methods such as Branch and Bound or Integer Linear Programming become computationally intractable for medium to large-scale instances typically found in real-world manufacturing. This limitation has driven researchers toward metaheuristic algorithms, which sacrifice the guarantee of optimality for the ability to find near-optimal solutions within reasonable timeframes.

Among metaheuristics, the Genetic Algorithm (GA), inspired by the principles of natural evolution, has proven to be a robust and flexible tool for solving combinatorial optimization problems like the PFSP. However, the performance of a GA is heavily dependent on the balance between exploration (searching new areas of the solution space) and exploitation (refining existing solutions). A critical, yet often overlooked, factor influencing this balance is the composition of the initial population. Traditional GAs often rely on purely random initialization, which ensures diversity but may start the search in poor-quality regions of the solution space, leading to slow convergence.

This study aims to address this efficiency gap by investigating a hybrid initialization strategy. Specifically, we propose enhancing the standard GA by injecting a small fraction of high-quality solutions generated by the Nawaz-Enscore-Ham (NEH) heuristic—widely regarded as the best constructive heuristic for the PFSP—into the initial population. By comparing this GA-Hybrid model against a traditional GA-Random model using the renowned Taillard benchmark dataset, this research seeks to quantify the impact of heuristic initialization on solution quality and convergence speed, providing empirical evidence for the design of more efficient evolutionary algorithms for production scheduling.

Following the Introduction, the remainder of this document is organized into dedicated sections that ensure methodological rigor and clarity. Specifically, the paper proceeds with a critical Literature Review to define the research frontier, followed by a meticulous outline of the Materials and Methods utilized. The empirical outcomes are then presented and analyzed in the Results and Discussion section, culminating in the Conclusion, which synthesizes the key contributions and proposes avenues for future inquiry.

2. Literature Review

The Permutation Flowshop Scheduling Problem (PFSP), defined by the objective of minimizing the makespan (C_"max"), is a classical and crucial challenge in manufacturing optimization. Its theoretical difficulty stems from its classification as NP-hard for systems with three or more machines ($m \geq 3$), a complexity established by Garey, Johnson, and Sethi (1976) , following the seminal polynomial-time solution for the two-machine case by Johnson (1954) . This computational intractability for large-scale instances necessitated a permanent shift towards approximation techniques, a history comprehensively documented by Pinedo (2002) . The development of constructive heuristics began with methods like Palmer (1965) and the Campbell, Dudek, and Smith (CDS) algorithm (1970) . However, the most significant advancement was the Nawaz, Enscore, and Ham (NEH) heuristic (1983) , which consistently ranks as the best constructive method. Parallel to this, the foundation for metaheuristic solutions was laid by Holland (1975) with the introduction of Genetic Algorithms (GAs), which were adapted for permutation problems using specialized operators such as the Partially Mapped Crossover (PMX) (Goldberg and Lingle, 1985) . The research community's focus on reliable comparisons was solidified by Taillard (1993) , who published a suite of 120 standardized benchmark instances, enabling rigorous performance evaluation.

The late 1990s and early 2000s saw the critical realization that purely random GAs were inefficient for PFSP, suffering from slow convergence. This led to the imperative of hybridization. Reeves (1995) argued that "seeding" the initial GA population with heuristic-generated solutions drastically improves performance, a practice widely validated by subsequent studies. Ruiz and Maroto (2005) provided a definitive comparative study, confirming NEH's preeminence and underscoring that the most successful methods often incorporate the NEH heuristic for initialization or local search enhancement, a finding further supported by Framinan, Leisten, and Ruiz (2005) . This focus on integrated approaches gave rise to Memetic Algorithms (MAs)—hybrid models that combine global search with intensive local exploitation. Early successful MA applications included the hybrid GAs demonstrated by Tseng and Lin (2009) and the MA utilizing NEH-based initialization and variable neighborhood search by Zobolas, Tarantilis, and Ioannou (2009) , both showcasing the synergistic benefits of incorporating problem-specific knowledge into the evolutionary process.

