UTS

Incident management using an integrated machine learning and dynamic traffic simulation modelling

SPEAKER'S NAME:

Adriana-Simona MIHAITA
University of Technology Sydney, Australia

Collaborators:
Sajjad Shafiei, Hoang Nguyen, Chen Cai, Data61|CSIRO
Michael Towke, Transport for New South Wales
Artur Grigorev, University of Technology Sydney

April 12th 2021

ITS Asia Pacific 2021

FM LAB

www.fmlab.org

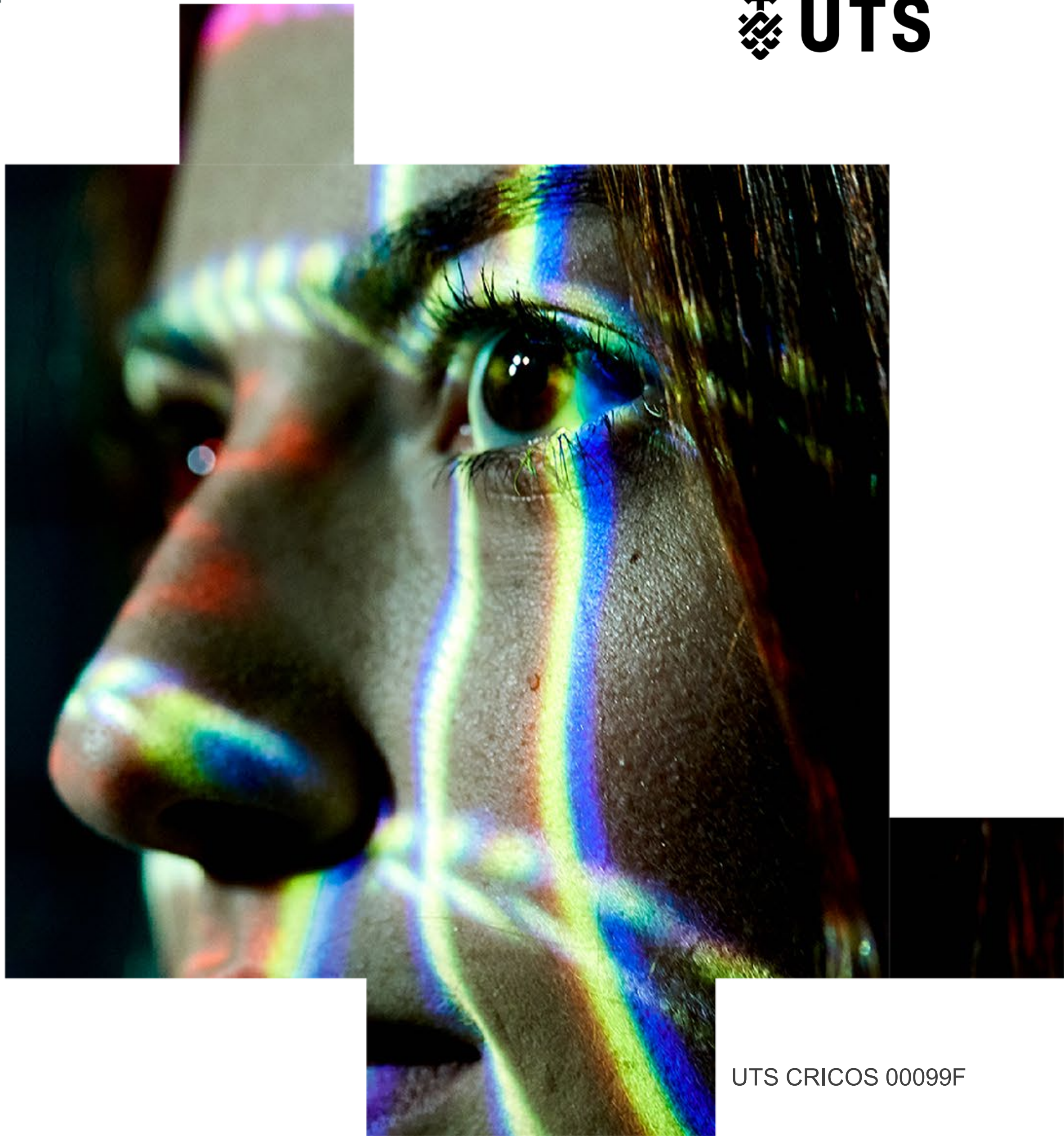
Summary

1. **Introduction**

2. **Methodology**
    1. Incident Management platform
    2. Demand Estimation and prediction via ML

3. **Case Studies and results**
    1. Sydney Victoria Rd
    2. Demand estimation and prediction results
    3. Incident impact Analysis
    4. San Francisco Incident duration prediction

4. **Conclusions**



Introduction

Why is hard to manage incidents?

a) random planned and unplanned events can severely disturb regular traffic conditions
b) the spatial structure and layout of the network can induce high complexity in the localisation of traffic count stations and the reported incidents
c) the spatial and temporal distribution of traffic flow can induce direct and indirect congestion propagation patterns
d) missing or erroneous data due to varying equipment functioning state, or inconsistent human reporting.
e) traffic forecasting is a necessary step for efficient network operation and is an integral part of intelligent transportation systems (ITS) applications.

Introduction

Main critical functionalities for an incident management platform:

a) Provides fast insights into the current traffic states

b) Predicts the evolution of traffic congestion under disruptions

c) Is able to simulate various scenarios of how to manage the disruption in real time

d) Is able to compare and optimise the best response plans that traffic operators need to take in order to clear the affected area as soon as possible

e) Can be used for BOTH intervention and planning purposes.

