Simulation-assisted machine learning for yard departure prediction: a benchmarking study between European and North American contexts

- Niloofar Minbashi
- PhD Candidate
- KTH Royal Institute of Technology

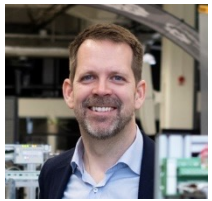
Minbashi, N., Zhao, J., Dick, C. T., Bohlin, M. (2023, April 25-28). Application of simulation-assisted machine learning for yard departure prediction. Presented at the 10th International Conference on Railway Operations Modelling and Analysis (ICROMA), RailBelgrade 2023, Belgrade, Serbia.

Agenda

- Motivation
- Overview
- Research Questions
- Method
- Results
- Conclusion

- PhD thesis: **Application of Predictive Analytics for Shunting Yard Delays**
- October 2018 – June 2023
- Licentiate defense November 2020

- Supervisors:

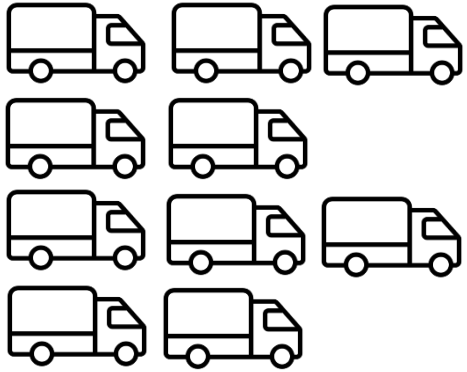


Markus Bohlin
Visiting Professor, KTH

Behzad Kordnejad
Researcher, KTH

Carl-William Palmqvist
Researcher, KTH

Motivation



$\approx 10\times10^{9}\ tonne-kilometers$

$79\times10^{6}\ tonnes\ \text{of}\ CO_2$

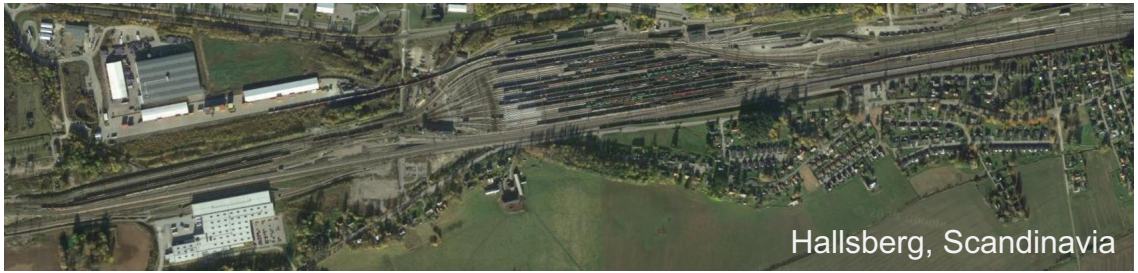
$\approx 20\times10^{9}\ tonne-kilometers$