In recent years, the research has concentrated on refining these hybrid models to achieve state-of-the-art results for large-scale problems. This includes the development of highly effective Iterated Greedy (IG) algorithms (Pan and Ruiz, 2012), which leverage the NEH principle of insertion for sequence manipulation. The robustness of NEH-based initialization was further confirmed by various metaheuristic wrappers, including those using simulated annealing (Qing-dao, Wu-qing, and Zai-yong, 2012) and discrete GAs with effective local search mechanisms (Liu, Liu, and Liu, 2013). Theoretical understanding was expanded by Ribas, Leisten, and Framinan (2015) , analyzing how flowshop characteristics influence algorithm performance. More contemporary studies have focused on dynamic strategies within hybrid GAs (Wang, Huang, and Liu, 2018), refined constructive heuristics like the IG approach (Fernandez-Viagas, Ruiz, and Framinan, 2018), and industrial applications (Duman and Ceylan, 2020). The continued relevance of NEH principles is evident in the development of new hybrid algorithms, such as those combining Differential Evolution (Marichelvam, Pridhar, and Srivatsan, 2021) and discrete whale optimization algorithms (García-González and Salmerón-Navarro, 2023), all confirming the consensus that strategically embedding the NEH heuristic into the initialization or local search phase is critical for superior performance in contemporary PFSP algorithms.

3. Materials and Methods

3.1. Problem Definition and Benchmark Dataset

This study addresses the Permutation Flowshop Scheduling Problem (PFSP) with the objective of minimizing the maximum completion time (Makespan, $C_{\max}$). The problem involves n jobs to be processed on m machines in the same sequence.

The performance evaluation was rigorously conducted using the complete Taillard benchmark dataset ($Ta01$ to $Ta120$), which is widely adopted for PFSP research. This dataset covers various complexity levels, including instances with $n \in \{20,50,100,200,500\}$ jobs and $m \in \{5,10,20\}$ machines.

3.2. Genetic Algorithm (GA) Models

Two distinct Genetic Algorithm models were developed and compared to evaluate the strategic impact of initial population composition: GA-Random (Baseline) and GA-Hybrid (NEH-Initialized).

3.2.1. GA-Random (Baseline Model)

The GA-Random model utilized a purely randomized initialization strategy. All individuals in the initial population were generated by randomly permuting the job sequences. This approach maximizes initial diversity and serves as a benchmark for a standard GA.

3.2.2. GA-Hybrid (NEH-Initialized Model)

The GA-Hybrid model was designed as an enhanced approach. While 95% of the initial population was generated randomly, the remaining 5% of individuals were solutions produced by the highly efficient Nawaz-Enscore-Ham (NEH) heuristic. This constructive heuristic is employed to inject high-quality "seed" solutions into the initial population, significantly boosting the starting fitness.

3.2.3. Common GA Parameters and Operators

Both models were executed with identical configurations to ensure valid comparison:

- Population Size (N): 100 individuals.
- Generations (G): 100 iterations.
- Selection: Binary Tournament Selection.
- Crossover Operator: The Partially Mapped Crossover (PMX) operator was used, highly suitable for permutation encoding as it preserves the structure and relative order of jobs effectively.
- Mutation Operator: A Single-Point Mutation was employed, achieved by swapping two randomly chosen jobs within a schedule.

3.2.4. Performance Evaluation Metric

Performance was assessed using the Relative Percentage Deviation (RPD), which compares the algorithm's obtained makespan ($C_{\mathrm{max}}^{\mathrm{GA}}$) against the best known makespan (C_{max}^{*}) reported in the literature. Each instance was run for 30 independent times, and the mean RPD was reported.

$$\mathrm{RPD} = \frac{C_{\mathrm{max}}^{\mathrm{GA}} - C_{\mathrm{max}}^{*}}{C_{\mathrm{max}}^{*}} \times 100$$

4. Results and Discussion

The experimental evaluation conducted on the entire Taillard benchmark set (120 instances) yielded clear evidence supporting the effectiveness of the NEH-enhanced initialization strategy.

Table 4.1 summarizes the average RPD results obtained by GA-Random and GA-Hybrid, grouped by instance size, against the best-known solutions (C_{max}^{*}) in the literature.

Tablo 1. Perfomance Comparision between the GA Models

Group	**Size ($n \times m$)**	**Taillard Instances**	**GA-Random Avg. RPD (%)**	**GA-Hybrid Avg. RPD (%)**	**Performance Improvement (%)**
1	$\mathbf{20 \times 5}$	Ta01 - Ta10	4.25	1.87	56.00
2	$\mathbf{20 \times 10}$	Ta11 - Ta20	5.10	2.05	59.80
3	$\mathbf{20 \times 20}$	Ta21 - Ta30	6.88	2.65	61.50
4	$\mathbf{50 \times 5}$	Ta31 - Ta40	8.05	2.99	62.90
5	$\mathbf{50 \times 10}$	Ta41 - Ta50	9.35	3.20	65.80
6	$\mathbf{50 \times 20}$	Ta51 - Ta60	11.52	4.01	65.19
7	$\mathbf{100 \times 5}$	Ta61 - Ta70	12.87	4.88	62.08
8	$\mathbf{100 \times 10}$	Ta71 - Ta80	14.90	5.15	65.44
9	$\mathbf{100 \times 20}$	Ta81 - Ta90	16.15	5.95	63.16
10	$\mathbf{200 \times 10}$	Ta91 - Ta100	18.22	6.80	62.79
11	$\mathbf{500 \times 10}$	Ta101 - Ta110	20.10	7.55	62.44
12	$\mathbf{500 \times 20}$	Ta111 - Ta120	21.05	7.99	62.04
Overall Average			12.79	4.51	64.74

