Review

Not peer-reviewed version

A Review on Reinforcement Learning in Production Scheduling: An Inferential Perspective

Vladimir Modrak * , Sudhakarapandian Ranjitharamasamy , Arunmozhi Balamurugan , Zuzana Soltysova

Posted Date: 7 May 2024

doi: 10.20944/preprints202405.0335.v1

Keywords: bibliometric analysis; production scheduling; reinforcement learning

Review

A Review on Reinforcement Learning in Production Scheduling: An Inferential Perspective

Vladimir Modrak [1,*], Sudhakarapandian Ranjitharamasamy [2], Arunmozhi Balamurugan [2] and Zuzana Soltysova [1]

[1] Faculty of Manufacturing Technologies, Technical University of Kosice, Slovakia, vladimir.modrak@tuke.sk, zuzana.soltysova@tuke.sk
[2] School of Mechanical Engineering, Vellore Institute of Technology, Vellore, India, sudhakarapandian.r@vit.ac.in
* Correspondence: vladimir.modrak@tuke.sk

Abstract: In this study, a comprehensive review on production scheduling based on reinforcement learning (RL) techniques using bibliometric analysis has been carried out. The aim of this work is, among other things, to point out the growing interest in this domain and to outline the influence of RL as a type of machine learning on production scheduling. To achieve this, the paper explores production scheduling using RL by investigating the descriptive metadata of all pertinent publications contained in the Web of Science database. The study focuses on a wide spectrum of publications spanning the years between 1998 and 2023. The findings and recommendations of this study can serve as new insights for future research endeavors in the realm of production scheduling using RL techniques.

Keywords: bibliometric analysis; production scheduling; reinforcement learning

1. Introduction

Production scheduling is considered as one of the most critical element of manufacturing management in aligning production activities with business objectives, in ensuring a smooth flow of goods resources, and in supporting company's ability to remain competitive in the marketplace. Scheduling algorithms have long been a subject of extensive research in various interdisciplinary domains, such as industrial engineering, automation, and management science, due to their important role in enhancing production efficiency and effectiveness [1]. The production scheduling tasks can be solved using three main types of step-by-step procedures such as exact algorithms, heuristic algorithms, and meta-heuristic algorithms [2, 3]. Although an exact algorithm can theoretically guarantee the optimum solution, the NP-hardness of major problems makes them impossible to address effectively and efficiently [4]. Heuristics use a set of rules to create scheduling solutions quickly and effectively without consideration of global optimization. Furthermore, the creation of rules is heavily reliant on a thorough comprehension of the particulars of the situation [5–8]. Whereas meta-heuristics can produce good scheduling solutions in a reasonable amount of computing time, the way search operators create them significantly depends on the specific situation at hand [9 -13]. In addition, the iterative search process poses challenges in terms of time consumption and applicability in real-time scenarios when dealing with large-scale problems. Reinforcement learning is a subfield within the broader domain of machine learning. RL is considered one of the most perspective approaches for robust cooperative scheduling, which allows production managers to interact with a complex manufacturing environment, learn from previous experience, and select optimal decisions. It involves the process of an agent autonomously selecting and executing actions to accomplish a given task. The agent learns via experience and aims to maximize the rewards it receives in certain scenarios. The primary goal of RL is to optimize the cumulative reward obtained by an agent through the evaluation and selection of actions within a dynamic environment [14–20]. The most current development in artificial intelligence technology has allowed successful application

 doi:10.20944/preprints202405.0335.v1

of RL in sequential decision-making problems with multiple objectives which are usable for robot scheduling and control [21, 22]. The research on production scheduling using RL since 1998 has been evolving as advancing optimization techniques compared to metaheuristics. RL significantly improves the computational efficiency of addressing scheduling problems. Numerous studies of RL on production scheduling have been undertaken since its inception in 1998, establishing a substantial and valuable foundation (see, e.g., [23–27]). For this reason, it appears useful to employ bibliometric analysis to illustrate the current advancements and trends in this particular domain. Moreover, gaining a deeper understanding of prominent authors and collaborative countries in the given research field can serve as a source of motivation for academics and practitioners in their forthcoming endeavors.

Accordingly, the objective of this study is to gain a comprehensive impression of the existing research and emerging developments in RL applied to production scheduling. The main research question that this paper will address is: What are the main emerging research areas, the most important methodologies, and typical implementation domains within RL applied to production scheduling?

For this purpose, a consistent methodological framework for assessing bibliometric data has been designed, whose structure is described in subsequent section of this paper.

2. Materials and Methods

The bibliometric approach employed here presents a quantitative instrument for monitoring and representing scientific progress by examining and visualizing scientific information. The growing acceptance of bibliometric methods in several academic disciplines indicates that their use brings expected effects [28–31]. The present investigation is conducted utilizing the procedure comprising of five coherent phases, as depicted in Figure 1.

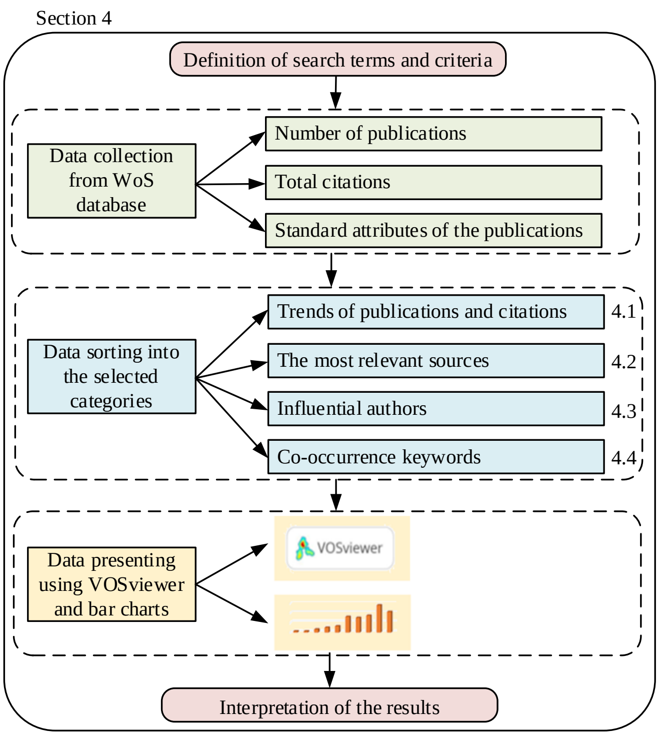

Figure 1. Research Methodology for inferential perspective.

