Plan Overview

A Data Management Plan created using DMPonline

Title: Models and algorithms for rolling stock (re)scheduling

Creator: Danny Zhu

Principal Investigator: Jia Hui Zhu

Data Manager: Jia Hui Zhu

Project Administrator: Dennis Huisman, Twan Dollevoet

Affiliation: Erasmus University Rotterdam

Template: Data Management Plan v4.6

ORCID iD: 0000-0003-1050-1928

Project abstract:

Rolling stock is the most expensive resource of a railway company. Therefore, it is important to schedule rolling stock in an efficient way. In addition, the availability of a seat is important for the passengers. Minimum service levels on seat availability are also set by the government in the concession of the railway company. Also, when there are disruptions, seat availability requirements have to be satisfied. As such this has to be taken into account in rolling stock rescheduling during the operations. In a previous research project (Hoogervorst), new solution methods for rolling stock scheduling and rescheduling have been developed. However, these methods do not take all operational details into account. When adding all these operational details, the mathematical models get extremely challenging to solve. In this project, we want to develop new algorithms to solve these challenging models.

ID: 180541

Start date: 01-10-2022

End date: 01-10-2026

Last modified: 17-07-2025

Models and algorithms for rolling stock (re)scheduling

General

Please tick the following boxes if you agree to act according to the following terms:

- I will check and, if necessary, update my data management plan a minimum of once a year
- I will discuss the data management plan with my research team
- I will answer all questions truthfully and to the best of my knowledge

Support in writing a data management plan is available through the faculty Data Stewards. Which research support professional is available for you?

- Data Steward of my own faculty - ESE

Scientific research must be conducted in line with existing guidelines on good research practices and integrity. Please tick the boxes if you have read and understand these guidelines and will act accordingly.

- The Netherlands Code of Conduct for Research Integrity (VSNU, 2018)
- The European Code of Conduct for Research Integrity (ALLEA, 2023)

Administration and Project Description

1. Provide the details of your project

Project title

Models and algorithms for rolling stock (re)scheduling

Project start date as intended

2022-10-01

Project duration in months as intended

48

Funding body (if applicable)

-

Grant number (if applicable)

-

Date of DMP Version 1

0020-08-07

Current DMP - Version [if other than version 1]

2

Current DMP - Date [if other than version 1]

2025-06-19

2. List the name and affiliation of all members of the research team.
List the researcher responsible for research data management first.
For PhD projects, please indicate the Promotor(s) and/or Daily Supervisor(s) with a (!)

	Name	Email	ORCID	Research Institution
1	Jia Hui Zhu	j.h.zhu@ese.eur.nl	0000-0003-1050-1928	ESE
2	Twan Dollevoet (!)	dollevoet@ese.eur.nl	0000-0003-3193-823X	ESE
3	Dennis Huisman (!)	huisman@ese.eur.nl	0000-0001-9114-658X	ESE
4				
5				
6				
7				
8				
9				
10				

3. Briefly summarize the project background and research question(s) to help others understand the purpose for which the data are being collected or created

Rolling stock is the most expensive resource of a railway company. Therefore, it is important to schedule rolling stock in an efficient way. In addition, the availability of a seat is important for the passengers. Minimum service levels on seat availability are also set by the government in the concession of the railway company. Also, when there are disruptions, seat availability requirements have to be satisfied. As such this has to be taken into account in rolling stock rescheduling during the

operations. In a previous research project (Hoogervorst), new solution methods for rolling stock scheduling and rescheduling have been developed. However, these methods do not take all operational details into account. When adding all these operational details, the mathematical models get extremely challenging to solve. In this project, we want to develop new algorithms to solve these challenging models.

4. Specify the research type and briefly describe the methodology, how the data will be collected, and the tools used for data collection, processing and analysis:

My research falls within the field of operations research, which focuses on designing mathematical models and algorithms to solve logistical problems. My research relates to trains and railways and is therefore in collaboration with Netherlands Railways, NS, who provide me with data sets that are based on real-life data. This data is confidential within NS. To process the data, programming code is written in Java, using the IDE Eclipse and IntelliJ. No data analysis is conducted.

5. Are additional (financial or time) resources required for data management in this project?

- No, I will use the services and resources provided by the EUR

Preparation: Legal Arrangements and Policy

6. With whom will you need to make legal arrangements?

- With third parties

7. List the agreements that you will initiate and with whom will you make them.

Who	Type of agreement
NS	Non-disclosure agreement

8. List the agreements or other data management policies that you need to uphold but did not initiate. If you are reusing existing data, list the terms of use under which you may re-use them.