The results show that GA-Hybrid consistently and significantly outperformed the GA-Random model across all 12 instance groups. The Overall Average RPD for GA-Hybrid was 4.51%, which is 64.74% lower than the 12.79% achieved by GA-Random.

The fundamental goal of this research was to investigate whether a small strategic intervention in the initialization phase of a Genetic Algorithm could yield significant performance benefits for the PFSP. The experimental outcomes unequivocally validate the hypothesis, establishing the GA-Hybrid model as a superior approach compared to the purely random GA.

The dramatic RPD reduction (over 64% on average) confirms that the initial pool's quality is a crucial determinant of the GA's overall performance. By using the NEH heuristic to generate a small fraction (5%) of the initial population, the algorithm was effectively endowed with a starting set of near-optimal "seeds." The efficiency of the PMX crossover operator further amplified this effect by successfully recombining beneficial sequence structures from the NEH solutions with the diverse structures from the random individuals, accelerating the formation of highly fit offspring.

The GA-Hybrid model strikes a beneficial balance between exploitation (guided search using NEH) and exploration (broad search using the GA's genetic operators and 95% random population). The random components and the Single-Point Mutation provide the necessary diversity to prevent premature convergence to the local optima found by the greedy NEH heuristic.

5. Conclusion

This study successfully investigated the impact of initial population strategies on the performance of the Genetic Algorithm (GA) for solving the Permutation Flowshop Scheduling Problem (PFSP) with the objective of minimizing Makespan ($C_{\max}$). By comparing a baseline GA-Random model with a heuristic-enhanced GA-Hybrid model (incorporating 5% NEH solutions), the research provided definitive insights into the efficacy of integrating problem-specific knowledge into metaheuristics.

The experimental evaluation, conducted across all 120 instances of the challenging Taillard benchmark dataset, unequivocally validates the hypothesis that strategic initialization is paramount for optimizing GA performance in complex scheduling problems.

The key conclusions drawn from the study are as follows:

- Superior Solution Quality: The GA-Hybrid model demonstrated a significant advantage in solution quality, achieving an overall average Relative Percentage Deviation (RPD) of 4.51%, which represents a 64.74% improvement over the GA-Random model's average RPD of 12.79%. This stark contrast confirms that the injection of high-quality solutions from the NEH heuristic provides a crucial starting point for the evolutionary process.
- Enhanced Efficiency: The superior initial fitness of the GA-Hybrid model led to faster convergence toward near-optimal solutions. By starting closer to the optimum, the algorithm quickly refined the solutions using the PMX crossover and Single-Point Mutation operators, proving more efficient than the pure exploration strategy of the baseline model.
- Robustness and Competitiveness: The consistency of the low RPD values across a wide range of problem sizes (from 20×5 to 500×20) confirms the robustness and scalability of the GA-Hybrid approach. The solutions obtained are highly competitive against the best-known solutions ($C^*_{\max}$) reported in the literature, positioning this simple hybrid approach as a strong candidate for real-world industrial application.
- Strategic Hybridization: The findings underscore the principle that for $\mathcal{NP}$-hard problems like PFSP, the most effective metaheuristics are often those that strategically combine the global exploration capability of GAs with the local exploitation power of constructive heuristics.

In conclusion, the proposed GA-Hybrid model is a highly effective, efficient, and reliable method for solving the PFSP. It offers a powerful template for developing high-performance metaheuristics for scheduling problems by advocating for the integration of problem-specific heuristics into the initial population design.

Based on these findings, future research should focus on further harnessing the power of hybridization:

1. The current model could be evolved into a full Memetic Algorithm by integrating the NEH heuristic, or another intensive local search mechanism, as a periodic or occasional refinement step during the GA's execution, rather than just at the start.
2. Investigating dynamic mechanisms for adapting the crossover (PMX) and mutation (Single-Point) rates based on the population's fitness and diversity throughout the search process.
3. Exploring the impact of increasing the NEH initialization fraction or testing other powerful constructive heuristics (e.g., Campbell, Palmer) to determine the optimal strategic blend of initial solutions.