The bibliometric analysis of RL in the context of production scheduling is focusing here on its theoretical foundations and practical implications, between the years 1998 and 2023. The first two phases involve the collection of bibliographic data, which was gathered via a Web of Science (WoS) database - Core Collection of Clarivate Analytics. The search was restricted to this database due to its prominence as a comprehensive repository of scientific literature and its frequent utilization in

 doi:10.20944/preprints202405.0335.v1

academic assessments. The inclusion criteria encompassed publications that contain the terms "reinforcement learning" and "manufacturing" and "scheduling" using search term "Topic" and document types "Article", "Review Article", and "Early Access". In the next two phase of the data sorting and presentation, the Microsoft Excel and VOSviewer software were employed to extract the necessary information, such as the annual scientific outputs, most relevant sources, most cited author, and keyword co-occurrence.

Moreover, this research follows an inferential approach, where the sample of population is explored to determine its characteristics [32, 33]. Moreover, the inferential concept of scientific representation proposed by Suárez [34] was applied here to formulate research outputs. Its essence is to employ alternative reasoning to reach results that differ from the isomorphic view of scientific representation in the sense that empirical knowledge plays an important role in inductive reasoning [35, 36].

Prior to the bibliometric analysis, a literature review on the recent development and research in the field is carried out in the following section.

3. Literature Review

RL techniques have gained significant popularity, hence affirming the level of interest in agent-based models. The objective of previous studies has focused on the application of deep reinforcement learning (DRL) in the domain of job shop scheduling. In recent years, there has been notable success in the application of RL to address many combinatorial optimization issues, including the Vehicle Routing Problem as well as the Traveling Salesman Problem [37–42]. Considering that a production scheduling task may be conceptualized as the environment within the framework of RL, an agent can acquire a policy of well-designed actions and states, and engage the extensive offline training through interaction with the environment. This innovative concept offers a fresh perspective on addressing scheduling challenges, particularly those characterized by uncertainty and dynamism, and necessitating stringent real-time constraints, as in the case of in a dynamic job shop scheduling problem [22, 43–47].

In the context of production scheduling problems, value-based approaches are commonly utilized RL algorithms. These algorithms, ranked from most to least frequently employed, include Q-learning, temporal difference TD(λ) algorithm, SARSA, ARL, informed Q-learning, dual Q-learning, approximate Q-learning, gradient descent TD(λ) algorithm, revenue sharing, Q-III learning, relational RL, relaxed SMART, and TD(λ)-learning. In the field of DRL, many value-based approaches have been employed, such as DQN (Deep Q-Learning Networks), loosely-coupled DRL, multiclass DQN, and the Q-network algorithm [48–58].

Qu et al. [25] applied multi-agent approximate Q-learning to address the issue of conducting numerical experiments to showcase the efficacy of the methodology in both static and dynamic environments, as well as in various scenarios of a flow shop. The objective was to develop and execute optimized manufacturing scheduling in a manufacturing setting, considering realistic interactions among a workforce's skill set and adaptive machines. Luo [59] employs Deep Reinforcement Learning (DRL) techniques to address the dynamic flexible job shop scheduling problem. The author is specifically focusing on scenarios including new work insertions, while his primary objective is to minimize the overall tardiness of the schedule. Luo et al. [60] were the first to employ hierarchical multi-agent proximal policy optimization (HMAPPO) as a solution for the constantly changing partial-no-wait multi-objective flexible job-shop problem (MOFJSP) with new job insertions and machine breakdowns.

Wang et al. [37] discusses that the production scheduling process involves manufacturing of several types of items using a hybrid production pattern that utilizes Multi-Agent Deep Reinforcement Learning (MADRL) model. Popper et al. [43] suggested that MADRL can be employed to optimize flexible production plants in a reactive manner, taking several criteria into account such as efficient and ecological target values. Du et al. [61] utilized the DQN algorithm to address the flexible task shop scheduling problem (FJSP) in the presence of varying processing rates, setup time, idle time, and task transportation. The method incorporates 34 state indicators and 9 actions to

 doi:10.20944/preprints202405.0335.v1

optimize the exploitation capabilities of the DQN component. Additionally, a problem-driven EDA component is integrated into the algorithm to augment the exploration capabilities.

Li et al. [62] utilizes a DRL approach for discrete flexible job shop problem with inter-tool reusability (DFJSP-ITR). To address the multi-objective optimization problem of minimizing the combined makespan and total energy consumption, a set of 26 generic state features, a genetic programming-based action space, and a reward function are proposed. Zhou et al. [63] suggests the utilization of online scheduling techniques that rely on RL with composite reward functions. This approach aims to enhance the efficiency and resilience of industrial systems. A novel algorithm for production scheduling based on deep reinforcement learning in complex job shops was applied in work of [18]. The authors in this article also presented the main advantages of this application such as better flexibility, global transparency, and global optimization. There are also other authors who paid attention to deep reinforcement learning, and developed diverse learning algorithms able to develop complex strategies and optimize production scheduling. For example, Luo et al. [26] developed on-line rescheduling framework for the dynamic multi-objective flexible job shop scheduling problem with new job insertions. This framework offers several advantages over others, e.g., it allows to minimize the total tardiness or maximize machine utilization rate. Another authors [27] applied deep Q-learning method to solve dynamic scheduling in smart manufacturing environment. Dynamic scheduling method on deep reinforcement learning has been also proposed in work [64], where proximal policy optimization was adopted to find the optimal policy of the scheduling in a real-world environment. It is also worth mentioning here that Wang and Usher [65] investigated in their work the application of widely used reinforcement Q-learning algorithm to be applied for agent-based production scheduling.

4. Research Findings and Results Description

Using the above-specified keywords and search criteria, 220 co-authored articles were found. The primary research areas that receive significant attention in the articles include Manufacturing Engineering, Industrial Engineering, Computer Science Interdisciplinary Applications, Operations Research Management Science, Electrical Electronic Engineering, Computer Science and Artificial Intelligence, Automation Control Systems, Computer Science and Information Systems, and other topics. The 220 publications in our sample were categorized into 34 distinct research topics. The eight primary research areas, along with their article distribution are displayed in Table 1.

Table 1. Associated research disciplines.