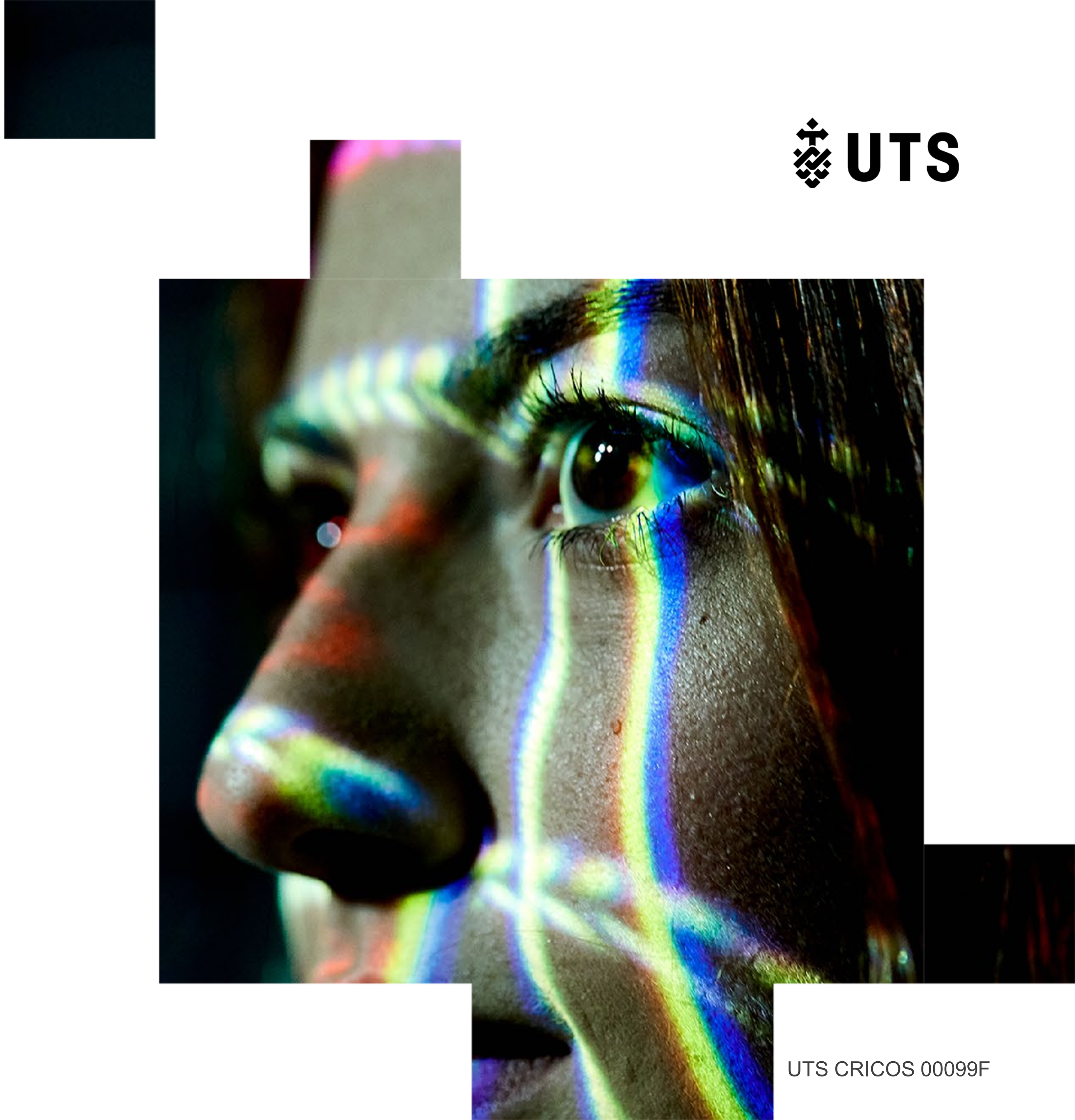
Summary

1. **Introduction**

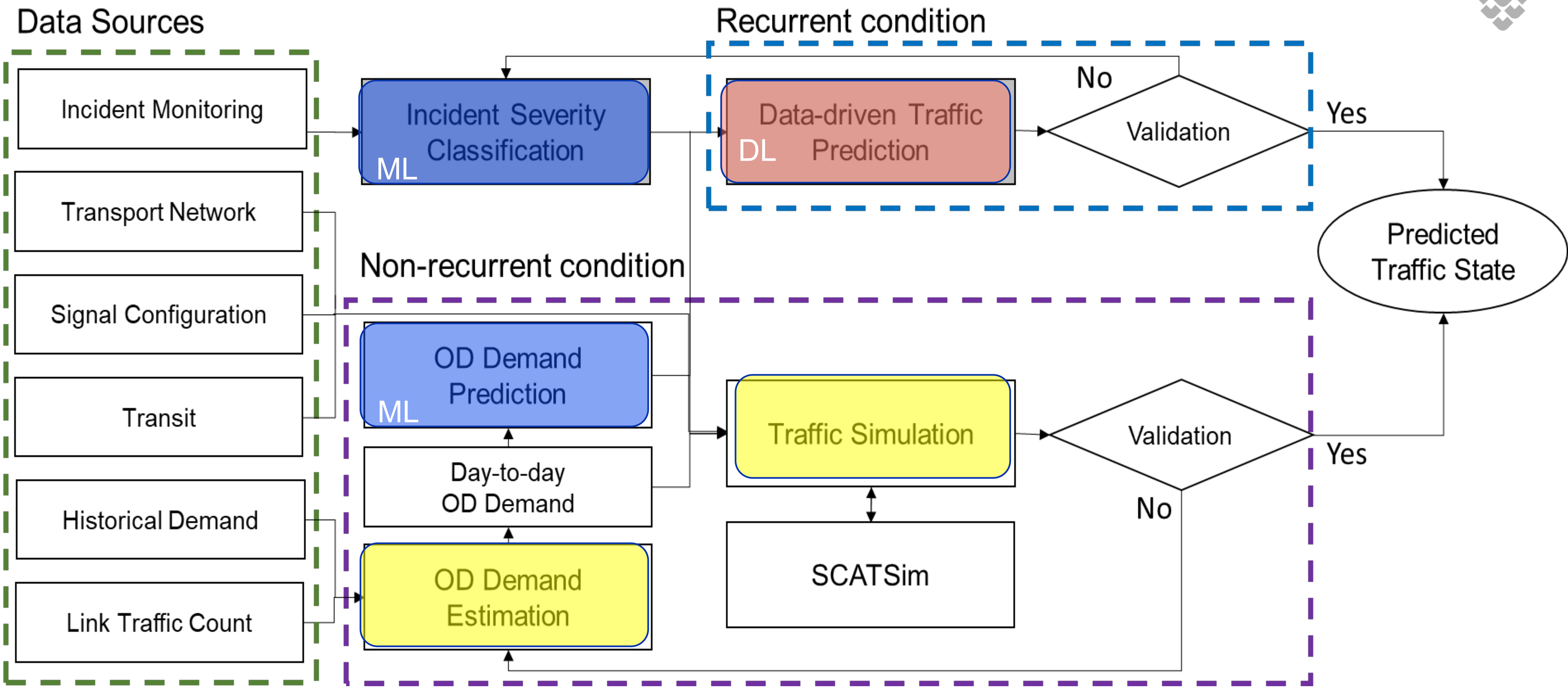
2. **Methodology**
    1. Incident Management platform
    2. Demand Estimation and prediction via ML
    3. Incident Duration Prediction via ML

3. **Case Studies and results**
    1. Sydney Victoria Rd
    2. Demand estimation and prediction results
    3. Incident impact Analysis
    4. San Francisco Incident duration prediction

4. **Conclusions**



Methodology



OD Demand Estimation

- The most crucial input for any DTA models is the origin-destination (OD) trip table.
- The success of the DTA application relies on the quality of OD demand matrices.
- Estimating the OD demand by using link traffic data is a popular approach and far superior to doing the conventional travel surveys which are slow and expensive.
- Many studies proposed a bi-level optimization formulation where the feedback of demand changes is evaluated by an assignment model iteratively .

We express the problem mathematically as follows:

$$\min \omega . \sum_{i\in I}\sum_{t=1}^{T}(x_i^t-\hat{x}_i^t)^2+(1-\omega).\sum_{a\in A}\sum_{t=1}^{T}(y_a^t-\hat{y}_a^t)^2 \qquad \textbf{(1)}$$

$$y_a^h=\sum_{i\in I}\sum_{t=1}^{h}p_{a,i}^{h,t}(X)x_i^t$$

where,

$\hat{x}_i^t$, x_i^t are the initial and estimated demand flow of OD pair *i* (*i* ϵ *I*) at time period *t*,

$\hat{y}_a^t, y_a^t$ are the observed and estimated link flow in link "*a*" at a time period *h* (*a* ϵ *A*, *h*=[1,T]),

$p_{a,i}^{h,t}$ is the assignment proportion of x_i^t that passes link "*a*" during a time period *h,*

ω is the reliability weight on the initial demand data.

OD Demand Prediction

Several Machine Learning models have been applied (Geron, 2015)

- **Support Vector Machines (SVM) –** used for both classification and regression and can deal with noisy data sets
- **Decision Trees (DT) –** are non-parametric supervised ML algorithms which divides the data sets in subsets based on specific thresholds
- **ARIMA –** traditional time series prediction using the autoregressive moving average
- **Extreme Gradient Boosting (XGBoost) –** is a tree based algorithm with a boosting enhancement on the objective function
- **Random Forests (RF) –** ensemble learning methods using a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean average prediction of individual trees (regression),
- **kNN (k-Nearest Neighbors)** – used for both classification and regression which analyses the k closest training examples in the data set,
- **Light Gradient Boosted Model (LGBM)** – uses tree based algorithms**, etc**.