$1800\times10^{6}\ tonnes\ \text{of}\ CO_2$

Source: The International Transport Forum's ITF Transport Outlook 2019

50 by 2050

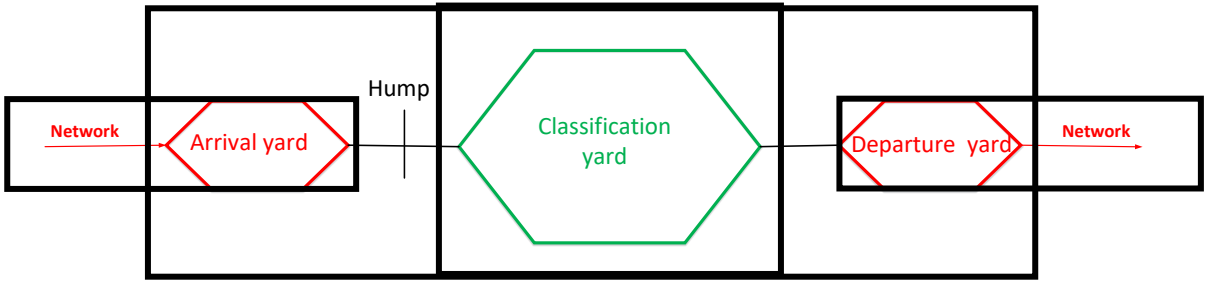
Yards


Hallsberg, Scandinavia


Maschen, Europe


Bailey, North America

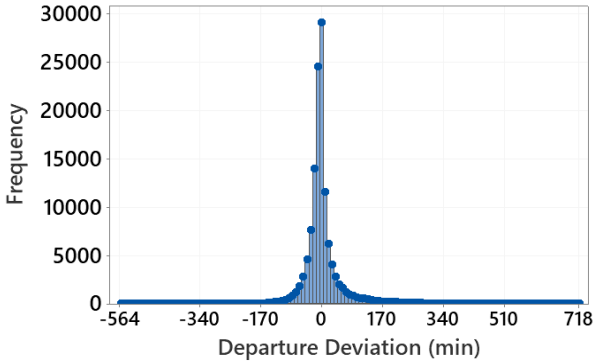
Research Problem

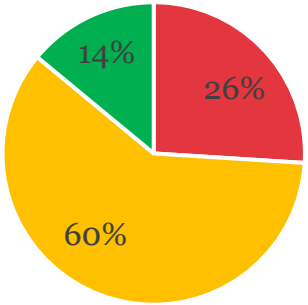


Research Gaps

1. Methods: Data-driven
    1. Descriptive Analytics
    2. Predictive Analytics
    3. Hybrid Modeling
2. Yard performance:
    1. Analyzing yard departure deviations
    2. Analyzing the network impact on departure delays
3. Yard models:
    1. Yard departure prediction models
    2. Hybrid yard and network models
    3. Hybrid yard models: event-driven and data-driven models

Research Problem (Swedish Context): Yard Departure Prediction





Infrastructure Manager
Yard Operator

Descriptive Analytics:
Probability Distribution

Contents lists available at ScienceDirect

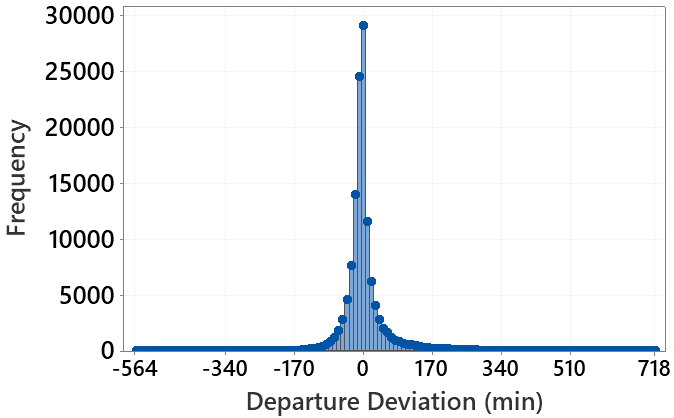
Journal of Rail Transport Planning & Management

journal homepage: http://www.elsevier.com/locate/jrtpm

Statistical analysis of departure deviations from shunting yards: Case study from Swedish railways

Niloofar Minbashi [a,*], Carl-William Palmqvist [b,c], Markus Bohlin [a], Behzad Kordnejad [a]

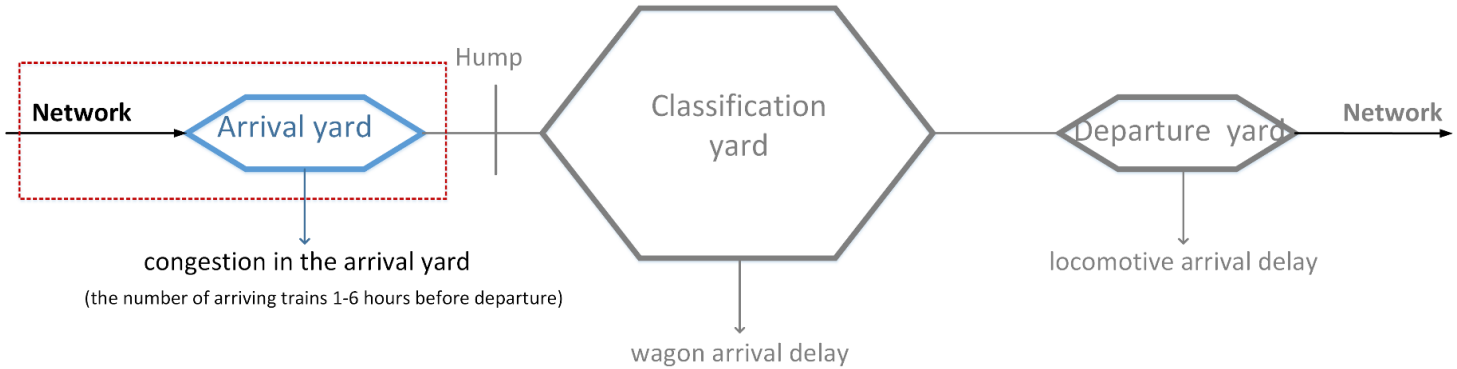
[a] Division of Transport Planning, School of Architecture and Built Environment, KTH, SE-100 44, Stockholm, Sweden
[b] Division of Transport and Roads, Department of Technology and Society, Lund University, P.O. Box 118, 221 00, Lund, Sweden
[c] K2 Swedish Knowledge Centre for Public Transport, Bruksgatan 8, 222 36, Lund, Sweden