Research Areas	Number of publications	Percentage
Manufacturing Engineering	66	12,4 %
Industrial Engineering	59	11 %
Computer Science Interdisciplinary Applications	51	9,6 %
Operations Research Management Science	47	8,8 %
Electrical Electronic Engineering	43	8 %
Computer Science and Artificial Intelligence	41	7,7 %
Automation Control Systems	36	6,7 %
Computer Science and Information Systems	24	4,5 %
Others	167	31,3 %

4.1. Trends of Publications and Citations

Numbers of published articles and their citations usually provide sufficiently reliable information to anticipate further development of examined research domain. Considering the numbers of publications and citations in the field of production planning using learning algorithms keeping around 26 years of data, the trend analysis graph has been derived. For this purpose, the same search terms and keywords were applied as in case of identification of major research categories, but the types of documents were extended to all the types. The reason of changing document types was to find out the initial research initiatives in this domain. The same search conditions were applied in the rest of the paper to include the larger sample of publications for the purpose of the investigation. The annual distribution of publications (out of the total 297 items) and their citations from the same database during the period from 1998 to 2023 is illustrated in Figure 2.

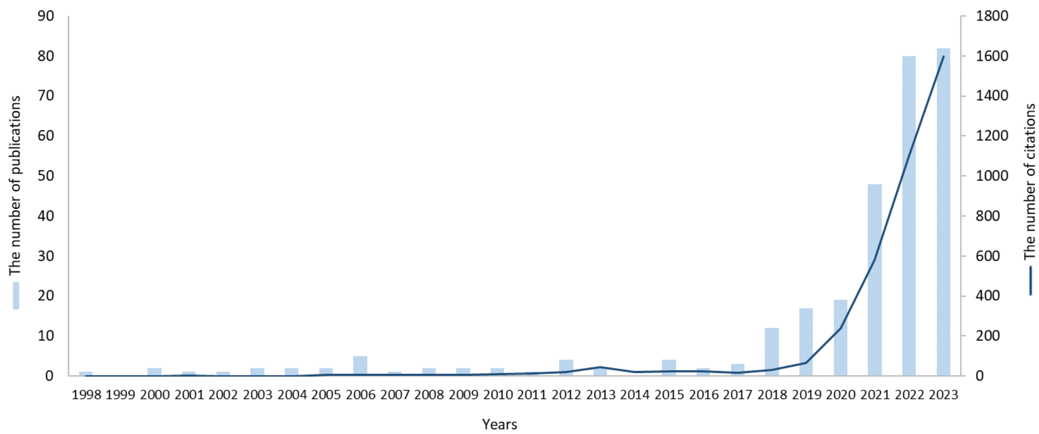

Figure 2. Publication and citation trends for RL on production scheduling field from 1998 to 2023.

An examination of the yearly scientific outputs between 1998 and 2017 amply demonstrated a relatively stable low number of articles published annually. Throughout these two decenniums, the need for RL in production scheduling in real conditions apparently did not appear. During the 2019–2023 era, there was a noticeable rise in the number of publications that were registered in the most recognized database for peer reviewed content. This phenomenon can be primarily attributed to the advancements in artificial intelligence. It is noteworthy to emphasize that if this exponential trend of increasing the number of publications continues, then one can anticipate that during the next decade the importance of RL in manufacturing scheduling will significantly increase.

4.2. Most Relevant Sources

An identification of the most relevant publications from an initial dataset presents common approach in bibliometric research since such sources usually publish influential research that attracts widespread interest. As a rule, the most productive journals have the greatest influence on the development of science in a particular field since they publish more articles and generate more citations [66–68]. As for as the relevancy of literature in the explored field, the top ten journals that have published the most articles, are identified here. Also, the top ten most cited scientific journals are mentioned in this sub-section. Figures 3 and 4 categorize journals according to these two criteria in order to show that they represent journals that exhibit the utmost relevance to RL in production scheduling.

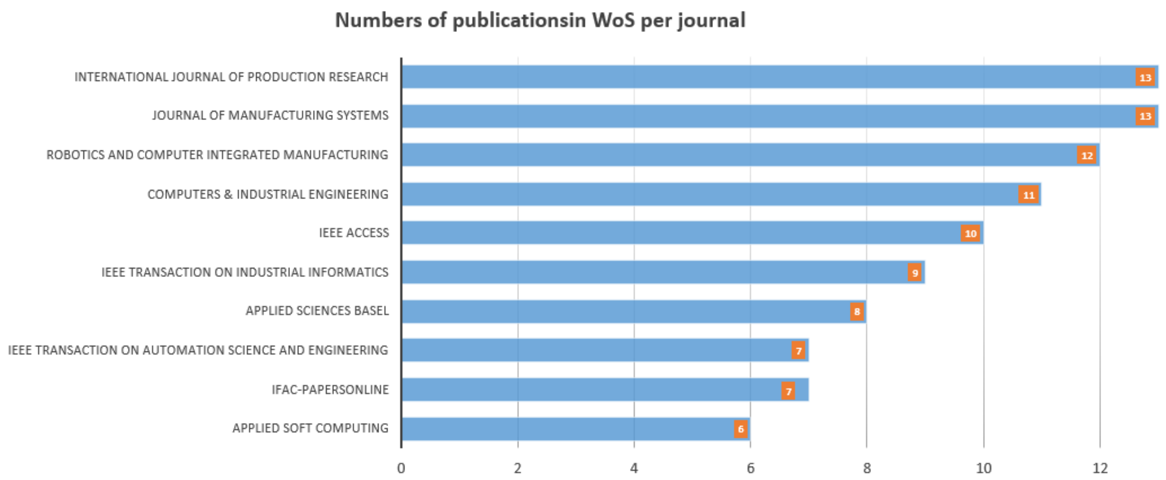

Figure 3. Top ten published journals for RL on production scheduling field from 1998 to 2023.

