A Data-driven Approach to Predict Railway In-train Forces

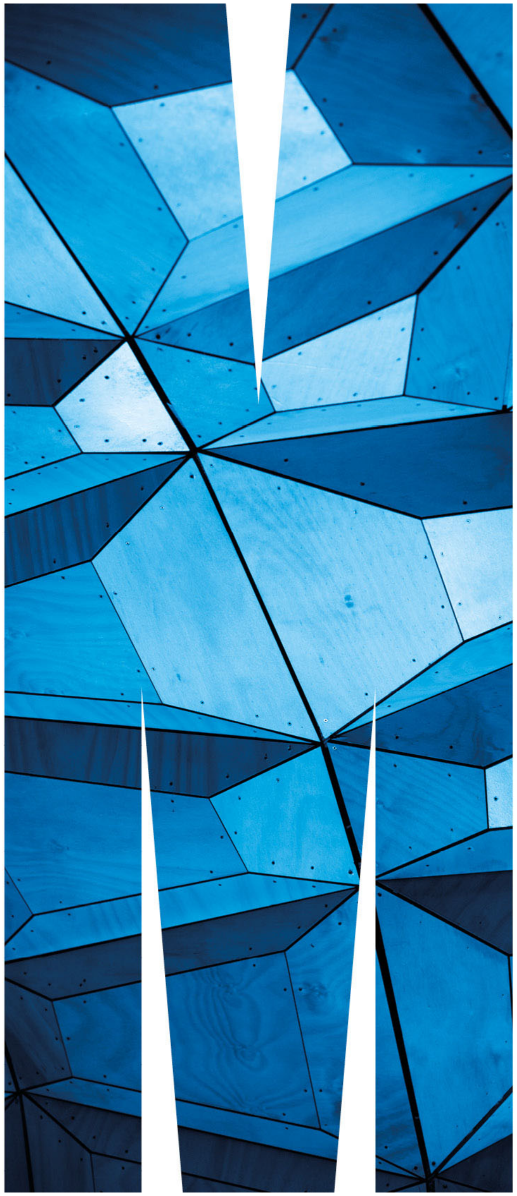
Presenter: Mr. Sheng Zhang (PhD student)

Main supervisor: Prof. Wenyi Yan

Department of Mechanical and Aerospace Engineering, Monash University

Associate supervisors: Dr. Ryan Huang, Mr. Tim Constable

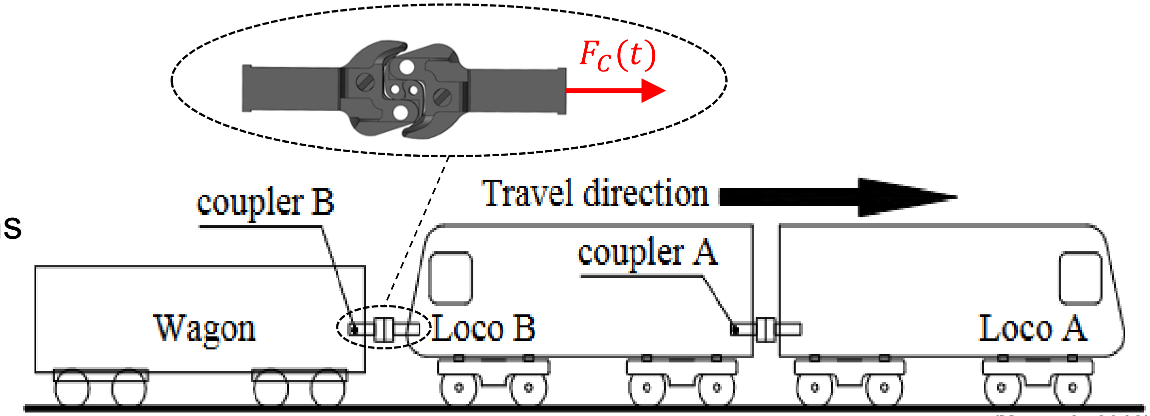
Institute of Railway Technology, Monash University



Introduction

Railway in-train forces

- Physical draft and buff forces
- Train operations and topographical conditions
- Getting larger and more complicated



(Xu et al., 2012)

In-train force related research

- Component integrity evaluation
- Wagon stability assessment
- Control strategy design
- Service quality improvement
- Train energy management

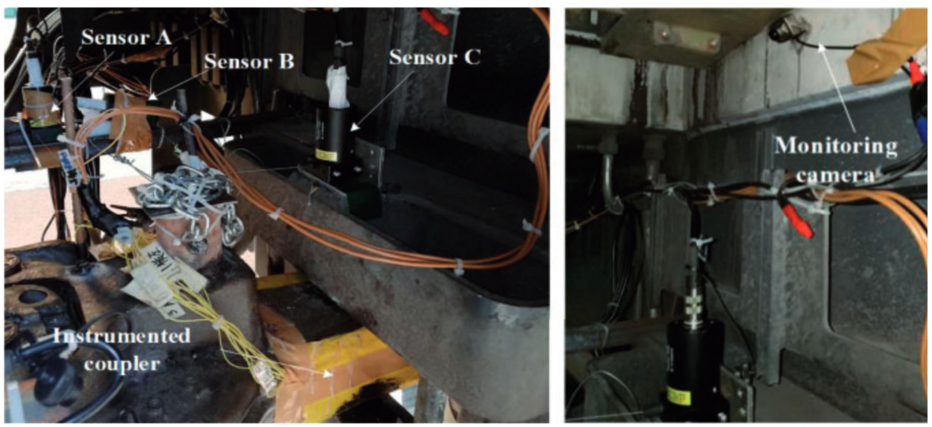

Failure of a railway coupler (Cookson & Mutton, 2014)

Introduction

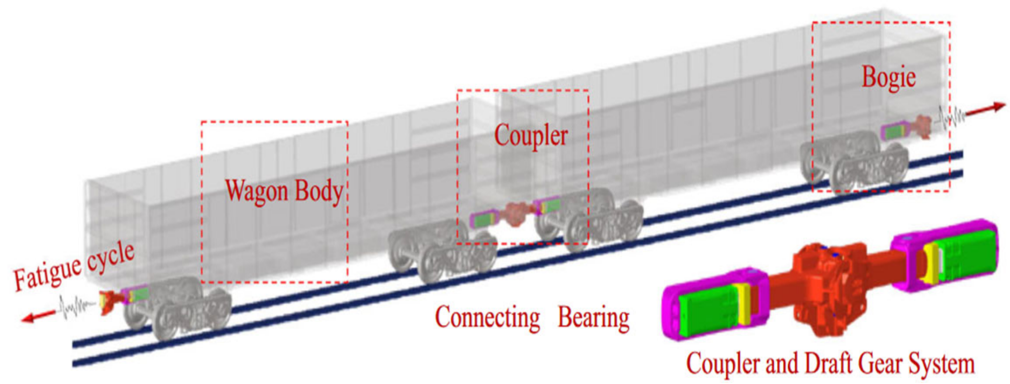
Current measurement/prediction methods

- Field measurement
  - ✓ Most reliable
  - × Enormous time and manpower
  - × Risk of sensing device damage