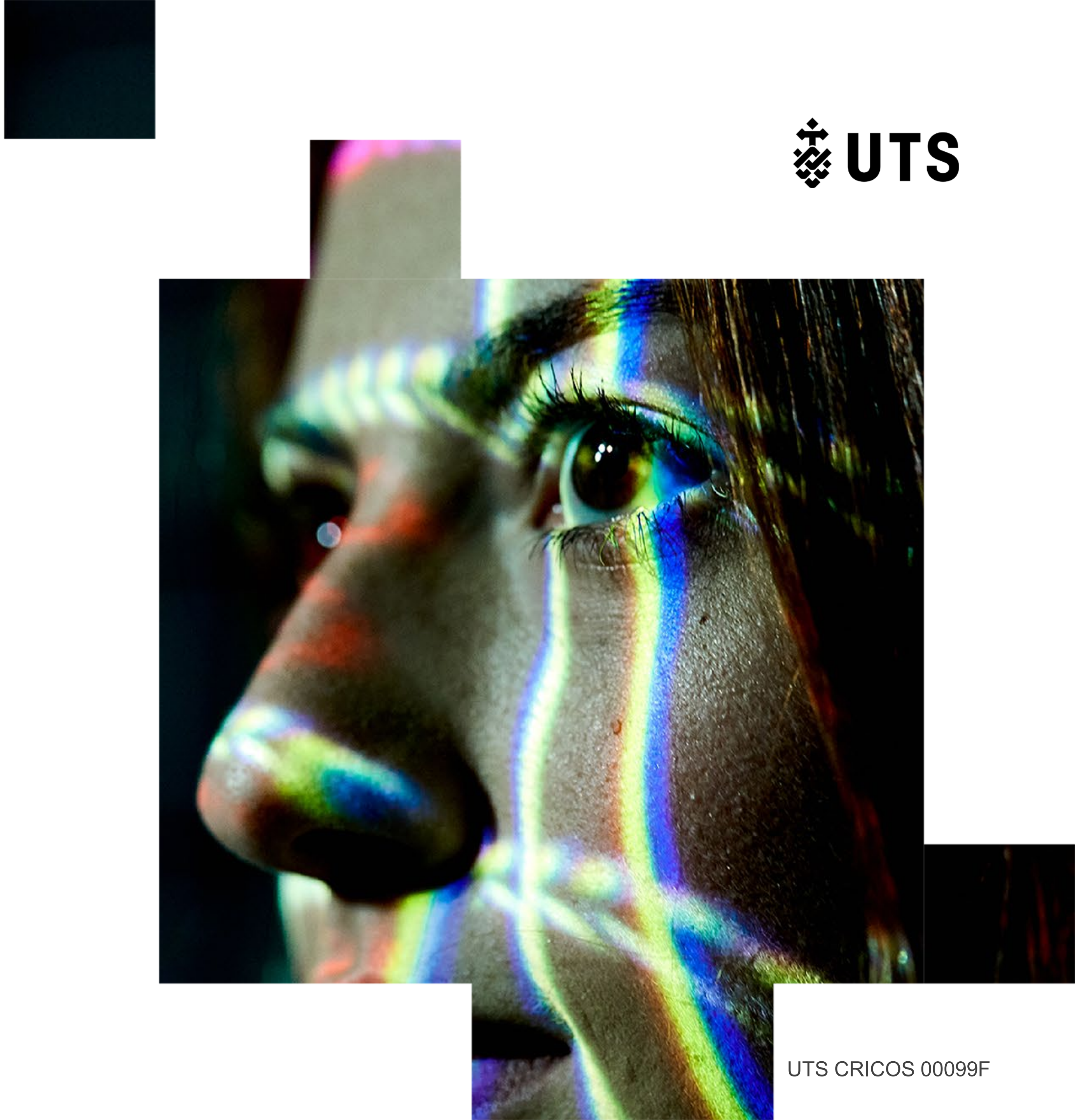
Summary

1. **Introduction**

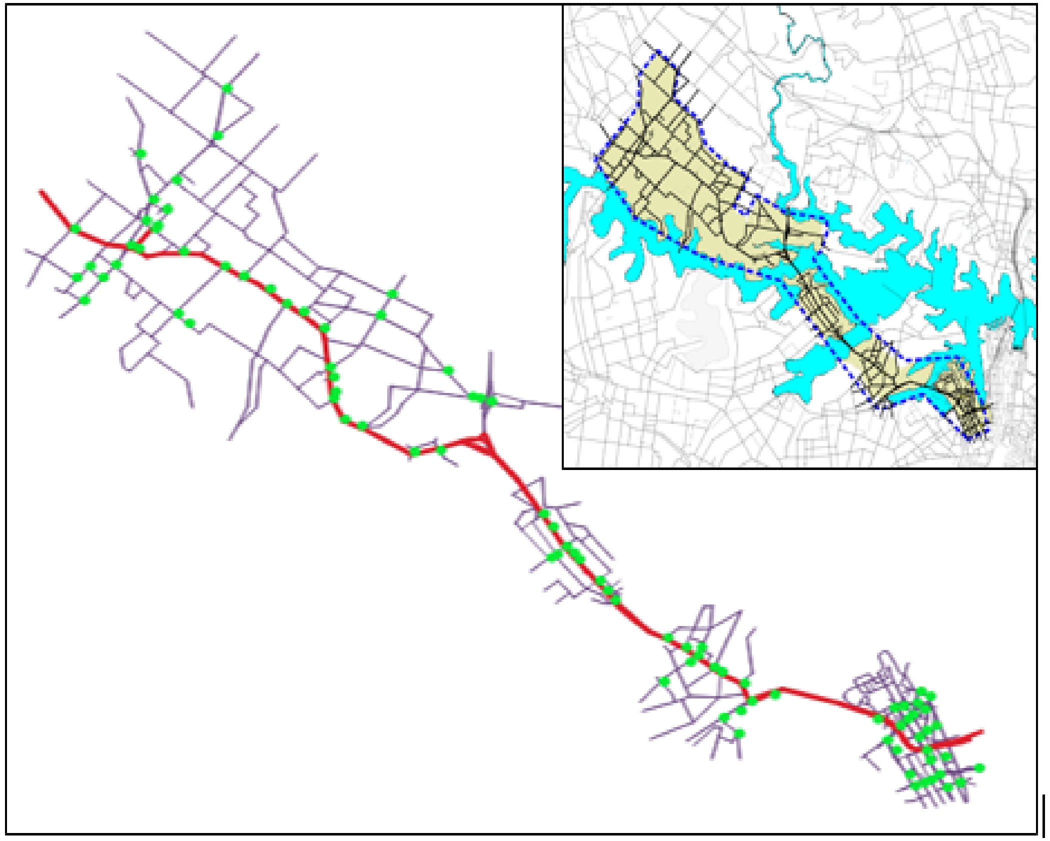
2. **Methodology**
   1. Incident Management platform
   2. Demand Estimation and prediction via ML
   3. Incident Duration Prediction via ML