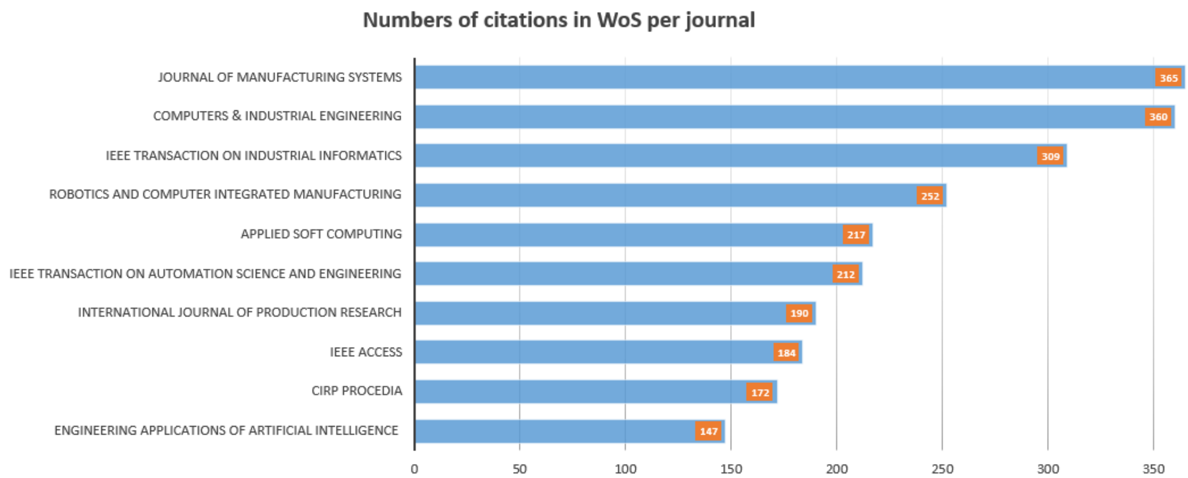

Figure 4. Top ten cited journals for RL on production scheduling field from 1998 to 2023.

The International Journal of Production Research, Journal of Manufacturing Systems and Computers & Industrial Engineering along with the nine others are considered the most relevant and respected scientific publications in this field. It can also be noted that all twelve journals listed in Figures 3 and 4, regardless of the results of the metrics used, can be empirically ranked as widely recognized for disseminating advanced research on reinforcement learning applied to production scheduling. In addition, their scientific rigor is also indicated by the fact that out of the twelve identified journals, eleven met high standards for quality and impact, and are indexed for Current Content Connect journals with a verifiable impact on steering research practices and behaviors [69, 70].

4.3. Most Cited Authors

Since the1998s, many authors have made significant contributions to the development of this field. In this sub-section, the intention is to present some of those authors who made significant intellectual contributions to the research. The analysis of the most cited authors was performed using data from WoS - Core Collection database. In addition, co-citation analysis was carried out for each publication source (out of the total 297 items) to reveal the network between the studies. For this purpose, VOSviewer software [71] has been applied. To obtain relevant information and clear graphic representation of complex relations, the following filters were employed. Filter 1: Maximum number of authors per document – 10; Filter 2: Minimum number of documents of an author – 3; Filter 3: Minimum number of citations of an author – 10. Moreover, the full counting method has been applied meaning that the publications that have co-authors from multiple countries are counted as a full

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 May 2024 doi:10.20944/preprints202405.0335.v1

publication for each of those countries. The co-citation network of the selected sample of scholars using these settings is visualized in Figure 5.

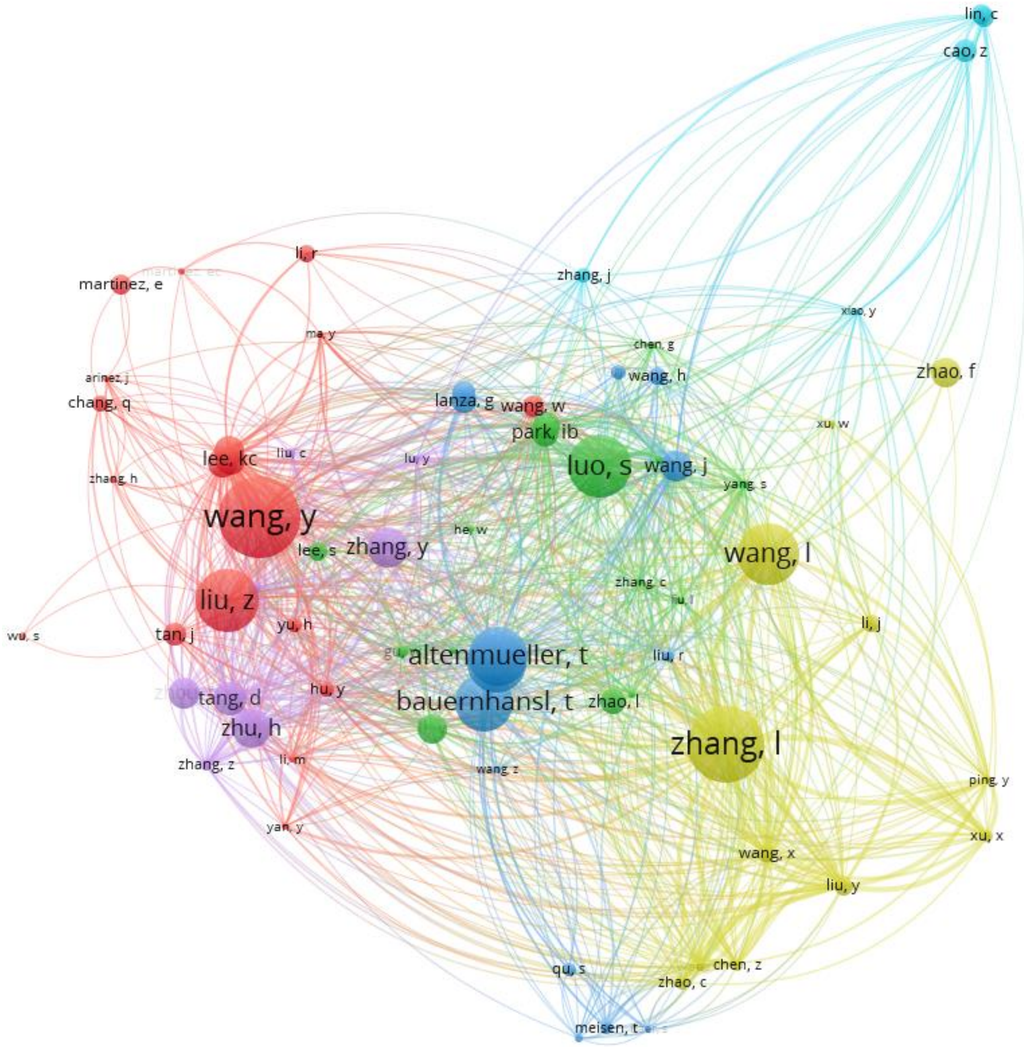

Figure 5. Co-citation network of the authors that have made significant contributions to the development of this field from 1998 to 2023.