- Multibody dynamics (MBD) simulation
  - ✓ Relatively cost-effective
  - × High level of domain knowledge required
  - × Complex numerical model
  - × Large computational and storage spaces



Field measurement (Ge et al., 2021)



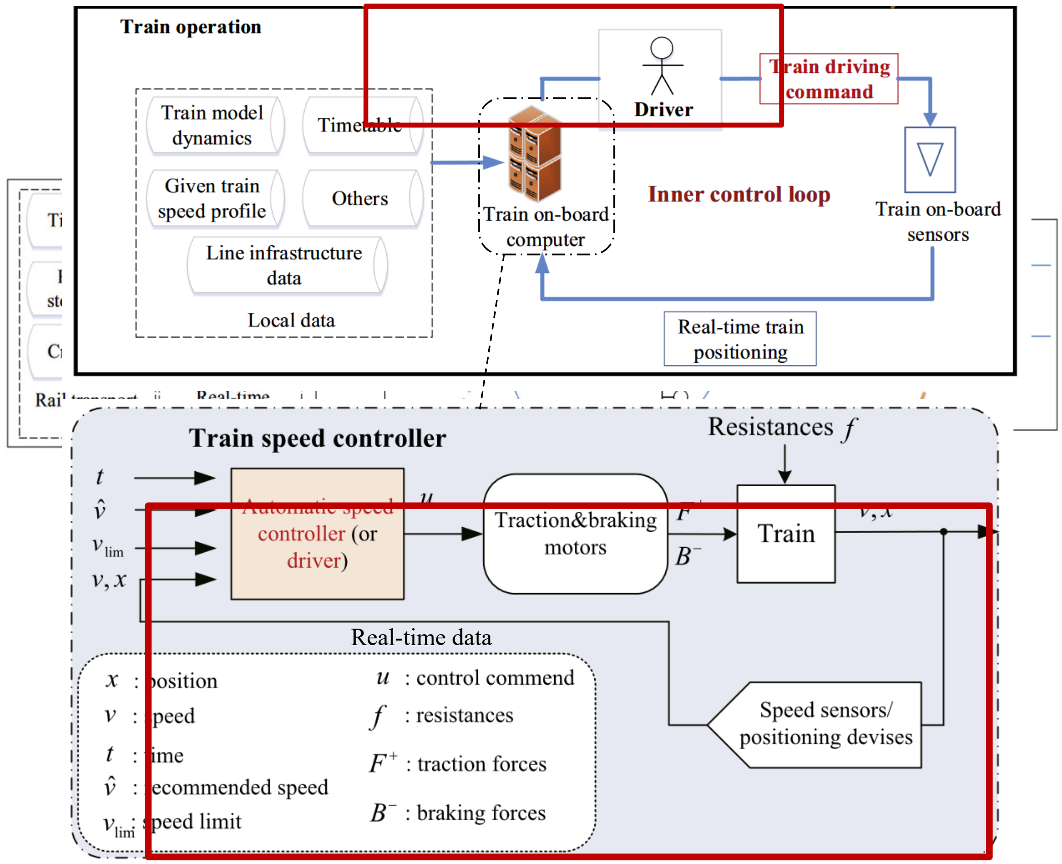
Multibody dynamics simulation (Ren et al., 2022)

Both methods can only be conducted for a specific service condition one at a time

Introduction

Automatic train operation (ATO) system

- Improve train's operations based on real-time measured data
- Structure:
  - Railway traffic management module
  - Train operation control module
- Onboard real-time measurements:
  - Train dynamic responses
  - Driving behaviours
  - Topographical information
- Do not integrate the function of real-time in-train force measurement/prediction

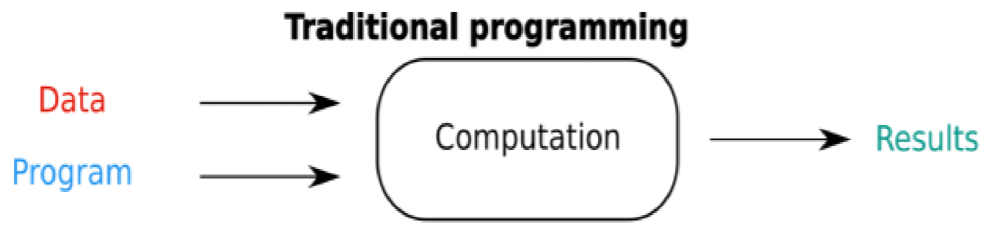


ATO train operation control module (Yin et al., 2017)

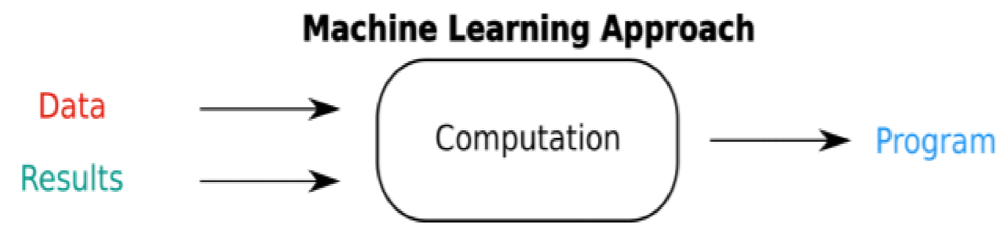
Introduction

Machine learning in engineering

- ▪ Teaching computers to learn from data and make decisions
- ✓ Solves complex problems that traditional programming cannot
- × Requires a large amount of high-quality data, lacks interpretability