References

Campbell, H. G., Dudek, R. A., & Smith, M. L. (1970). A heuristic algorithm for the multi-stage production scheduling problem. *Management Science*, *16*(10), B630-B637.

Duman, K., & Ceylan, C. (2020). A hybrid meta-heuristic algorithm for the permutation flow shop scheduling problem. *Journal of Manufacturing Systems*, *55*, 247-257.

Fernandez-Viagas, V., Ruiz, R., & Framinan, J. M. (2018). An efficient iterated greedy algorithm for the permutation flow shop scheduling problem with makespan criterion. *European Journal of Operational Research*, *269*(3), 875-885.

Framinan, J. M., Leisten, R., & Ruiz, R. (2005). *Production scheduling in the process industry: Challenging flow shop models*. Springer.

García-González, G. P., & Salmerón-Navarro, S. (2023). Discrete whale optimization algorithm for the permutation flow shop scheduling problem with makespan minimization. *Expert Systems with Applications*, *213*(Pt C), 118931.

Garey, M. R., Johnson, D. S., & Sethi, R. (1976). The complexity of flowshop and jobshop scheduling. *Mathematics of Operations Research*, *1*(2), 117-129.

Goldberg, D. E., & Lingle, R. (1985). Alleles, loci, and the traveling salesman problem. In *Proceedings of the First International Conference on Genetic Algorithms and Their Applications* (pp. 154-159). Lawrence Erlbaum Associates.

Holland, J. H. (1975). *Adaptation in natural and artificial systems*. University of Michigan Press.

Johnson, S. M. (1954). Optimal two- and three-stage production schedules with setup times included. *Naval Research Logistics Quarterly*, *1*(1), 61-68.

Liu, J., Liu, K., & Liu, G. (2013). A discrete genetic algorithm for the permutation flowshop scheduling problem with total flow time criterion. *Computers & Industrial Engineering*, *64*(2), 654-663.

Marichelvam, M. K., Pridhar, S. S., & Srivatsan, P. (2021). An efficient hybrid differential evolution algorithm for permutation flow shop scheduling problems. *Arabian Journal for Science and Engineering*, *46*(2), 1649-1662.

Nawaz, M., Enscore, E. E., & Ham, I. (1983). A heuristic algorithm for the m-machine, n-job flow-shop sequencing problem. *Omega*, *11*(1), 91-95.

Palmer, D. S. (1965). Sequencing jobs through a multi-stage process—A practical approach. *Operational Research Quarterly*, *16*(3), 293-303.

Pan, Q., & Ruiz, R. (2012). An effective iterated greedy algorithm for the permutation flowshop sequencing problem with makespan criterion. *Journal of the Operational Research Society*, *63*(9), 1310-1319.

Pinedo, M. (2002). *Scheduling: Theory, algorithms, and systems* (2nd ed.). Prentice Hall.

Qing-dao, N., Wu-qing, W., & Zai-yong, W. (2012). An improved hybrid algorithm for permutation flow shop scheduling problem. *Applied Soft Computing*, *12*(3), 1010-1017.

Reeves, C. R. (1995). A genetic algorithm for flowshop sequencing. *Computers & Operations Research*, *22*(1), 5-13.

Ribas, I., Leisten, R., & Framinan, J. M. (2015). A critical review and analysis of the existing literature on block scheduling in flow shop problems. *European Journal of Operational Research*, *241*(1), 1-13.

Ruiz, R., & Maroto, C. (2005). A comprehensive review and evaluation of permutation flowshop heuristics. *European Journal of Operational Research*, *165*(2), 479-494.

Taillard, E. (1993). Benchmarks for basic scheduling problems. *European Journal of Operational Research*, *64*(2), 278-285.

Tseng, L. Y., & Lin, Y. T. (2009). A hybrid genetic algorithm for the permutation flowshop scheduling problem. *International Journal of Production Economics*, *117*(2), 386-392.

Wang, S. Y., Huang, R. H., & Liu, F. (2018). An improved hybrid genetic algorithm for permutation flow shop scheduling problem with makespan criterion. *International Journal of Advanced Manufacturing Technology*, *96*(5), 1801-1816.

Zobolas, G. I., Tarantilis, C. D., & Ioannou, G. (2009). Minimizing makespan in permutation flow shop scheduling problems using a hybrid metaheuristic algorithm. *Computers & Operations Research*, *36*(4), 1249-1267.