This co-citation network map shows, among other things, scholars that have received the highest number of citations in the last 26 years. Of the 744 cited authors, 73 meet the above mentioned criteria. Each scholar from each included publication is represented by a node in this network. The size of each node indicates frequency of citation of the subject's scholarly works. An edge is drawn between two nodes if the two scholars were cited by a common document. To rank influential scholars in the given domain based on their citation rates, the ten most cited authors were selected. Those ten authors are listed in Table 2.

Table 2. Most influential authors.

Name	Number of citations	Country	Name	Number of citations	Country
Wang, Y.	347	Taiwan	Altenmueller, T.	252	Germany
Zhang, L.	325	China	Waschenck, B.	252	Germany
Luo, S.	272	China	Bauernhansl, T.	242	Germany
Liu, Z.	265	China	Tang, D.	161	China
Wang, L.	262	China	Zhu, H.	159	China

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 7 May 2024 doi:10.20944/preprints202405.0335.v1

4.4. Co-Occurrence Keywords

The goal of this part of the article is to identify areas of research in which the issue of production scheduling based on RL is of interest. For this purpose, the keyword analysis application has been employed to help separate important research themes which has received a high interest of researchers from less important ones. To obtain relevant and representative categories not including less significant ones, the following setting has been used: Minimum number of occurrence of keywords – 14. Based on this restriction, 20 keywords meet the threshold from the total 1 075 keywords. Those main keywords that were produced automatically from the titles in the papers on production scheduling based on RL along with their occurrence are shown in Figure 6.

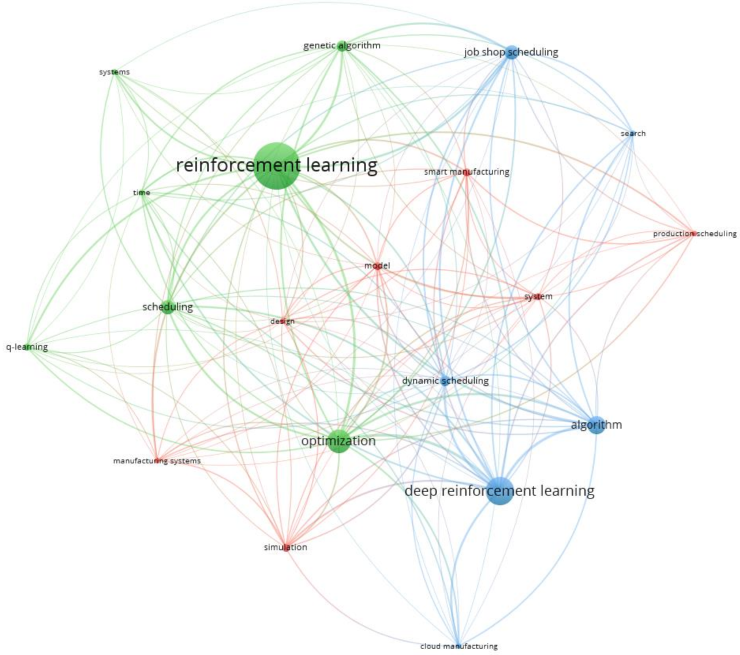

Figure 6. Co-occurrences keywords for RL on production scheduling field from 1998 to 2023.

As can be seen, topics can be divided into three clusters based on a computer algorithm, while each cluster has a different color as shown in Figure 6. When the colors are mixed, it means than the algorithm couldn't make clear differences between the clusters, and the clustering results does not meet the analytical target from the practical perspective. In this case, the co-occurrence keywords network map allows to identify relevant research topics and their mutual relationships. Based on the obtained bibliometric results extracted from VOSviewer as well as our empirical experiences, ten related topics that are very close to the explored research domain were identified as shown in Table 3.

Table 3. Most related methodologies and implementation areas of production scheduling based on RL.

Related terms/topics	Number of co-occurrence	Related terms/topics	Number of co-occurrence
Deep reinforcement learning	69	Dynamic scheduling	24
Optimization	58	Scheduling	20
Algorithm	44	Simulation	19
Job shop scheduling	35	Model	19
Genetic algorithm	26	Smart manufacturing	19

From this table one can easily find that the related terms closely correspond to the associated research disciplines that are identified in Table 1. Moreover, the results of this subsection seems to be consistent with practical observations and extant literature. For example, in recent years, there has been evidently increased interest in using deep reinforcement learning for optimization of real-time job scheduling tasks [72- 75]. This fact can be correlated with the continuing trend of mass customization in the production of consumer goods [76, 77]. As known, for mass customization is characteristic to meet dynamically changing user requirements in time, while customized products need to be completed by different deadlines. Accordingly, efficient real-time job-scheduling algorithms based on DRL become essential. The second important method that is ranked among the top 10 co-occurrences keywords is simulation. Its importance results from the fact that only operation times are required to produce an optimal schedule. But in some practical application other components (such as, e.g., dispatch time and suspend time) may play important role. Therefore, validation of proposed solutions through simulation is often used (see, e.g., [78–81]). The next important co-occurred keyword in Table 3 - smart manufacturing represents the implementation domain of production scheduling based on RL. Despite the fact that smart manufacturing has also become a buzzword, which also has its drawbacks, this conception is gradually being established as the new manufacturing paradigm. On the other hand, complexity of smart manufacturing network infrastructures becomes higher and higher, and the uncertainty of such manufacturing environment becomes a serious problem [25]. These facts lead to the necessity of applying advanced dynamic planning solutions that also includes production scheduling using RL. This paragraph simultaneously answers to the main research question formulated in Section 1.

5. Conclusion

The examination of bibliometric findings frequently indicates that an increase in the quantity of published articles is associated with recognition of progressive trends in the subject. Also for this reason, a bibliometric analysis is becoming more and more beneficial in a variety of academic fields since it makes mapping scientific information and analyzing research development objective and repeatable. The use of this method enables us to identify the networks of scientific collaboration, to establish connections between novel study themes and research streams, as well as show the connections between citations, co-citations and published productivity in the field.