Data – results relationship (program) sometimes hard to be defined



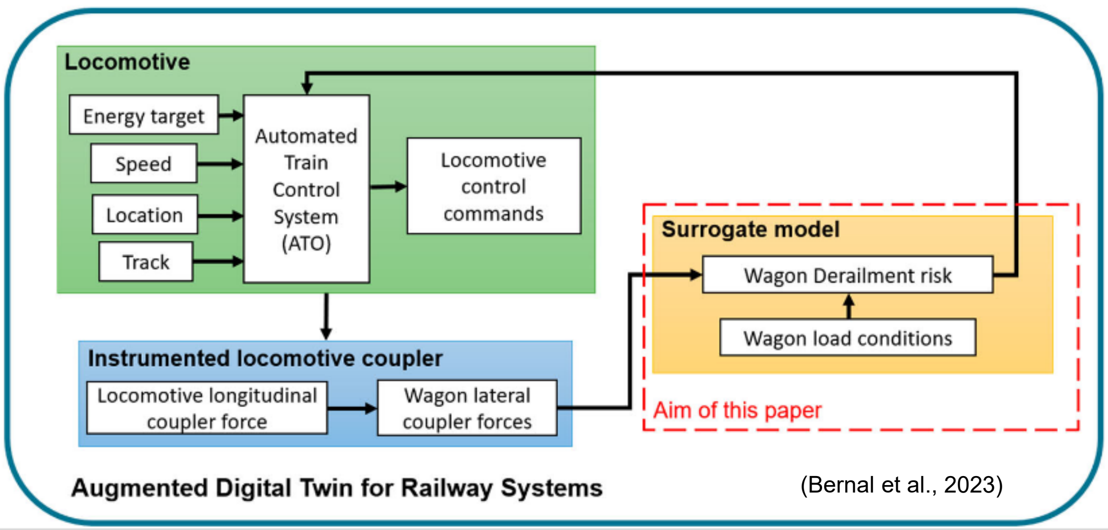
Literature Review -- Machine learning

Title	Author, Year	Research	Method	Key Finding
A data-driven dynamics simulation framework for railway vehicles	Nie et al., 2018	Train crashworthiness	Multibody dynamics; Finite element method; Decision tree model	The time spent in co-simulation is less than MBD simulations with higher accuracy.
Artificial neural networks applied to the measurement of lateral wheel-rail contact force: A comparison with a harmonic cancellation method	Urda et al., 2020	Wheel lateral force prediction	Harmonic cancellation method Multibody dynamics; Neural networks	ANNs are reliable alternatives for both the harmonic cancellation method and MBD, but ANNs have a shorter predicting time.
MBSNet: A deep learning model for multibody dynamics simulation and its application to a vehicle-track system	Ye et al., 2021	System dynamic responses prediction	Vehicle-track system dynamics; CNN-LSTM neural network	The deep learning model has high robustness in different inputs and can quickly achieve long-term predicted dynamic responses.

Machine learning based methods have a high accuracy with shorter prediction time

Literature Review -- Digital twin

Title	Author, Year	Research	Method	Key Finding
Application of machine learning techniques to build digital twins for long train dynamics simulations	Bosso et al., 2023	Safety index and wheel-rail forces prediction	LTD simulation Machine learning	Surrogate models accurately predict safety indexes with low calculation time.
Vehicle system dynamics in digital twin studies in rail and road domains	Bernal et al., 2023	Real-time derailment risk prediction	Field measurement Machine learning	Effective prediction of derailment risk, improving railway operations.



ML model is used to replace lateral/vertical vehicle dynamics simulation

Literature Review -- International Benchmark

Longitudinal train dynamics (LTD) simulation

- **International benchmarking of longitudinal train dynamics simulators: Benchmarking questions.**
  Spiryagin, M., Wu, Q., & Cole, C. (2017). *Vehicle System Dynamics*

- **International benchmarking of longitudinal train dynamics simulators: Results.**
  Wu et al., (2018). *Vehicle System Dynamics*

- **Method**
  9 LTD simulators were compared through 4 different train configurations

- **Findings**
  - All simulators had an agreement in simulations
  - The major differences lie in the draft gear models

Research gaps & Objective

Research gaps

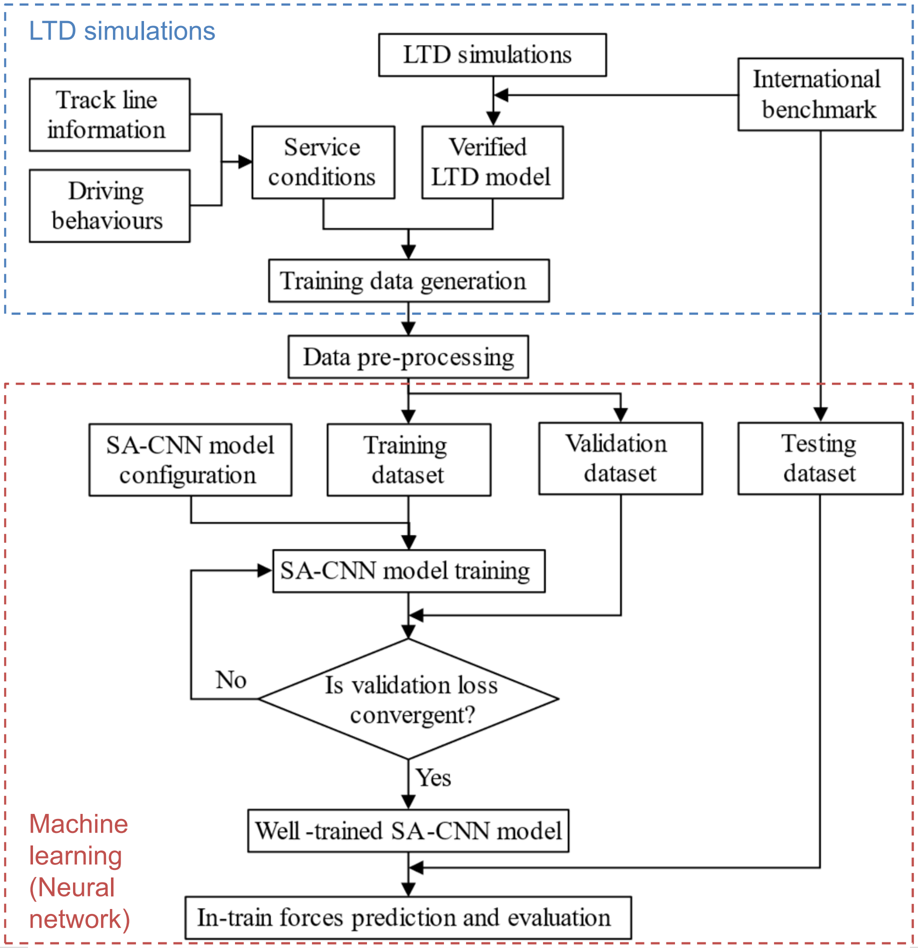
- Traditional in-train forces acquisition methods are either time-consuming or expensive
- ATO systems cannot measure in-train forces
- Machine learning has not been applied to predict the in-train forces
- Gap between the ATO measurements and in-train force prediction

Research objective

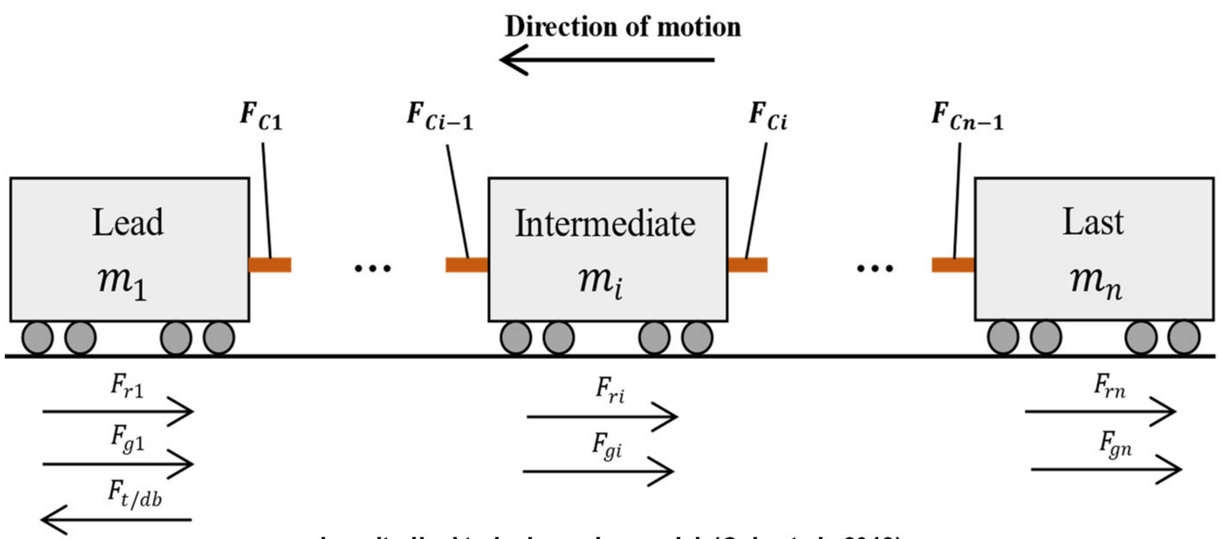
Development of a data-driven approach by combining ATO measurements with a machine learning/neural network model to achieve a real-time or on-board prediction of railway in-train forces.