3. **Case Studies and results**
   1. Sydney Victoria Rd
   2. Demand estimation and prediction results
   3. Incident impact Analysis
   4. San Francisco Incident duration prediction

4. **Conclusions**



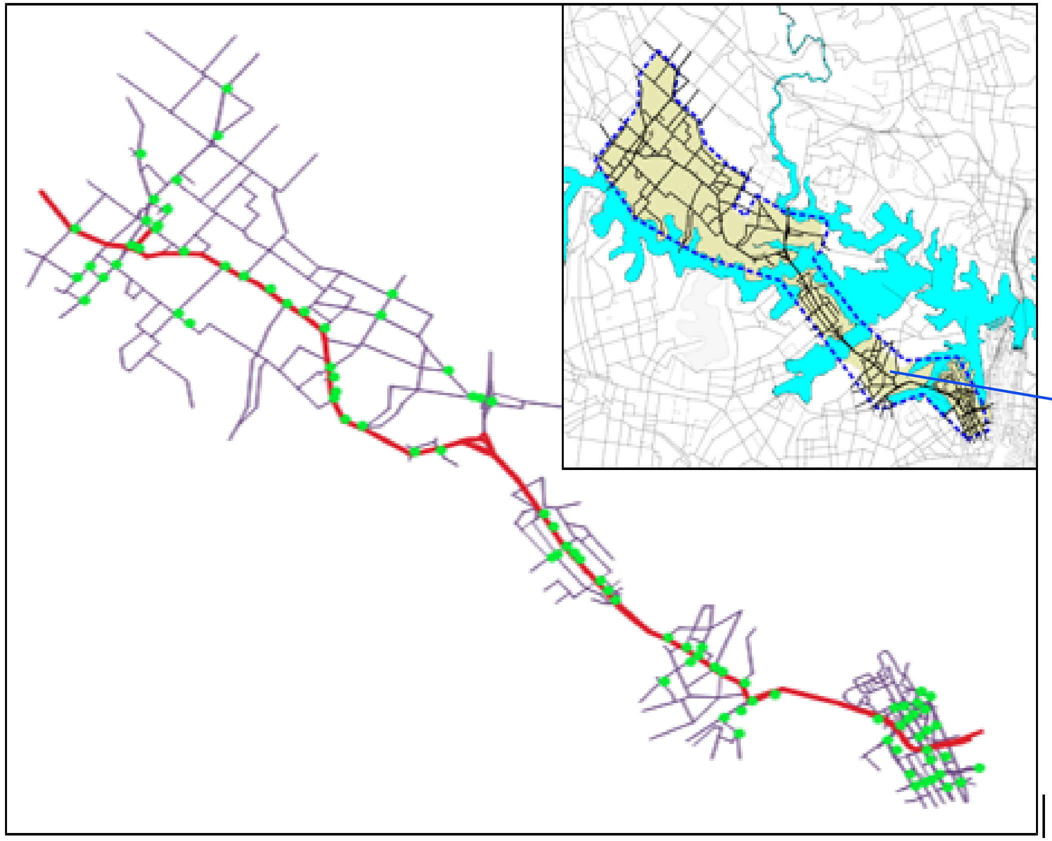
Case studies

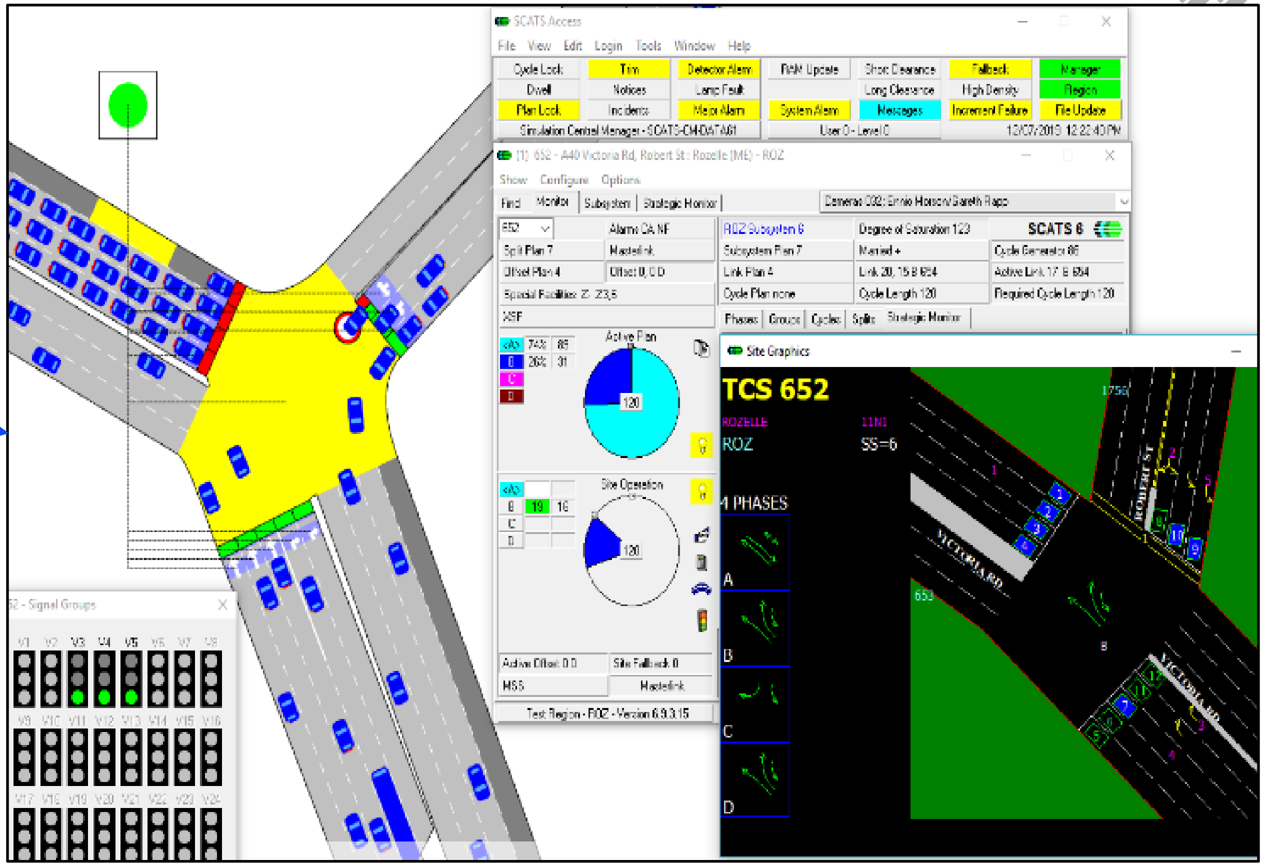


Victoria corridor sub-network. Green points show signals equipped with SCATS count detectors (measured traffic data). Red line demonstrates the main Victoria corridor

- 1,310 links and 428 nodes
- 81 signalized intersections with the adaptive SCATS control system running
- 4-h morning peak hours from 6:00 to 10:00 AM.
- 1,262 OD pairs with various demand profiles – one ML for each OD pair
- modified multinomial logit model as an advanced stochastic route choice model (Aimsun, 2013)
- Maximum five shortest paths are calculated using Dijkstra's label-setting algorithm.
- Aimsun modelling tool.

Case studies