Descriptive Analytics:
Network Impact

Minbashi, N., Bohlin, M., Kordnejad, B. (2020). A departure delay estimation model for freight trains. TRA 2020, the 8th Transport Research Arena: Rethinking transport – towards clean and inclusive mobility, Helsinki, Finland.

Minbashi, N., Bohlin, M., Kordnejad, B. (2021, February 3-4). Analysis of railyard congestion and departure delays relationship: a case study from Swedish railways. hEART 2020: 9th Symposium of the European Association for Research in Transportation, Lyon, France.



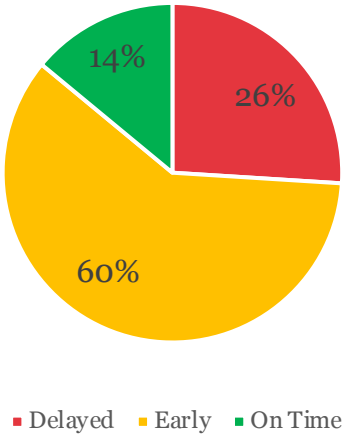
Predictive Analytics:
Yard Departure Prediction Model

Research Article

The Application of Tree-Based Algorithms on Classifying Shunting Yard Departure Status

Niloofar Minbashi[1], **Markus Bohlin**[1], **Carl-William Palmqvist**[2], **and Behzad Kordnejad**[1]

[1]*Division of Transport Planning, KTH Royal Institute of Techonology, 100 44 Stockholm, Sweden*
[2]*Division of Transport and Roads, Lund University, P.O. Box 118, 221 00 Lund, Sweden*



Hybrid Modeling: Integrated Yard-Network

Contents lists available at ScienceDirect

Journal of Rail Transport Planning & Management

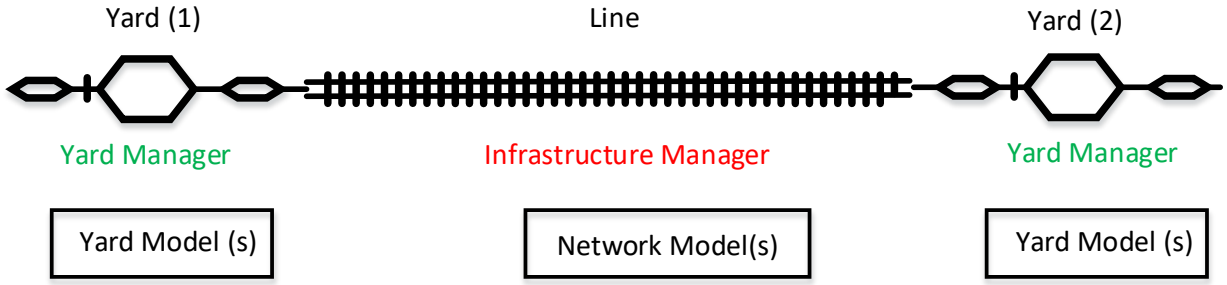
journal homepage: www.elsevier.com/locate/jrtpm