Methodology -- Overall workflow

- LTD simulations: generating training data due to:
  1. In-train force measurement are not included in ATO systems
  2. Measurements on serval routes cannot stand for all the general service conditions
  3. Randomly generated features (only ML training stage) are difficult to be implemented in real-world
- SA-CNN network: learning the underlying relationship



Methodology -- Longitudinal train dynamics (LTD) modelling



Longitudinal train dynamics model (Cole et al., 2012)

For the lead vehicle: $F_{C1} = F_{t/db1} - F_{g1} - F_{cr1} - F_{pr1} - m_1 a_1$

For the *i*th vehicle: $F_{Ci-1} + F_{Ci} = F_{t/dbi} - F_{gi} - F_{cri} - F_{pri} - m_i a_i$

For the *n*th or last vehicle: $F_{Cn-1} = F_{t/dbn} - F_{gn} - F_{crn} - F_{prn} - m_n a_n$

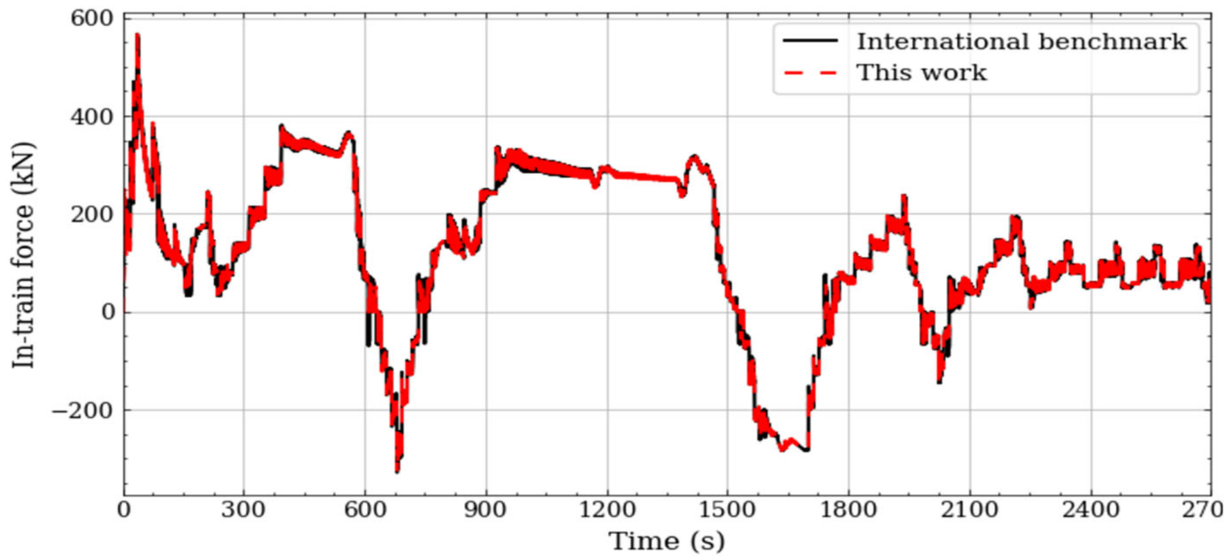
For the *i*-th vehicle:

- F_{Ci-1} : front in-train force
- F_{Ci} : rear in-train force
- m_i : mass
- a_i : acceleration
- $F_{t/dbi}$: traction/DB effort
- F_{gi} : gravitational component
- F_{cri} : curving resistance
- F_{pri} : propulsion resistances (rolling/air resistances)

LTD model verification

Heavy haul rain information (Spiryagin et al., 2017)

Vehicle	Axle-load (tonne)	Axle number	Total mass (tonne)	Overall length (m)
Locomotive	22.33	6	134	22.95
Wagon	32	4	128	15
Configuration	Head-end train: 2 locos + 50 wagons			



In-train forces on the 10th coupler

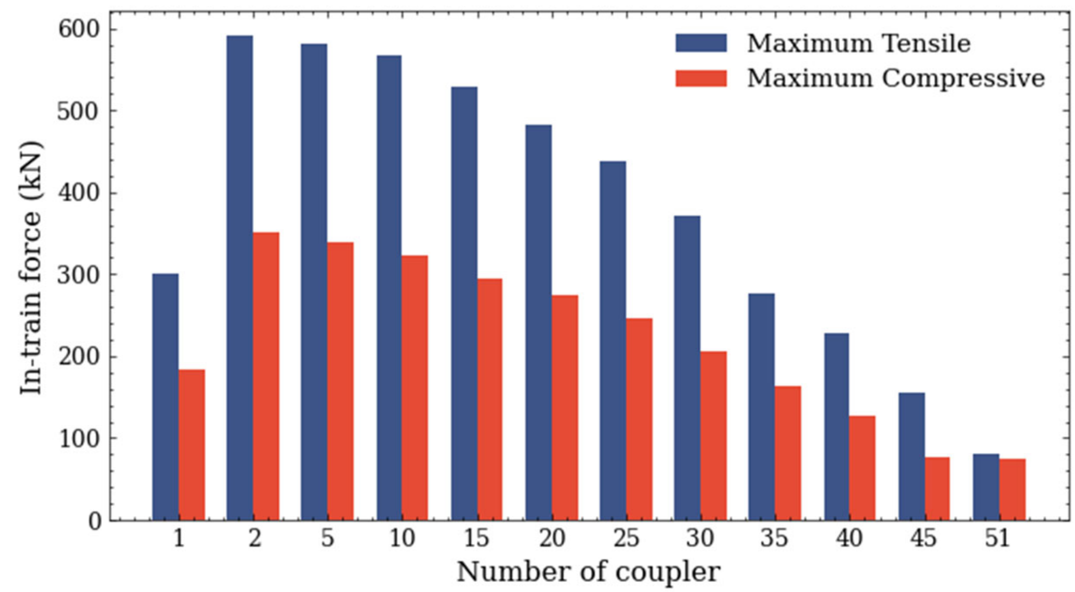
Methodology -- Service conditions & Target coupler