The existing body of research on production scheduling primarily consists of studies conducted within the domains of Engineering Manufacturing, Industrial Engineering, Operation Research, Computer Science, Artificial intelligence, and Engineering Electrical Electronic. These scientific disciplines, notably 'Engineering Manufacturing´ and ´Industrial Engineering´ have exerted significant influence on development of production scheduling based on RL.

In addition to the above mentioned findings it would be needed to focus on other challenges to be considered in the future such manufacturing process planning with integrated support for knowledge sharing, increasing demand for improvements in ubiquitous "smartness" in manufacturing processes including designing and implementing smart algorithms, and the need for robust scheduling tools for agile collaborative manufacturing systems.

Author Contributions: Conceptualization, V.M.; methodology, S.R. and A.B.; validation, Z.S.; formal analysis, V.M. and S.R; investigation, V.M, and A.B.; writing—original draft preparation, A.B., and V.M.; writing—review and editing, S.R. and V.S.; visualization, Z.S.; project administration, V.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been funded by the project SME 5.0 with funding received from the European Union's Horizon research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 101086487 and the KEGA project No. 044TUKE-4/2023 granted by the Ministry of Education of the Slovak Republic.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Pinedo, M. Planning and scheduling in manufacturing and services. Springer (New York) 2005.
2. Beheshti, Z.; Shamsuddin, S. M. H. A review of population-based meta-heuristic algorithms. *Int. j. adv. soft comput. Appl* 2013, Volume 5(1), 1-35.
3. Xhafa, F.; Abraham, A. (Eds.). Metaheuristics for scheduling in industrial and manufacturing applications, Volume 128, 2008, Springer.
4. Abdel-Kader, R. F. Particle swarm optimization for constrained instruction scheduling. *VLSI design* 2008.
5. Balamurugan, A.; Ranjitharamasamy, S.P. A Modified Heuristics for the Batch Size Optimization with Combined Time in a Mass-Customized Manufacturing System. *International Journal of Industrial Engineering: Theory Applications and Practice* 2023, Volume 30, 1090–1115.
6. Ghassemi Tari, F.; Olfat, L. Heuristic rules for tardiness problem in flow shop with intermediate due dates. *The International Journal of Advanced Manufacturing Technology* 2014, Volume 71, 381-393.
7. Modrak, V.; Pandian, R. S. Flow shop scheduling algorithm to minimize completion time for n-jobs m-machines problem. *Tehnički vjesnik* 2010, Volume 17(3), 273-278.
8. Thenarasu, M.; Rameshkumar, K.; Rousseau, J.; Anbuudayasankar, S.P. Development and analysis of priority decision rules using MCDM approach for a flexible job shop scheduling: A simulation study. *Simul Model Pract Theory*, 2022, Volume 114.
9. Pandian, S.; Modrak, V. Possibilities, obstacles and challenges of genetic algorithm in manufacturing cell formation. *Advanced Logistic systems* 2009, Volume 3(1), 63-70.
10. Abdulredha, M. N.; Bara'a, A. A.; Jabir, A. J. Heuristic and meta-heuristic optimization models for task scheduling in cloud-fog systems: A review. *Iraqi Journal for Electrical and Electronic Engineering* 2020, Volume 16(2), 103-112.
11. Modrak, V.; Pandian, R. S.; Semanco, P. Calibration of GA parameters for layout design optimization problems using design of experiments. *Applied Sciences* 2021, Volume 11(15), 6940.
12. Keshanchi, B., Souri, A., & Navimipour, N. J. (2017). An improved genetic algorithm for task scheduling in the cloud environments using the priority queues: formal verification, simulation, and statistical testing. Journal of Systems and Software, 124, 1-21.
13. Jans, R., & Degraeve, Z. (2007). Meta-heuristics for dynamic lot sizing: A review and comparison of solution approaches. European journal of operational research, 177(3), 1855-1875.
14. Han, B.A.; Yang, J.J. A deep reinforcement learning based solution for flexible job shop scheduling problem. *International Journal of Simulation Modelling* 2021, Volume 20, 375–386.
15. Shyalika, C.; Silva, T.; Karunananda, A. Reinforcement Learning in Dynamic Task Scheduling: A Review, 2020.
16. Wang, X.; Zhang, L.; Ren, L.; Xie, K.; Wang, K.; Ye, F.; Chen, Z. Brief Review on Applying Reinforcement Learning to Job Shop Scheduling Problems. *Journal of System Simulation* 2022, Volume 33(12), 2782-2791.
17. Dima, I. C.; Gabrara, J.; Modrak, V.; Piotr, P.; Popescu, C. Using the expert systems in the operational management of production. In 11th WSEAS International Conference on Mathematics and Computers in Business and Economics (MCBE'10)-Published by WSEAS Press 2010.
18. Waschneck, B.; Reichstaller, A.; Belzner, L.; Altenmüller, T.; Bauernhansl, T.; Knapp, A.; Kyek, A. Optimization of global production scheduling with deep reinforcement learning. *Procedia CIRP* 2018, 1264–1269. Elsevier B.V.
19. Yan, J.; Liu, Z.; Zhang, T.; Zhang, Y. Autonomous decision-making method of transportation process for flexible job shop scheduling problem based on reinforcement learning. Proceedings - 2021 International Conference on Machine Learning and Intelligent Systems Engineering, MLISE 2021, 234–238. Institute of Electrical and Electronics Engineers Inc.
20. Modrak, V.; Pandian, R. S. Operations management research and cellular manufacturing systems. Hershey, USA, 2010, IGI Global.
21. Huang, Z.; Liu, Q.; Zhu, F. Hierarchical reinforcement learning with adaptive scheduling for robot control. Engineering Applications of Artificial Intelligence 2023, 126, 107130.
22. Arviv, K.; Stern, H.; Edan, Y. Collaborative reinforcement learning for a two-robot job transfer flow-shop scheduling problem. International Journal of Production Research 2016, Volume 54(4), 1196-1209.
23. Wang, L.; Pan, Z.; Wang, J. A Review of Reinforcement Learning Based Intelligent Optimization for Manufacturing Scheduling. Complex System Modeling and Simulation 2021, Volume 1, 257–270.
24. Aydin, M.E.; Öztemel, E. Dynamic job-shop scheduling using reinforcement learning agents, 2000.
25. Qu, S.; Wang, J.; Govil, S.; Leckie, J.O. Optimized Adaptive Scheduling of a Manufacturing Process System with Multi-skill Workforce and Multiple Machine Types: An Ontology-based, Multi-agent Reinforcement Learning Approach. Procedia CIRP 2016, 55–60, Elsevier B.V.
26. Luo, S.; Zhang, L.; Fan, Y. Dynamic multi-objective scheduling for flexible job shop by deep reinforcement learning. Comput Ind Eng. 2021, Volume 159, 107489.
27. Zhou, L.; Zhang, L.; Horn, B.K.P. Deep reinforcement learning-based dynamic scheduling in smart manufacturing. Procedia CIRP 2020, 383–388, Elsevier B.V.

