Software Impacts 22 (2024) 100720

Contents lists available at ScienceDirect

Software Impacts

journal homepage: www.journals.elsevier.com/software-impacts

Original software publication

PostgreSQL: Relational database structures application on capacitated lot-sizing for pharmaceutical tablets manufacturing processes

Michael Simonis [a,b,*], Stefan Nickel [a]

[a] Department of Economics and Management, Karlsruhe Institute of Technology, Kaiserstraße 89, Building 05.20-4A 76133 Karlsruhe, Germany
[b] Camelot Consulting Group, Theodor-Heuss-Anlage 12, 68165 Mannheim, Germany

ARTICLE INFO

Keywords:
Relational database structures
Capacitated lot-sizing
PostgreSQL
MLCLSP-L-B
Probabilistic demand
Pharmaceutical tablets manufacturing

ABSTRACT

Multi-level capacitated lot-sizing problems with linked lot sizes and backorders (MLCLSP-L-B) are used in pharmaceutical tablets manufacturing processes to right-size material production lots so that costs are kept at a minimum, production resource capacities are not exceeded, and customer demand is fulfilled. Uncertain demand behavior characterizes today's global tablets market. Pharmaceutical companies request solution approaches that solve the MLCLSP-L-B with probabilistic demand. Implementing this model in industrial applications for tablets manufacturing systems requires efficient data processing due to the amount of data and the capability to store simulated demand scenarios. This paper covers the first integration of the MLCLSP-L-B with probabilistic demand and Relational Database Structures (RDS). Modeling techniques for the RDS to process massive data are outlined. A virtual environment provides the implementation software PostgreSQL and infrastructure environment. Additionally, numerical experiments with research data are used to evaluate the agility and efficiency of the developed RDS.

Code metadata

Current code version	v1.1
Permanent link to code/repository used for this code version	https://github.com/SoftwareImpacts/SIMPAC-2024-255
Permanent link to Reproducible Capsule	
Legal Code License	GLP-3.0 license
Code versioning system used	git
Software code languages, tools, and services used	Ubuntu 22.04.1 system, PostgreSQL 14.9, Python 3.10.12
Compilation requirements, operating environments & dependencies	All necessary requirements are listed in https://gitlab.kit.edu/mn2681/mlclsplb_eerm/-/tree/software_impacts/README.md
If available Link to developer documentation/manual	https://gitlab.kit.edu/mn2681/mlclsplb_eerm/-/blob/software_impacts/Development%20Documentation.pdf
Support email for questions	michael.simonis@partner.kit.edu

1. Impact overview

Motivation in pharmaceutical tablets manufacturing

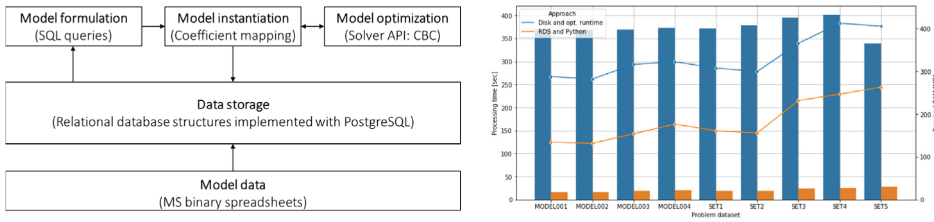
Population profiles of almost all countries are becoming older at an increasing pace. By 2030, [1] estimated that 1 in 6 people in the world will be aged 60 years or over. Thus, the increasing prevalence of chronic symptoms, investment in healthcare systems, and the incidence of novel viral diseases reshaped the global pharmaceutical drug markets. [2] observed that the pharmaceutical tablets segment has a global revenue share position of approximately 25% with a 11% compound annual growth rate from 2021 to 2028. Hence, the pharmaceutical companies request industrial applications that support decision-making for production lot-sizing to protect competitive advantages and companies' revenue targets.

Link to existing research papers

The MLCLSP-L-B is a MIP (Mixed-Integer Program) operating on massive structured production data. The model minimizes setup, inventory, and backorder costs so that demand is fulfilled in time and capac-

* Corresponding author at: Department of Economics and Management, Karlsruhe Institute of Technology, Kaiserstraße 89, Building 05.20-4A 76133 Karlsruhe, Germany.
E-mail addresses: michael.simonis@partner.kit.edu (M. Simonis), stefan.nickel@kit.edu (S. Nickel).

https://doi.org/10.1016/j.simpa.2024.100720
Received 26 October 2024; Accepted 10 November 2024



(a) Architecture design (b) RDS results

Fig. 1. Software impacts.

ities are not exceeded. This paper focuses on the data processing and utilization perspective of [3–5]. The authors presented optimization procedures for the industrial application of the MLCLSP-L-B for manufacturing pharmaceutical tablets. They highlighted that their developed optimization procedures are promising in terms of algorithmic performance. However, they provided fewer details on data architecture and processing strategies for industrial applications.