LTD input parameters:

- Heavy haul train model:
  - Tran configurations
  - Rolling stock models
  - Coupling system
- Resistance formulas:
  - Propulsion resistance
  - Curving resistance
  - Gravitational resistance
- Driving behaviours:
  - Traction/Dynamic braking effort
  - Target speed
- Track line conditions:
  - Curvature
  - Gradient
  - Curve length

LTD output results:

- In-train forces for each coupler

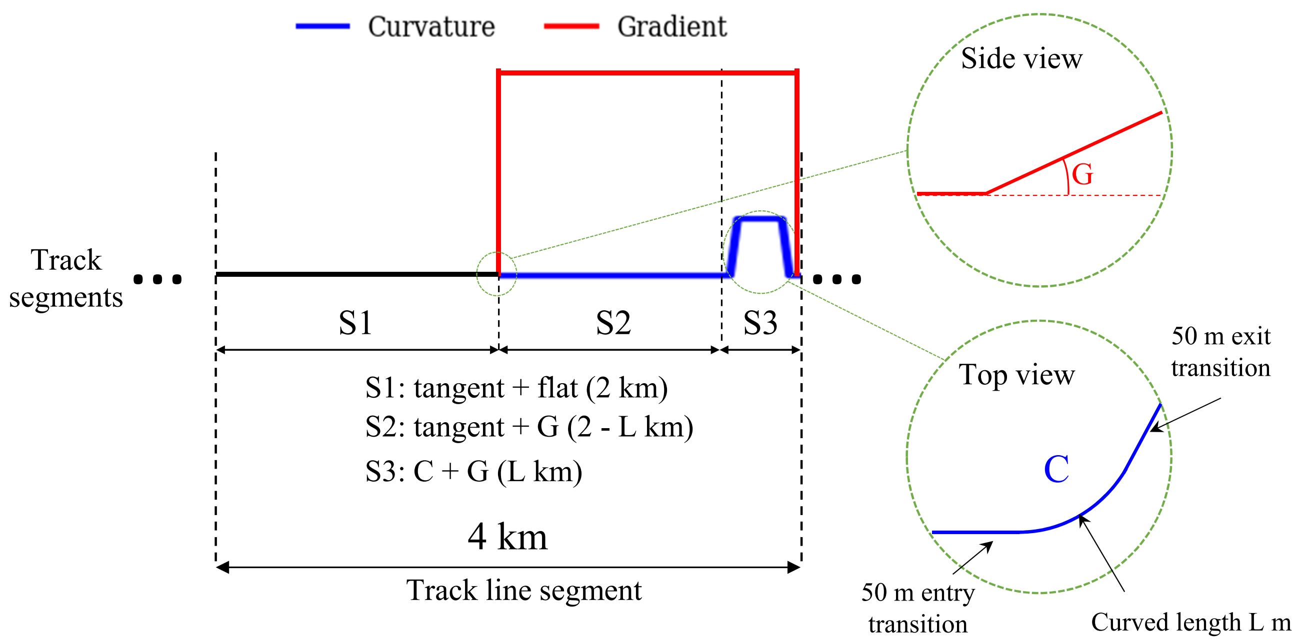


Critical coupler:

- No.2 (behind the locomotives)
- Under the severest working conditions

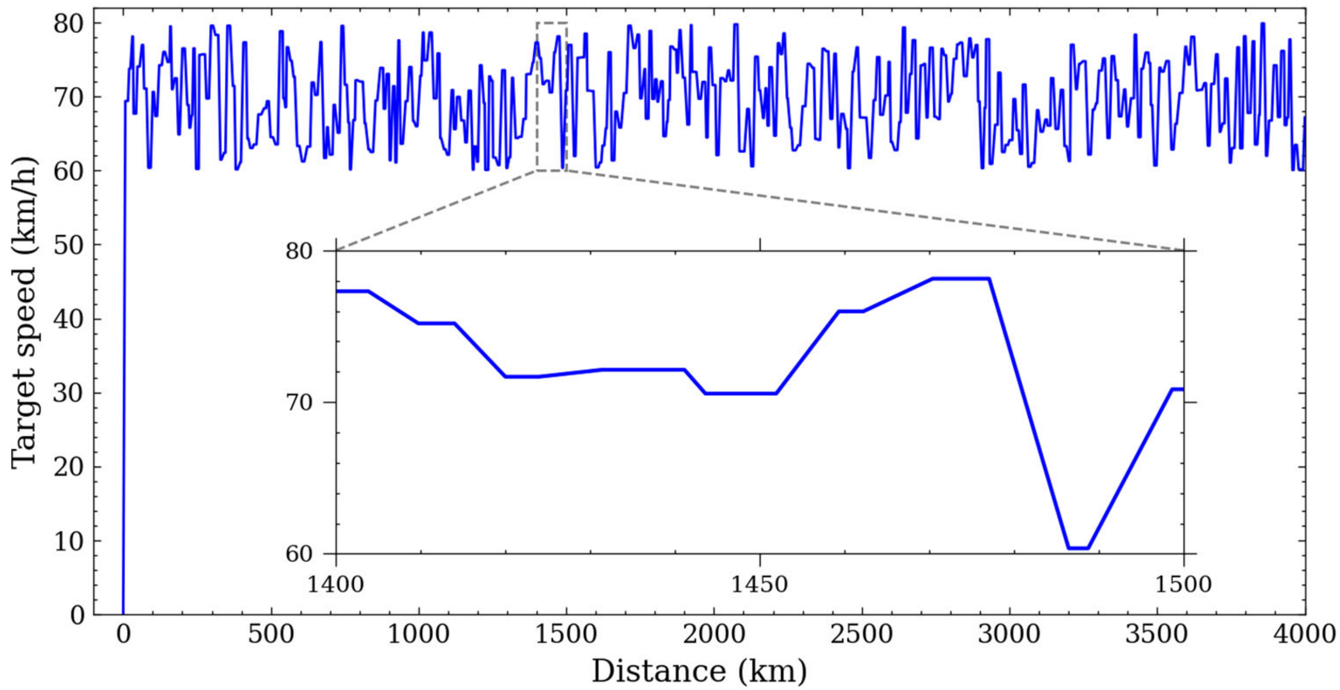
Methodology -- Training data generation

Service conditions	Track line conditions			Driving behaviours			
Parameters	Curve radii (R) (m)	Gradients (G)	Curve length (L) (m)	Low *V* (km/h)	Middle *V* (km/h)	High *V* (km/h)	Extra high *V* (km/h)
Range	[200, 8000]	[-1:100, +1:100]	[200, 900]	[20, 40]	[40, 60]	[60, 80]	[80, 100]



Methodology -- Training data generation (contd.)