Victoria corridor sub-network. Green points show signals equipped with SCATS count detectors (measured traffic data). Red line demonstrates the main Victoria corridor

Case studies

Input transport data for the microsimulation model

- **Traffic count**: SCATS (arterials) traffic counts,
- **Public transport plans** – using GTFS data,
- **Signal control plans** –using SCATS signal timing,
- **Travellers demand:** Origin-destination of private vehicle users,
- **Incident data logs**: includes the incident location ([x,y] coordinates),

Outputs of the microsimulation model

- network or individual link travel times,
- assigned link traffic volumes,
- network/link delay
- speed/density, etc.

Case studies

Demand Estimation Results and model validation

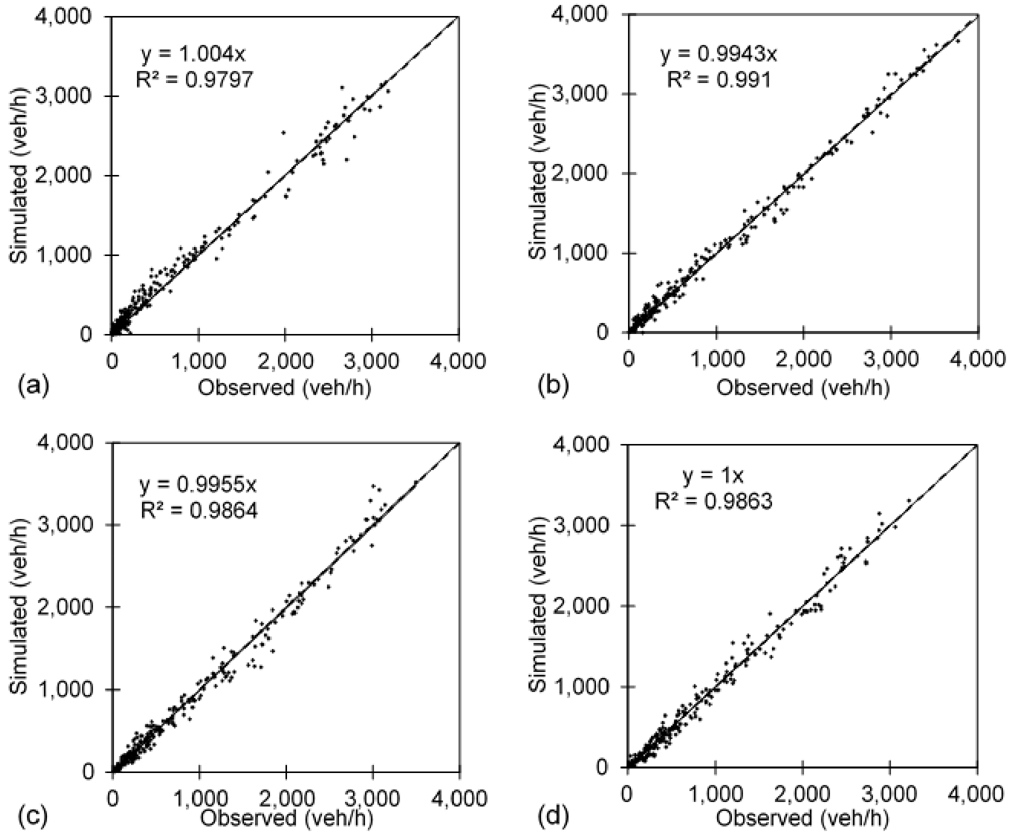


Figure 6. Simulated versus observed traffic flows after the rolling horizon OD estimation (a) 6:00-7:00 am, (b)7:00-8:00 am, (c) 8:00-9:00 am and (d) 9:00-10:00 am