Who	Type	Version and Date
EUR	RDM policy of Erasmus University Rotterdam	Version 1.0 [August 14th, 2020]
EUR	Internet and ICT facilities policy	July 13th, 2021
Eclipse	Terms and conditions	01-10-2022
IntelliJ	Terms and conditions	01-05-2025

9. Do you need to obtain ethical approval for your research project?

- No, my project does not require ethical approval

10. If you have obtained ethical approval, list the reference number

Question not answered.

During research: Collecting and analyzing

11. Specify what data you will be collecting and indicate format, estimated size, and whether this is data that you will be generating or existing data that you will be re-using.

Type	Data Classification	Format	Estimated size	Generate or Re-use
NS train timetable	confidential	.txt	<1 GB	Re-use
Data on rolling stock	confidential	.txt	<1 GB	Re-use
Initial rolling stock schedule	confidential	.txt	<1GB	Generate
Programming code	confidential	.txt	1-5GB	Generate

12. Will you be collecting or re-using (sensitive) personal data?

- No – My research does not include human participants

13. If you collect or re-use (sensitive) personal data, how will you protect the privacy of participants?

- Not applicable - I do not collect or re-use personal data

14. Please elaborate on your anonymization/ pseudonymization plans. If you are working with multiple datasets, please specify which datasets will be anonymized and which will be pseudonymized.

None

15. Will you be collecting or re-using non-personal sensitive data?

- Yes

16. Where will you store your data during the project? You can select multiple options.

- EUR OneDrive
- Other (please specify where the data will be stored in additional information box)

NS github

18. What hardware and software do you use? Select all applicable options.

- Private software or freeware [e.g. private DropBox]
- EUR supported software as found in the software catalog
- Private hardware [e.g. personal laptop, private external hard-drive]
- EUR supported hardware [e.g. @wEURk laptop, @wEURk workstation]

19. If you use private hardware, software, or freeware, please specify what and for what reason:

Personal PC for running programming code, Eclipse IDE for writing the programming code

20. Are regular backups made of your data?

- Yes, manually (please specify WHO makes the backups and HOW OFTEN backups are made in the additional information box).

I create a back up of my programming code each time I make changes, by uploading it to the NS github.

21. Who manages access to the data?

- Researcher responsible for research data management

22. Who will have access to the data (during the project)?

- Only researchers as indicated under ‘Administration & Project description’

23. How are you going to make sure your data will be accessible in case of staff changes, illness, etc?

- There is a clear procedure in place in my research team, department, or faculty

24. Have you and your research team agreed on a way to name and order project folders and files?

- Yes - I am working on the documentation

25. Have you and your research team agreed on how to handle versioning of files?

- Yes - I am working on the documentation

Research Publication: Data sharing and re-use

26. What data (and code) will be shared in a research data repository?

- A selection of the data (and code)

27. Please specify why you are unable to share (all) data (and code)

I signed an NDA with NS, so the confidential datasets from NS cannot be shared.

28. List the data (and code) that you plan to share in a research data repository. Also list the information / documentation / metadata that you will include to make the data package self-explanatory and re-usable in the future (for other researchers and yourself)

Data	Format	Size
Programming code	.java	<1 GB
Readme text file (code documentation)	.txt	<1 GB

29. In which repository will you place the metadata, data, and/or code that are associated with your paper?

- Zenodo
- EUR Data Repository (EDR)

30. What metadata standard will you use to document your research?

- DCMI [Dublin Core Metadata Initiative] (Note: Default within the EUR Data Repository)

31. Will you place any restrictions on re-using the data you plan to share?

- No

34. Under what license will you make your data available for re-use?

- License for specific types of data (e.g. MIT software license, please specify in Q.35)

35. Please specify which license

MIT license

After research: Archiving

36. You may be obliged to destroy some data before archiving. Do any of such obligations apply to you?

- No

37. List the data and all documentation you will be archiving. These data constitute your archival package.

Data	Format	Size		
NDA	.pdf	<1 GB		
Programming code	.java	<1 GB		
Readme text file (code documentation)	.txt	<1 GB		
NS train timetable	.txt	<1 GB		
Data on rolling stock	.txt	<1 GB		
Initial rolling stock schedule	.txt	<1 GB		
RDM policy of Erasmus University Rotterdam	.pdf	<1 GB		
Internet and ICT facilities policy	.pdf	<1 GB		
Eclipse Terms and conditions	.pdf	<1 GB		
IntelliJ Terms and conditions	.pdf	<1 GB		

38. Where will you be archiving your data?

- EUR Yoda Vault (EUR Archive) [retention period min. 10 years] --> You have reached the end of the DMP

The train data that I use is stored in NS' company repository, to which my supervisor and my other colleagues at NS have access. In case it is needed for academic verification, it can be retrieved easily by my supervisor.