Check for updates

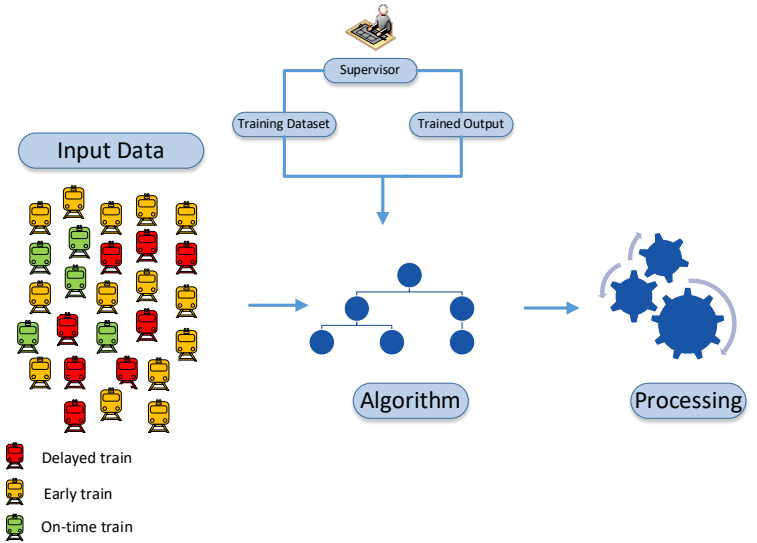
Machine learning-assisted macro simulation for yard arrival prediction

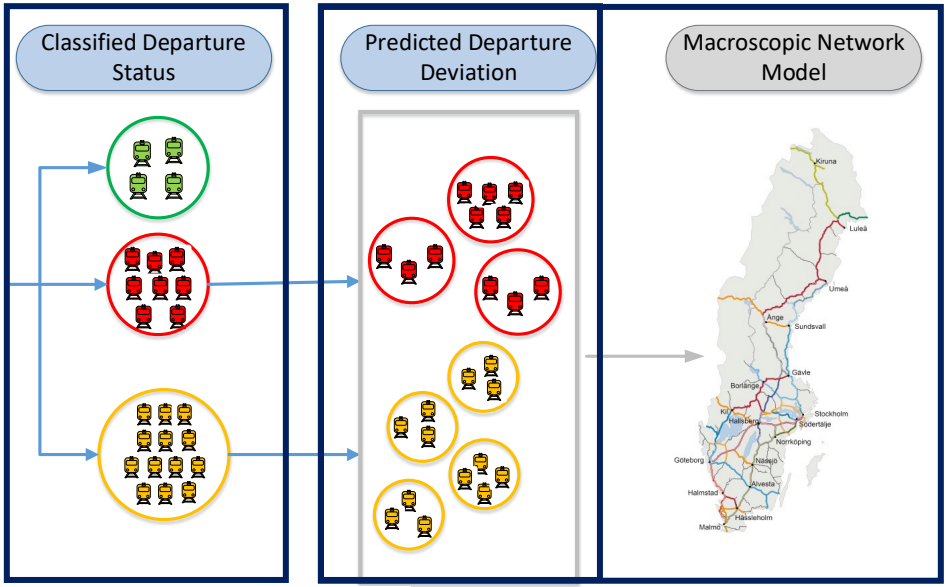
Niloofar Minbashi *, Hans Sipilä, Carl-William Palmqvist, Markus Bohlin, Behzad Kordnejad

Division of Transport Planning, KTH Royal Institute of Technology, Brinellvägen 23, SE-100 44 Stockholm, Sweden



Model Framework





Application of Simulation-assisted Machine Learning for Yard Departure Prediction

Niloofar Minbashi [a,1], Jiaxi Zhao [b], C. Tyler Dick [b], Markus Bohlin [a,c]
[a] Division of Transport Planning, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden
[b] Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX, USA
[c] School of Innovation, Design and Technology, Mälardalen University, Eskilstuna, Sweden
[1] E-mail: minbashi@kth.se, Phone: +46 (8) 790 70 54

Research Article

The Application of Tree-Based Algorithms on Classifying Shunting Yard Departure Status

Niloofar Minbashi [1], Markus Bohlin [1], Carl-William Palmqvist [2], and Behzad Kordnejad [1]
[1]Division of Transport Planning, KTH Royal Institute of Techonology, 100 44 Stockholm, Sweden
[2]Division of Transport and Roads, Lund University, P.O. Box 118, 221 00 Lund, Sweden

Contents lists available at ScienceDirect
Journal of Rail Transport Planning & Management
journal homepage: www.elsevier.com/locate/jrtpm