Already in the early 2000er, [6] highlighted that modern data models must ensure a certain data quality standard, efficiently process vast amounts of data and be implementable in industrial applications. Influenced by this methodology, researchers and practitioners applied RDS to run optimization models in industrial environments. [7] studied the design and use of mathematical programming (MP) based on RDS within the American steel industry. MP approaches were impacted by the description, manipulation, and display of data, so the author recommended the implementation of their optimization model via RDS. [8] extended the work of [7] by database principles for multi-period environment. The authors applied RDS to different MP approaches to scale optimization models across diverse industries.

Significance and research impact

This paper contributes to the existing literature in three aspects. First, it complements the studies of [3–5] by the first RDS application for the MLCLSP-L-B with probabilistic demand following the guidelines of [8]. Second, modeling techniques are described in the repository [9] to derive a consolidated, extended entity-relationship (EER) model processing vast amounts of structured data. Third, it grants practitioners unfettered access to problem instances and the RDS code developments in [9]. It also imparts valuable managerial insights to decision-makers regarding lot size optimization in the pharmaceutical tablet manufacturing industry.

2. Software details

[9] provides open access to the repository of the software and experimental environment, see Fig. 1(a). A *Dockerfile* sets up a Ubuntu 22.04.1 system with PostgreSQL 14.9 database and a Python 3.10.12 runtime environment. *init.sql* contains the SQL script that creates the EER model. *docker-compose.yml* lists all services that should run simultaneously. The file maps the Dockerfile on the *lot_sizing* service and the Adminer data management software to the *adminer* service. [9] provides the *README.md* and *Development Documentation.pdf* file to document how to start and use the docker environment with the datasets from *Supplementary Material A*.

3. Conclusions and future work

The execution of the docker environment and it's functional capabilities provided by [9] observe that RDS significantly reduce data processing time and RAM usage to instantiate the MLCLSP-L-B for all problem datasets. Fig. 1(b) summarizes the results from *Supplementary Material B2* for the anonymized data from [3–5]. RDS and Python (orange) required 94.36% less processing time and 45.60% less RAM than a pure Python runtime (blue) instantiation. Moreover, four different research datasets were successfully migrated into the data entry point. Data maintenance was simplified, and even the required disk space was reduced by 82.05% on average compared to the storage approach provided by the researchers (see *Supplementary Material B1*) Besides those results, faster data processing and more available RAM also positively impact the efficiency of optimization procedures of optimization model solver software (see *Supplementary Material B3*).

The primary insights of the software application are promising. Nonetheless, several open research issues remain to be audited. First, the authors focused on the RDS and the MIP instantiation. However, RDS are often loaded into advanced interactive modeling systems to build a decision support system. Optimization procedures are executed in a scalable enterprise environment, and a user interface is provided to visualize optimization results for end users. Second, the MLCLSP-L-B uses only some of the information the data entry point provides. Further MIP extensions that increase the optimization model complexity might be analyzed, and the RDS might be adapted. Such extensions integrate pharmaceutical drug shelf-life behavior, uncertain machine capacities, or random production yield into the MIP. Third, the paper focuses mainly on pharmaceutical companies. However, other processing industries implemented the MLCLSP-L-B with probabilistic demand in industrial applications (paper industry, chemistry, or food industry). The studies can be applied to these industries with their data characteristics, and industry-specific benchmarks might be created.

CRediT authorship contribution statement

Michael Simonis: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Stefan Nickel:** Conceptualization, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.simpa.2024.100720.

References

[1] World Health Organization, Ageing and Health Report, Global Health Report, World Health Organization, Avenue Appia, Geneva 27, Switzerland, 2022, Accessed: 2023-06-01, Published: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.

[2] Grand View Research, Pharmaceutical Manufacturing Market Size, Share & Growth Analysis Report, Market Analsis Report GVR-4-68039-014-2, Grand View Research Inc., 201 Spear Street 1100, San Francisco, United States, 2021, Published: https://www.grandviewresearch.com/industry-analysis/pharmaceutical-manufacturing-market. (Accessed: 8 August 2023).
[3] M. Simonis, S. Nickel, Generalized data model for real-world capacitated lot-sizing problems with linked lot sizes and backorders, Data Brief 49 (2023) 109440.
[4] M. Simonis, S. Nickel, A simulation–optimization approach for a cyclic production scheme in a tablets packaging process, Comput. Ind. Eng. 181 (2023) 109304.
[5] M. Simonis, S. Nickel, Integrated shelf-life rules for multi-level pharmaceutical tablets manufacturing processes, Int. J. Prod. Res. (2024) 1–21.
[6] D.L. Moody, Metrics for evaluating the quality of entity relationship models, in: International Conference on Conceptual Modeling, Springer, 1998 pp. 211–225.
[7] R. Fourer, Database structures for mathematical programming models, Decis. Support Syst. 20 (4) (1997) 317–344.
[8] G. Dutta, R. Fourer, Database structure for a class of multi-period mathematical programming models, Decis. Support Syst. 45 (4) (2008) 870–883.
[9] M. Simonis, Postgresql: Relational database structures application on capacitated lot-sizing for pharmaceutical tablets manufacturing processess, 2024, https://gitlab.kit.edu/mn2681/mlclsplb_eerm/-/tree/software_impacts v1.1.