28. Broadus R. N. Toward a Definition of "Bibliometrics." *Scientometrics* 1987, Volume 12, 373–379.
29. Arunmozhi, B.; Sudhakarapandian, R.; Sultan Batcha, Y.; Rajay Vedaraj, I.S. An inferential analysis of stainless steel in additive manufacturing using bibliometric indicators. *Mater Today Proc.* 2023.
30. Randhawa, K., Wilden, R., & Hohberger, J. (2016). A bibliometric review of open innovation: Setting a research agenda. Journal of product innovation management, 33(6), 750-772.
31. van Raan, A. (1996). Advanced bibliometric methods as quantitative core of peer review based evaluation and foresight exercises. Scientometrics, 36(3), 397-420.
32. Brandom, R. B. Articulating reasons: An introduction to inferentialism. Harvard University Press, 2001.
33. Kothari, C. R. Research methodology: Methods and techniques. *New Age International* 2004.
34. Suárez, M. An inferential conception of scientific representation. *Philosophy of science* 2004, Volume 71(5), 767-779.
35. Contessa, G. Scientific representation, interpretation, and surrogative reasoning. *Philosophy of science* 2007, Volume 74(1), 48-68.
36. Govier, T. Problems in argument analysis and evaluation (Vol. 6). University of Windsor, 2018.
37. Wang, S.; Li, J.; Luo, Y. Smart Scheduling for Flexible and Hybrid Production with Multi-Agent Deep Reinforcement Learning. In: Proceedings of 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence, ICIBA 2021, 288–294. Institute of Electrical and Electronics Engineers Inc.
38. Tang, J.; Haddad, Y.; Salonitis, K. Reconfigurable manufacturing system scheduling: a deep reinforcement learning approach. *Procedia CIRP* 2022, 1198–1203. Elsevier B.V.
39. Shahrabi, J.; Adibi, M.A.; Mahootchi, M. A reinforcement learning approach to parameter estimation in dynamic job shop scheduling. *Comput Ind Eng.* 2017, Volume 110, 75–82.
40. Yang, J., You, X., Wu, G., Hassan, M. M., Almogren, A., & Guna, J. (2019). Application of reinforcement learning in UAV cluster task scheduling. Future generation computer systems, 95, 140-148.
41. Yuan, X., Pan, Y., Yang, J., Wang, W., & Huang, Z. (2021, February). Study on the application of reinforcement learning in the operation optimization of HVAC system. In Building Simulation (Vol. 14, pp. 75-87). Tsinghua University Press.
42. Kurinov, I., Orzechowski, G., Hämäläinen, P., & Mikkola, A. (2020). Automated excavator based on reinforcement learning and multibody system dynamics. IEEE access, 8, 213998-214006.
43. Popper, J.; Motsch, W.; David, A.; Petzsche, T.; Ruskowski, M. Utilizing multi-agent deep reinforcement learning for flexible job shop scheduling under sustainable viewpoints. International Conference on Electrical, Computer, Communications and Mechatronics Engineering 2021, ICECCME 2021. Institute of Electrical and Electronics Engineers Inc.
44. Xiong, H.; Fan, H.; Jiang, G.; Li, G. A simulation-based study of dispatching rules in a dynamic job shop scheduling problem with batch release and extended technical precedence constraints. *Eur J Oper Res.* 2017, Volume 257, 13–24.
45. Palacio, J.C.; Jiménez, Y.M.; Schietgat, L.; Doninck, B. Van; Nowé, A. A Q-Learning algorithm for flexible job shop scheduling in a real-world manufacturing scenario. *Procedia CIRP* 2022, 227–232. Elsevier B.V.
46. Chang, J.; Yu, D.; Hu, Y.; He, W.; Yu, H. Deep Reinforcement Learning for Dynamic Flexible Job Shop Scheduling with Random Job Arrival. *Processes* 2022, Volume 10.
47. Liu, R.; Piplani, R.; Toro, C. Deep reinforcement learning for dynamic scheduling of a flexible job shop. *Int J Prod Res*. 2022, Volume 60, 4049–4069.
48. Esteso, A.; Peidro, D.; Mula, J.; Díaz-Madroñero, M. Reinforcement learning applied to production planning and control, 2023.
49. Samsonov, V.; Kemmerling, M.; Paegert, M.; Lütticke, D.; Sauermann, F.; Gützlaff, A.; Schuh, G.; Meisen, T. Manufacturing control in job shop environments with reinforcement learning. ICAART 2021 - Proceedings of the 13th International Conference on Agents and Artificial Intelligence 2021, 589–597. SciTePress.
50. Cunha, B.; Madureira, A. M.; Fonseca, B.; Coelho, D. Deep reinforcement learning as a job shop scheduling solver: a literature review. In: Madureira, A.M.; Abraham, A.; Gandhi, N.; and Varela, M.L. (eds.) 18th International Conference on Hybrid Intelligent Systems (HIS 2018). Springer International Publishing, Cham 2018.
51. Wang, X.; Zhang, L.; Lin, T.; Zhao, C.; Wang, K.; Chen, Z. Solving job scheduling problems in a resource preemption environment with multi-agent reinforcement learning. *Robot Comput Integr Manuf*. 2022, Volume 77.
52. Oh, S.H.; Cho, Y.I.; Woo, J.H. Distributional reinforcement learning with the independent learners for flexible job shop scheduling problem with high variability. *J Comput Des Eng*. 2022, Volume 9, 1157–1174.
53. Zhang, Y.; Zhu, H.; Tang, D.; Zhou, T.; Gui, Y. Dynamic job shop scheduling based on deep reinforcement learning for multi-agent manufacturing systems. *Robot Comput Integr Manuf.* 2022, Volume 78.
54. Kuhnle, A.; May, M.C.; Schäfer, L.; Lanza, G. Explainable reinforcement learning in production control of job shop manufacturing system. *Int J Prod Res.* 2022, Volume 60, 5812–5834.