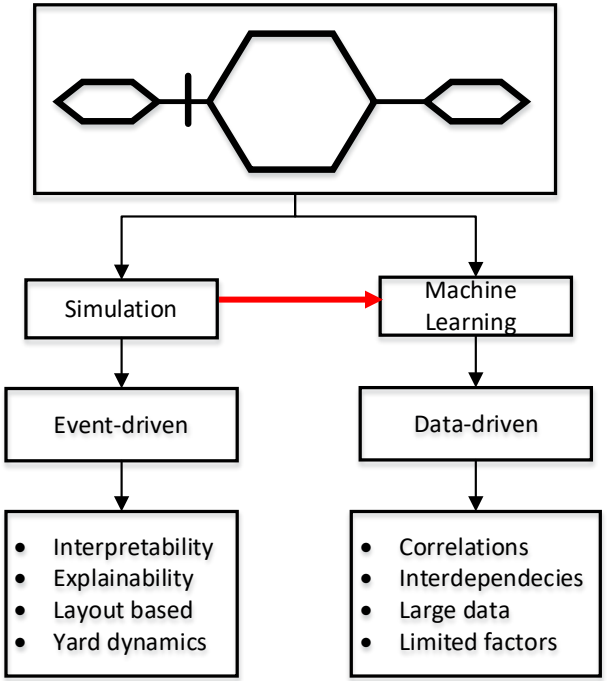
Machine learning-assisted macro simulation for yard arrival prediction

Niloofar Minbashi *, Hans Sipilä, Carl-William Palmqvist, Markus Bohlin, Behzad Kordnejad
Division of Transport Planning, KTH Royal Institute of Technology, Brinellvägen 23, SE-100 44 Stockholm, Sweden

Hybrid Modeling:

Application of Simulation-assisted Machine Learning for Yard Departure Prediction

Niloofar Minbashi [a,1], Jiaxi Zhao [b], C. Tyler Dick [b], Markus Bohlin [a,c]
[a] Division of Transport Planning, KTH Royal Institute of Technology, SE-100 44, Stockholm, Sweden
[b] Department of Civil, Architectural and Environmental Engineering, University of Texas at Austin, Austin, TX, USA
[c] School of Innovation, Design and Technology, Mälardalen University, Eskilstuna, Sweden
[1] E-mail: minbashi@kth.se, Phone: +46 (8) 790 70 54



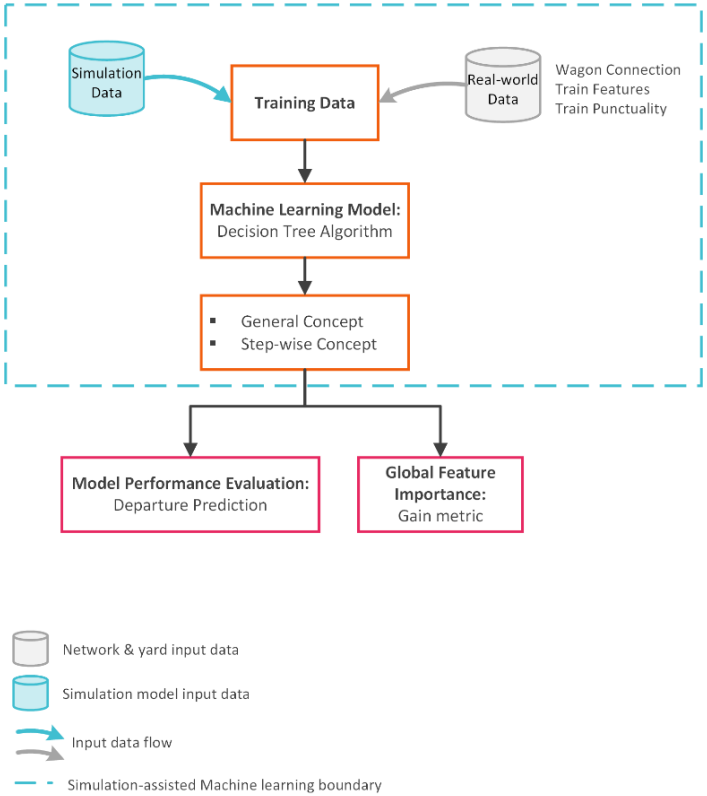
Research Questions

1) Can synthetic data from simulation in a North American context validate a machine learning model developed on real-world yard data from the European context?

2) Can simulation data improve the machine learning by providing access to more operational factors?

3) Can a model developed from simulation data inform the collection of new performance metrics to improve a model developed from the real-world data?