Service conditions	Track line conditions			Driving behaviours			
Parameters	Curve radii (R) (m)	Curve length (L) (m)	Gradients (G)	Low *V* (km/h)	Middle *V* (km/h)	High *V* (km/h)	Extra high *V* (km/h)
Range	[200, 8000]	[200, 900]	[-1:100, +1:100]	[20, 40]	[40, 60]	[60, 80]	[80, 100]



Methodology -- Training data generation (contd.)

Data collected:

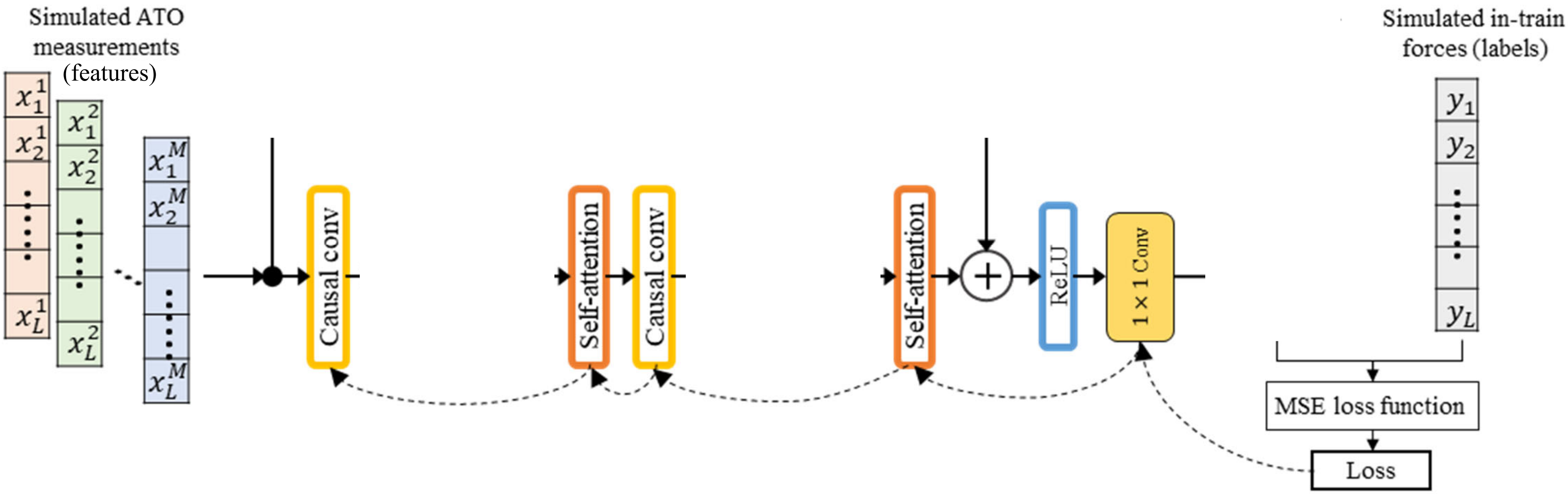
- Number of data: 6,431,681 samples (16,000 km)
- 1 sample = 6 features ($\boldsymbol{x}$, $\boldsymbol{v}$, $\boldsymbol{a}$, $\boldsymbol{F_{t/db}}$, $\boldsymbol{G}$, $\boldsymbol{C}$) and 1 label (in-train force on the 2nd coupler)
- Collection time interval: 0.2 s (discretely)

Speed range	Biased training data		Unbiased training data	
	Track length (km)	Number of samples (percentage)	Track length (km)	Number of samples (percentage)
Low V	4,000	$\boxed{\text{2,668,870 (41.5\%)}}$	1,296	800,000 (25%)
Middle V	4,000	1,650,597 (25.66%)	2,164	800,000 (25%)
High V	4,000	1,230,608 (19.13%)	3,050	800,000 (25%)
Extra high V	4,000	$\boxed{\text{881,606 (13.71\%)}}$	3,630	800,000 (25%)
Total	**16,000**	**6,431,681 (100%)**	**10,140**	**3,200,000 (100%)**

Methodology -- Data pre-processing

Row training data	Features/label pair at discrete time points
Step1: Data segmentation	Sequences with *L* time steps
Step 2: Dataset division	Training (90%); validation (10%)
Step 3: Data nominalisation	Normalise the data values to [-1, 1]
Step 4: Shuffling and Batching	Sequences shuffled and grouped into batches with size of N
Pre-processed data	Features (*N, L, M*); label (*N, L,* 1)

Methodology -- SA-CNN neural network



Causal convolution operation: $output_{conv} = \sigma(b + W * x_{t-k+1})$

Self-attention operation: $output_{atten}(I, W^q, W^k, W^v) = softmax\left(\frac{(I*W^q)\cdot(I*W^k)^T}{\sqrt{d_k}}\right) \cdot (I * W^v)$

Residual connection: $output_{RB} = f(x) + x,$ where $f(x) = atten(conv(atten(conv(x))))$

MSE loss function: $Loss = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$

Results -- Performance comparison with neural networks

The comparative networks are all 4-layer networks

	RMSE (kN)	MAE (kN)	R^2	Training time (s)	Inference time (s)
CNN	5.75	4.16	0.99895	156	4.5
LSTM	4.44	2.53	0.99937	277	3.7
CNN-LSTM	4.88	3.12	0.99924	185	4.8
TCN	4.74	2.83	0.99929	157	4.7
SA-CNN	**4.13**	**2.12**	**0.99946**	**169**	**4.7**

- **RMSE**: root of mean square error w.r.t LTD
- **MAE**: mean absolute error w.r.t LTD
- **R^2**: coefficient of determination w.r.t LTD
- **Training time**: time to train the ML model
- **Inference time**: time to make prediction by the well-trained model

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \qquad R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2}$$