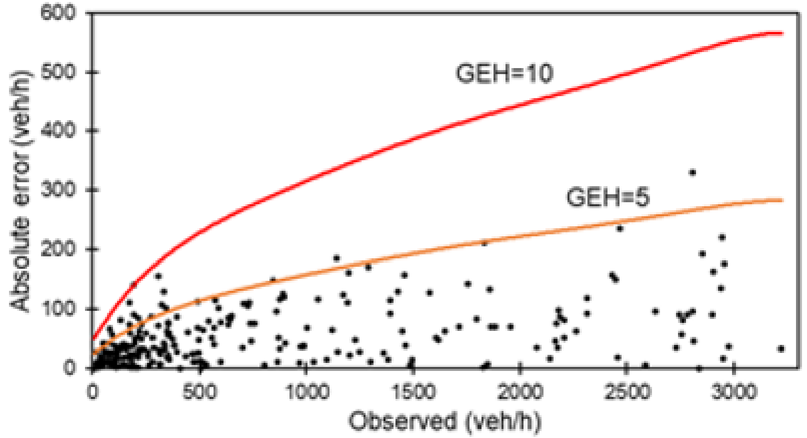


Figure 7. Absolute flow errors versus observed flows.

Case studies

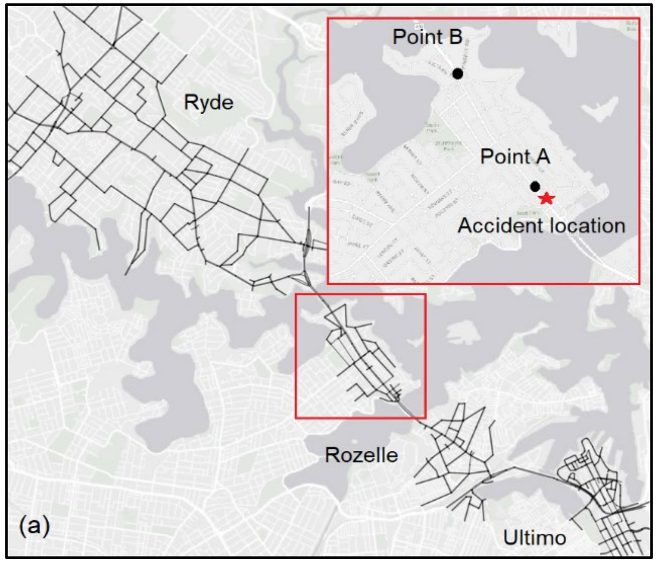
Demand Prediction results

Table 2. Predicted demand using different predictors.

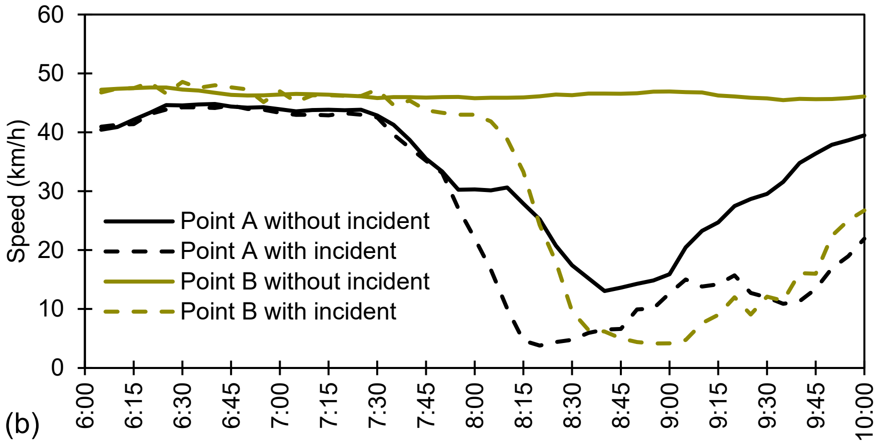
Predictor		Prediction window							
		15 min		30 min		45 min		60 min	
		Total error	MAE	Total error	MAE	Total error	MAE	Total error	MAE
Baseline									
		1283	1.37	2785	1.49	4519	1.61	6953	1.85
DT									
	Max=2	762	0.81	2785	0.79	2278	0.81	3077	0.82
	Max=5	606	0.65	1152	0.61	1743	0.62	2533	0.68
	unconstraint	660	0.70	1216	0.65	1816	0.65	2645	0.70
SVR									
	RBF (C=0.1)	1381	1.47	2613	1.40	3738	1.33	4797	1.28
	Sigmoid (C=0.1)	1592	1.71	3017	1.67	4519	1.54	5545	1.48
	Linear (C=0.1)	832	0.89	1676	0.89	2462	0.87	3327	0.89
	Linear (C=1.0)	852	0.91	1693	0.90	2490	0.89	3376	0.90
ARMA									
	(1,0,0)	869	0.93	1652	0.88	2459	0.87	3377	0.90
	(2,0,0)	883	0.94	1678	0.90	2517	0.89	3509	0.93
	(0,0,1)	1009	1.08	1876	1.00	2762	0.98	3864	1.03
XGBoost									
	Tree	556	0.59	1033	0.55	1579	0.56	2237	0.60
	Linear	1141	1.21	2336	1.24	3471	1.23	4640	1.24