Method:
Simulation-assisted Machine Learning



Method:
Machine Learning

- Original model extended from (Minbashi et al., 2021 & 2023)
- Decision Tree
- 10-fold cross-validation

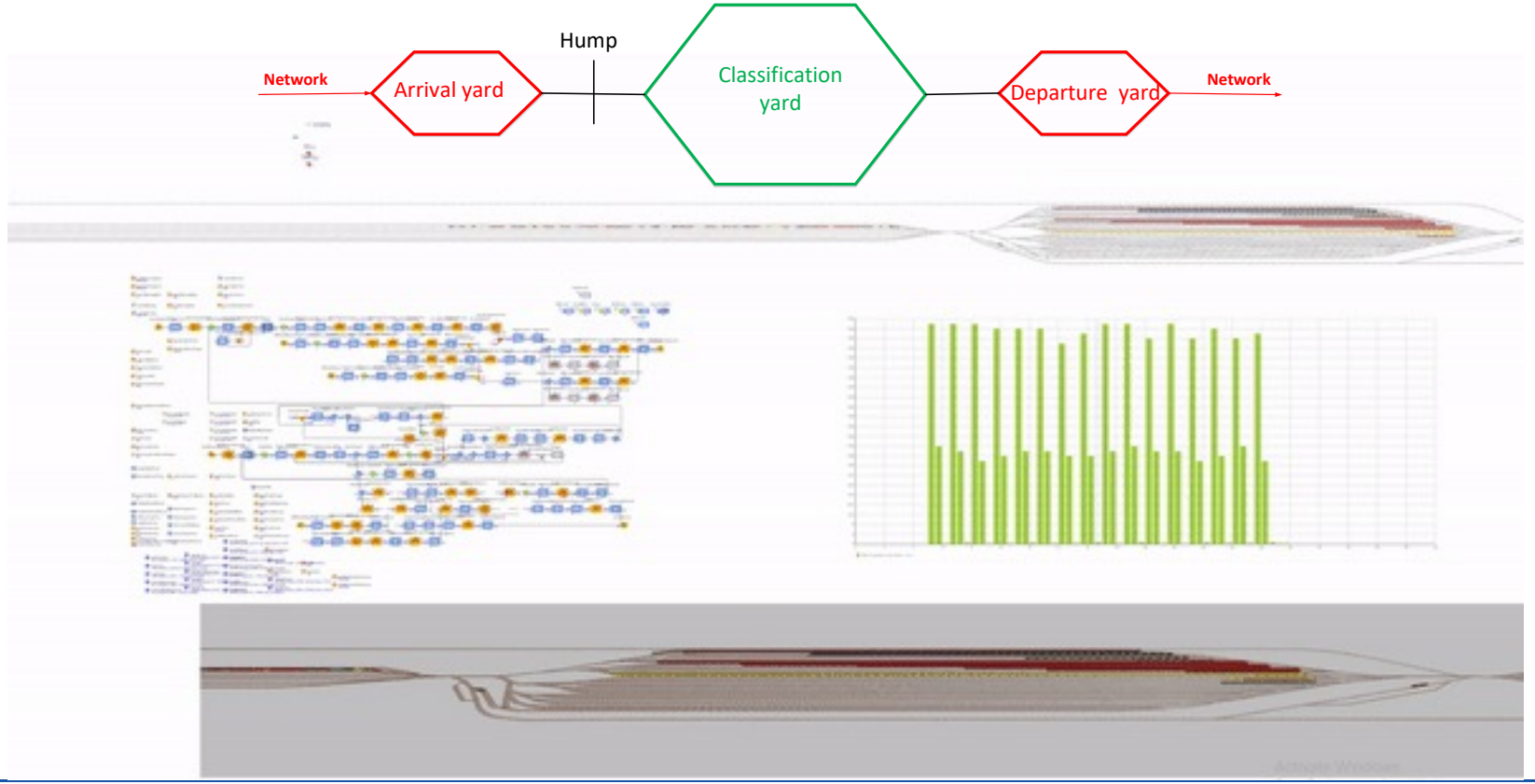
- Evaluation Criteria:
  - R-Squared
  - Mean absolute error
  - Root mean squared error
  - Mean signed difference

Method:
ML - Predictors

Schedule Data
Departing Train Number
Departure Hour
Departure Week-Day
Departure Month
Min Wagon Dwell Time
Max Wagon Dwell Time
Number of Wagons
Number of Arriving Trains
Maximum Planned Weight
Maximum Planned Length

Operational Data
Maximum Operated Length
Maximum Operated Weight
Min Arriving Deviation
Mean Arriving Deviation

Method:
Discrete Event Simulation



Method:
Simulation - Predictors

Data Type	Predictors
Schedule	Scheduled departure hour
	Train length
	Number of arriving trains
	Number of wagons
Arrival yard	Minimum hump duration
	Maximum hump duration
	Receiving utilization
	Hump utilization
	Idle time waiting for hump
Classification yard	Minimum bowl idle
	Maximum bowl idle
	Pulldown utilization
Departure yard	Departure utilization
	Departure inspection time
	Idle time waiting for departure