Results -- Performance comparison with LTD simulations

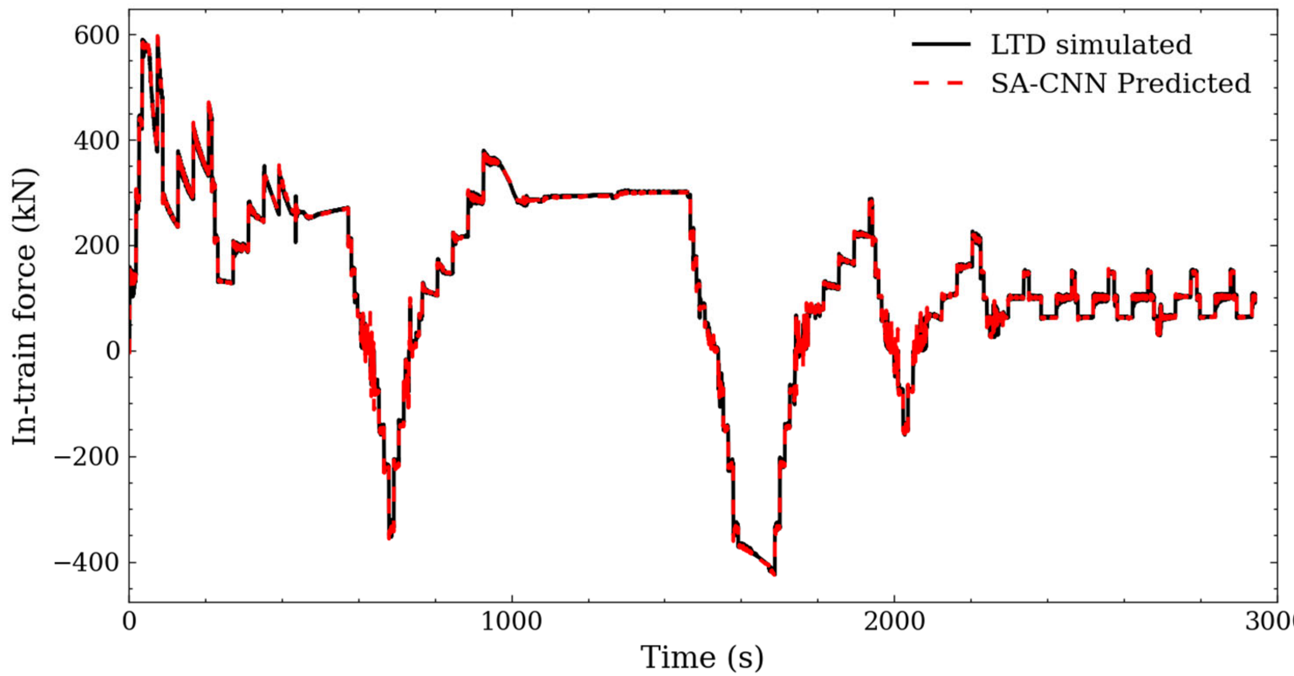
Service condition
(Case 1)

- Driving behaviours:

Controlled by throttle positions:
generated by speed optimisation software

- Track line conditions:

Real-word measured line condition



Results -- Performance comparison with LTD simulations (contd.)

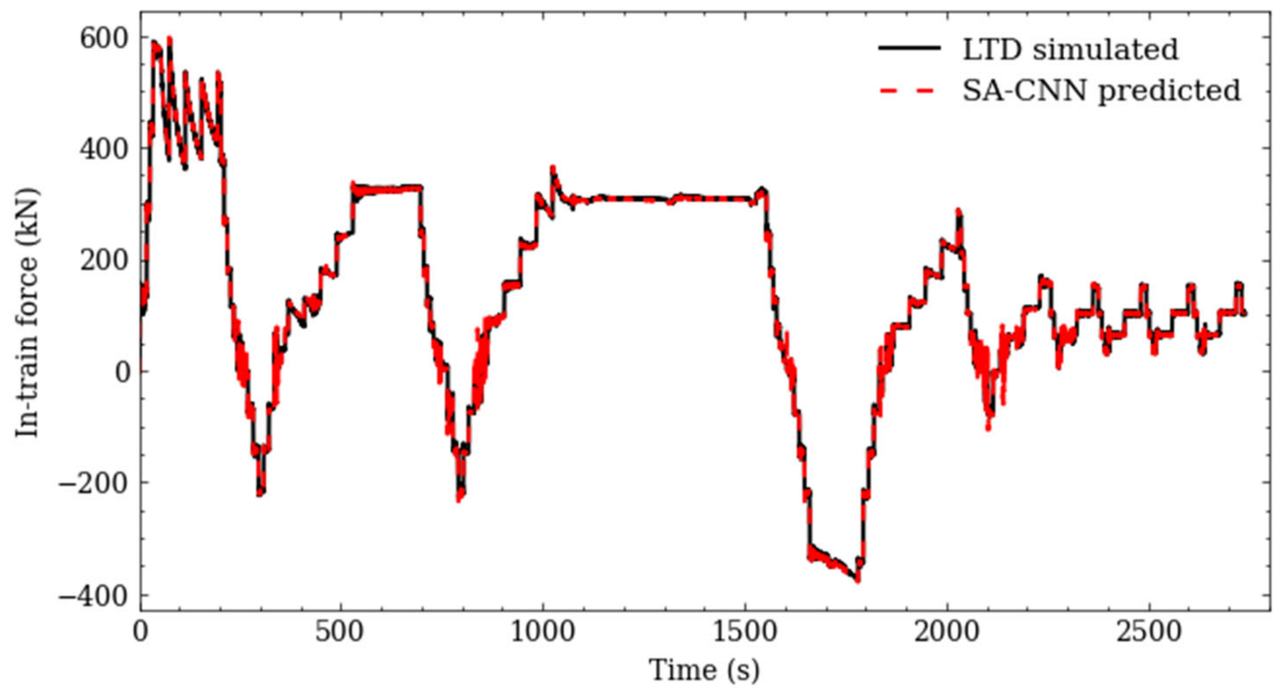
Service condition
(case 2)

- Driving behaviours:

Controlled by throttle positions:
generated by speed optimisation software

- Track line conditions:

Reversed direction of the track line used in the benchmark



Results -- Performance comparison with LTD simulations (contd.)

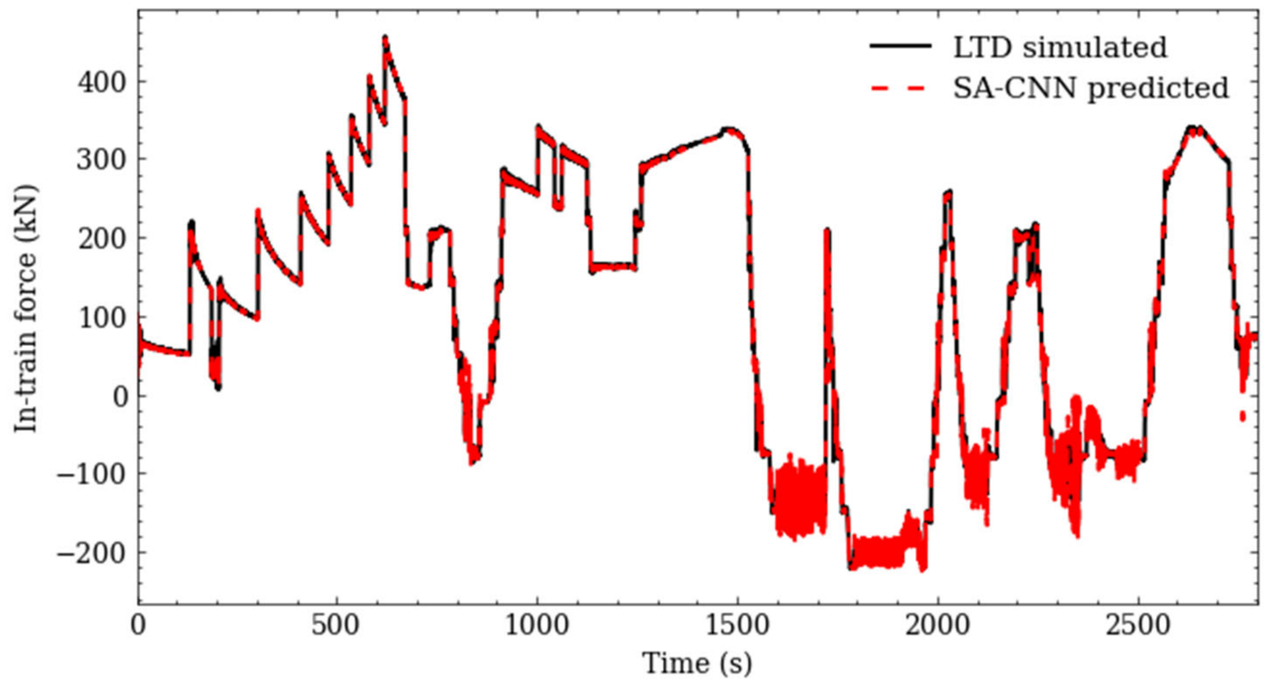
Service condition (case 3)

- Driving behaviours:

Controlled by speed (throttle positions): randomly generated target speed profile

- Track line conditions:

Randomly generated



Results -- Performance comparison with LTD simulations (contd.)