55. Liang, Y.; Sun, Z.; Song, T.; Chou, Q.; Fan, W.; Fan, J.; Rui, Y.; Zhou, Q.; Bai, J.; Yang, C.; Bai, P. Lenovo Schedules Laptop Manufacturing Using Deep Reinforcement Learning. *Interfaces (Providence)* 2022, Volume 52, 56–68.
56. Chen, Y., Guo, W., Liu, J., Shen, S., Lin, J., & Cui, D. (2024). A multi-setpoint cooling control approach for air-cooled data centers using the deep Q-network algorithm. Measurement and Control, 00202940231216543.
57. Théate, T., & Ernst, D. (2021). An application of deep reinforcement learning to algorithmic trading. Expert Systems with Applications, 173, 114632.
58. Sanaye, S., & Sarrafi, A. (2021). A novel energy management method based on Deep Q Network algorithm for low operating cost of an integrated hybrid system. Energy Reports, 7, 2647-2663.
59. Luo, S. Dynamic scheduling for flexible job shop with new job insertions by deep reinforcement learning. *Applied Soft Computing Journal* 2020, Volume 91.
60. Luo, S.; Zhang, L.; Fan, Y. Real-Time Scheduling for Dynamic Partial-No-Wait Multiobjective Flexible Job Shop by Deep Reinforcement Learning. IEEE *Transactions on Automation Science and Engineering* 2022, Volume 19, 3020–3038.
61. Du, Y.; Li, J.Q.; Chen, X.L.; Duan, P.Y.; Pan, Q.K. Knowledge-Based Reinforcement Learning and Estimation of Distribution Algorithm for Flexible Job Shop Scheduling Problem. *IEEE Trans Emerg Top Comput Intell*. 2023, Volume 7, 1036–1050.
62. Li, Y.; Gu, W.; Yuan, M.; Tang, Y. Real-time data-driven dynamic scheduling for flexible job shop with insufficient transportation resources using hybrid deep Q network. *Robot Comput Integr Manuf.* 2022, Volume 74.
63. Zhou, T.; Zhu, H.; Tang, D.; Liu, C.; Cai, Q.; Shi, W.; Gui, Y. Reinforcement learning for online optimization of job-shop scheduling in a smart manufacturing factory. *Advances in Mechanical Engineering* 2022, Volume 14.
64. Wang, L.; Hu, X.; Wang, Y.; Xu, S.; Ma, S.; Yang, K.; Wang, W. Dynamic job-shop scheduling in smart manufacturing using deep reinforcement learning. *Computer Networks* 2021, Volume 190, 107969.
65. Wang, Y.; Usher, J. M. Application of reinforcement learning for agent-based production scheduling. *Engineering applications of artificial intelligence* 2005, Volume 18(1), 73-82.
66. Cancino, C. A.; Merigó, J. M.; Coronado, F. C. A bibliometric analysis of leading universities in innovation research. *Journal of Innovation & Knowledge* 2017, Volume 2(3), 106-124.
67. Varin, C.; Cattelan, M.; Firth, D. Statistical modelling of citation exchange between statistics journals. *Journal of the Royal Statistical Society Series A: Statistics in Society* 2016, Volume 179(1), 1-63.
68. Moral-Muñoz, J. A., Herrera-Viedma, E., Santisteban-Espejo, A., & Cobo, M. J. (2020). Software tools for conducting bibliometric analysis in science: An up-to-date review. Profesional de la información/Information Professional, 29(1).
69. Curry, S. Let's move beyond the rhetoric: it's time to change how we judge research. *Nature* 2018, Volume 554(7690), 147-148.
70. Al-Hoorie, A.; Vitta, J.P. The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. *Language Teaching Research* 2019, Volume 23 (6), 727–744.
71. Waltman, L.; Van Eck, N.J.; Noyons, E.C.M. A Unified Approach to Mapping and Clustering of Bibliometric Networks. *J. Informetr.* 2010, Volume 4, 629–635.
72. Cheng, F.; Huang, Y.; Tanpure, B.; Sawalani, P.; Cheng, L.; Liu, C. Cost-aware job scheduling for cloud instances using deep reinforcement learning. *Cluster Computing* 2022, 1-13.
73. Sutton, R.S.; Barto, A.G. Reinforcement Learning: An Introduction. MIT, New York 2018.
74. Thaipisutikul, T.; Chen, Y.-C.; Hui, L.; Chen, S.-C.; Mongkolwat, P.; Shih, T.K. The matter of deep reinforcement learning towards practical AI applications. *Proceedings on 12th International Conference on Ubi-Media Computing* 2019, 24–29.
75. Yan, J.; Huang, Y.; Gupta, A.; Gupta, A.; Liu, C.; Li, J.; Cheng, L. Energy-aware systems for real-time job scheduling in cloud data centers: A deep reinforcement learning approach. *Computers and Electrical Engineering* 2022, 99, 107688.
76. Piller, F. T. Mass customization: reflections on the state of the concept. *International journal of flexible manufacturing systems* 2004, Volume 16, 313-334.
77. Suzić, N.; Forza, C.; Trentin, A.; Anišić, Z. Implementation guidelines for mass customization: current characteristics and suggestions for improvement. *Production Planning & Control* 2018, Volume 29(10), 856-871.
78. Steinbacher, L. M.; Ait-Alla, A.; Rippel, D.; Düe, T.; Freitag, M. Modelling framework for reinforcement learning based scheduling applications. *IFAC-PapersOnLine* 2022, Volume 55(10), 67-72.
79. Sheng, J.; Cai, S.; Cui, H.; Li, W.; Hua, Y.; Jin, B.; Wang, X. Vmagent: Scheduling simulator for reinforcement learning 2021, arXiv preprint arXiv:2112.04785.
80. Wagle, S.; Paranjape, A. A. Use of simulation-aided reinforcement learning for optimal scheduling of operations in industrial plants. In 2020 Winter Simulation Conference (WSC) 2020, 572-583, IEEE.

doi:10.20944/preprints202405.0335.v1

81. Al-Hashimi, M. A. A.; Rahiman, A. R.; Muhammed, A.; Hamid, N. A. W. Fog-cloud scheduling simulator for reinforcement learning algorithms. *International Journal of Information Technology* 2023, 1-17.