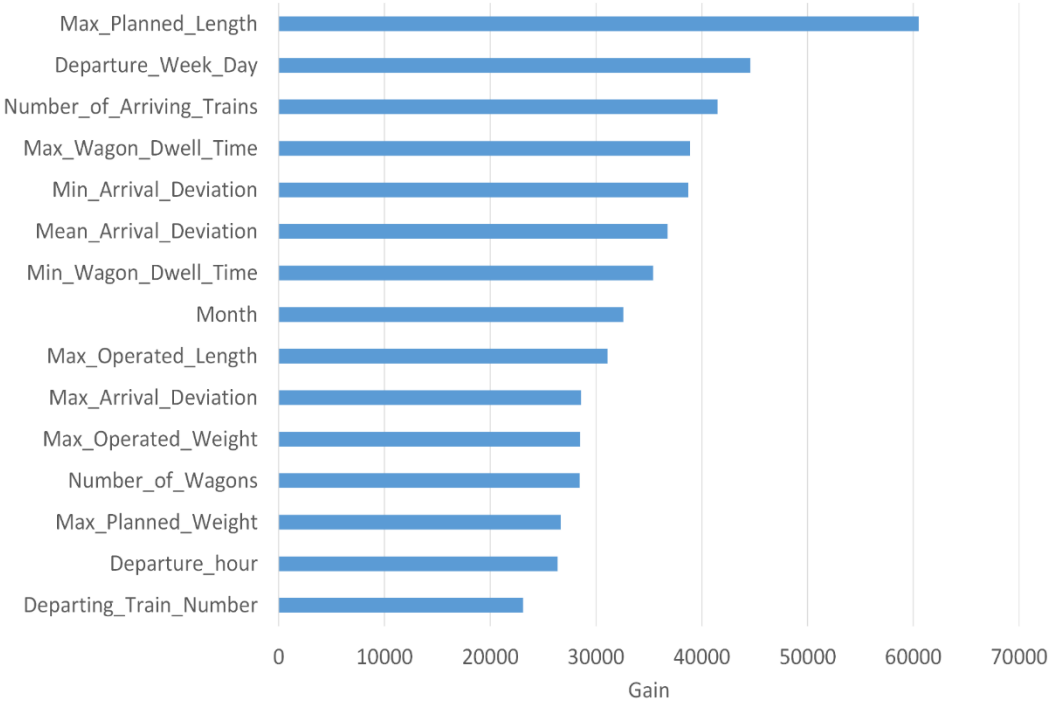
Results:
General Concept

Model (General Concept)	R-Squared	Mean Absolute Error (min)	Root Mean Squared Error (min)	Mean Signed Difference
Real-world	0.90	3.4	14.6	-0.004
Baseline simulation	0.87	3.0	6.6	0.002
Ultimate Randomness simulation	0.70	6.6	12.6	0..002