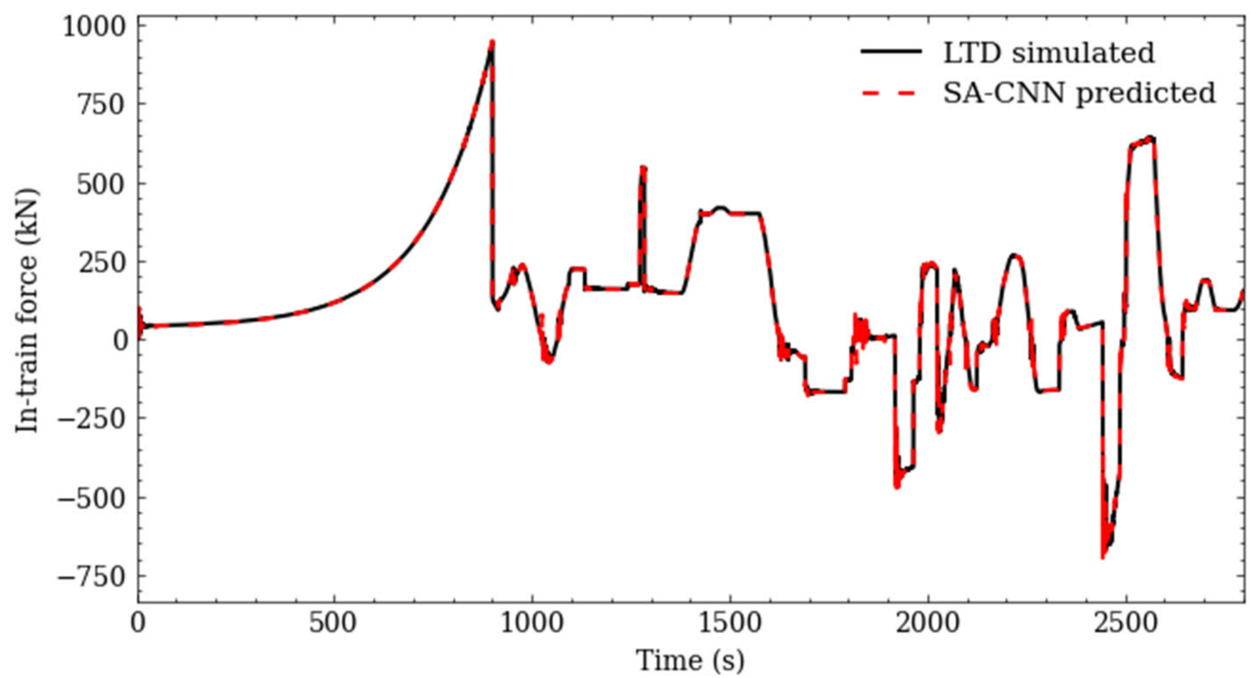
Service condition (case 4)

- Driving behaviours:

Controlled by speed (control force): randomly generated target speed profile

- Track line conditions:

Randomly generated



Results -- Performance comparison with LTD simulations (contd.)

The well-trained SA-CNN model has the same accuracy as LTD simulations but with significantly reduced prediction time

Case	RMSE (kN)	MAE (kN)	R2	SA-CNN training time (s)	SA-CNN inference time (s)	LTD simulation time (s)
Case1	4.13	2.12	0.99946	169	4.7	204
Case 2	4.76	2.41	0.99938		4.9	192
Case 3	7.32	3.9	0.99816		5	225
Case 4	9.53	2.5	0.99848		4.8	307

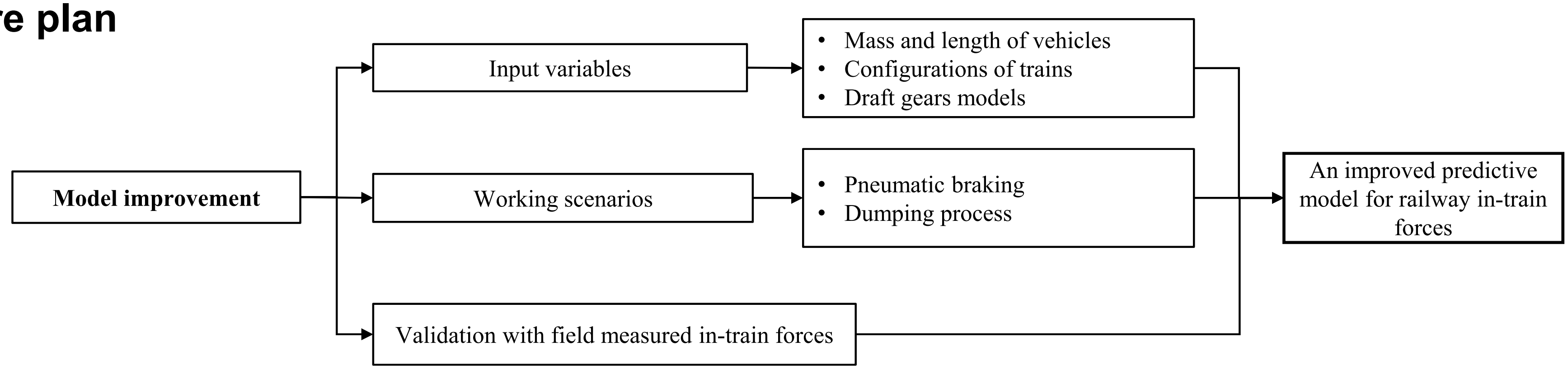
Real-time in-train force monitoring is feasible

Follow-up Researches

Limitations

- Lack of field validation
- Limited coupler considered
- Insufficient working scenarios

Future plan



Conclusions

A data-driven approach (ATO measurements + ML model) was proposed to predict the railway in-train forces

- LTSs were used for establishing the relationship between ATO measurements and in-train forces
- A SA-CNN was developed to learn the relationship between features and labels considering time dependency
- The well-trained SA-CNN are accurate with a quick prediction time
- Proposed approach has the potential to replace the traditional in-train force acquisition methods

MONASH University

Thank you.