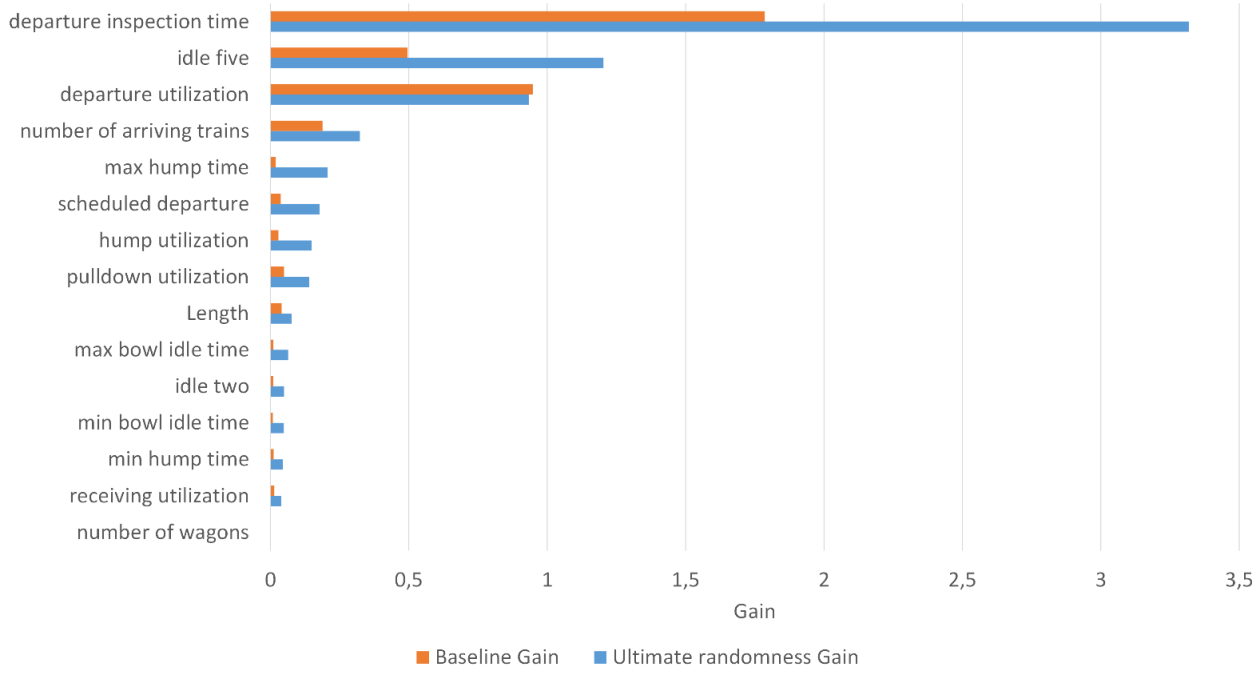
Results:
Step-wise Concept

Model	Step-wise Concept	R^2	MAE	RMSE	MSD
Real-world	Step 0 Schedule	0.90	3.5	15.18	0.05
	Step 1 Arrival	0.88	3.6	16.4	0.11
	Step 2 Classification	0.90	3.5	15.31	0.06
Baseline simulation	Step 0 Schedule	0.80	4.2	7.80	0.00
	Step 1 Arrival	0.79	4.2	8.40	0.00
	Step 2 Classification	0.79	4.2	8.40	0.002
	Step 3 Departure	0.87	3.0	6.00	0.001
Ultimate Randomness simulation	Step 0 Schedule	0.57	9.0	16.00	0.003
	Step 1 Arrival	0.60	8.0	15.00	0.003
	Step 2 Classification	0.62	9.0	15.00	0.001
	Step 3 Departure	0.70	7.0	13.00	0.003

Results:
Global Feature Importance



Results:
Global Feature Importance



Conclusion

- Machine learning performs well for both contexts.
- Simulation data can help in validating model performance.
- Simulation data can be beneficial where data is sensitive.
- Schedule predictors have enough information for accurate departure predictions.
- Departure yard predictors add higher information for ýard departure prediction.
- Adding randomness to the simulation should be explored in explainability terms.

Acknowledgement

Questions?

Niloofar Minbashi

minbashi@kth.se

Jiaxi Zhao

jiaxi.zr@utexas.edu

Stockholm, 2022-06-13

List of papers:

I. Minbashi, N., Palmqvist, C. W., Bohlin, M., Kordnejad, B. (2021). Statistical analysis of departure deviations from shunting yards: case study from Swedish railways. Journal of Rail Transport Planning & Management, 18, 100248.

II. Minbashi, N., Bohlin, M., Kordnejad, B. (2020). A departure delay estimation model for freight trains. Conference Proceedings of 8th Transport Research Arena TRA, Helsinki, Finland.

III. Minbashi, N., Bohlin, M., Palmqvist, C. W., Kordnejad, B. (2021). The application of tree-based algorithms on classifying shunting yard departure status. Journal of Advanced Transportation, 2021, 1-10.

IV. Minbashi, N., Sipilä, H., Palmqvist, C. W., Bohlin, M., Kordnejad, B. (2023). Machine learning-assisted macro simulation for yard arrival prediction. Journal of Rail Transport Planning & Management, 25, 100368.

V. Minbashi, N., Zhao, J., Dick, C. T., Bohlin, M. (2023). Application of Simulation-assisted machine learning for yard departure prediction. Accepted at the 10th International Conference on Railway Operations Modelling and Analysis (ICROMA). RailBelgrade 2023.

VI. Minbashi, N., Bohlin, M., Kordnejad, B. (2021, February 3-4). Analysis of railyard congestion and departure delays relationship: a case study from Swedish railways. hEART 2020: 9th Symposium of the European Association for Research in Transportation, Lyon, France.