Case studies

Incident Impact analysis



(X, Y)	(9684462, 4425168)
Date	11/10/2017
Start Time Plain	7:58:19 AM
End Time Plain	8:23:00 AM
Incident Description	Accident: Accident
Location Description	VICTORIA RD PARK AVE DRUMMOYNE 2047 CANADA BAY (LGA) NSW
Direction	BOTH DIRECTIONS
Affected Lanes	ALL LANES
Operator	----
ICEMS Suburb	DRUMMOYNE
rowid	A6C4B8F8-5617-4E12-B170-97323



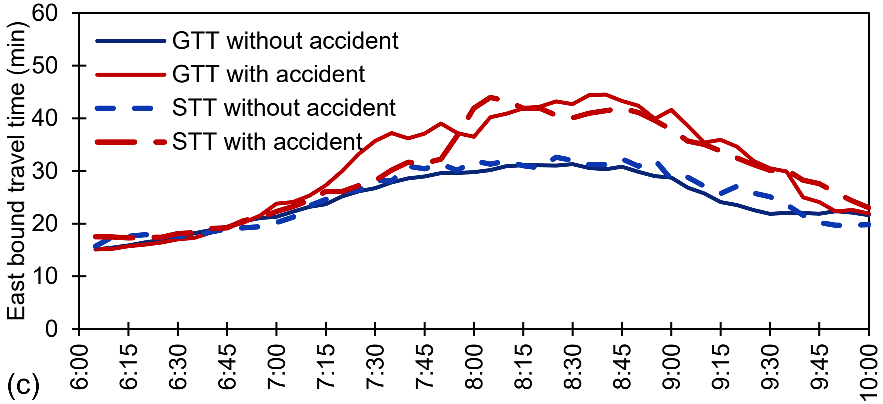
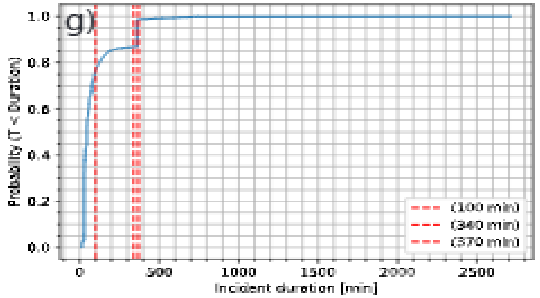
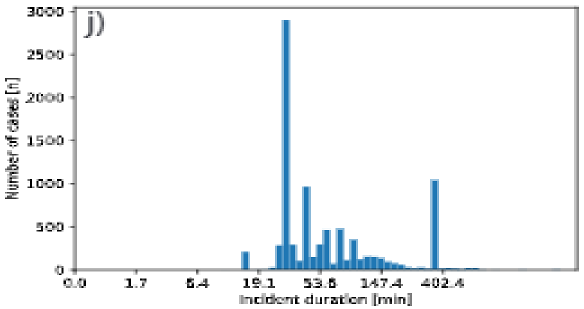


Figure 8. (a) Accident location (b) Speed profiles for Point A and B. (c) Eastbound journey times with and without the incident for both Google Travel Time and Experienced Travel Time calculation.

Case studies

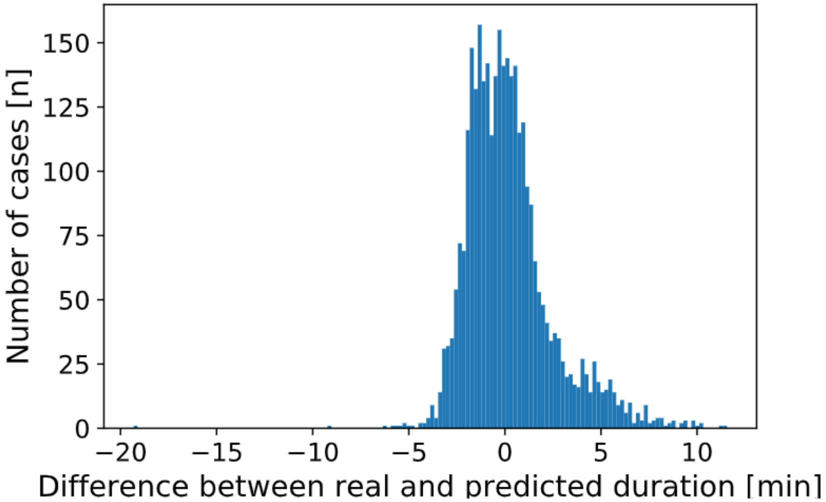
San Francisco Incident Duration Prediction





San-Francisco, All roads, All-to-all

Model	baseline	LT	iIF	eIF	iLOF	eLOF	LT-iIF	LT-eIF	LT-iLOF	LT-eLOF	Winner
LGBM	30.6	29.6	32.2	30.9	31.0	33.7	30.1	29.7	30.3	29.1	LT-eLOF
RF	38.2	28.4	39.3	38.9	36.1	36.4	29.2	28.9	29.1	28.5	LT
LR	131.4	70.9	133.0	133.1	127.6	127.3	71.5	71.3	70.8	69.3	LT-eLOF
GBDT	43.4	30.1	45.3	44.9	43.9	42.2	32.1	31.4	32.7	31.4	LT
KNN	64.0	47.3	64.0	63.9	61.7	62.8	47.8	46.7	47.0	46.2	LT-eLOF
XGB	38.0	31.7	36.8	35.6	33.7	38.6	33.1	31.2	33.0	31.4	LT-eIF
Best	LGBM	RF	LGBM	LGBM	LGBM	LGBM	RF	RF	RF	LGBM	RF-LT-eLOF



Our models can predict with an accuracy of +/- 5 min the duration of an accident in San Francisco

Contributions

- Proposed a prototype for an incident management platform using integrated data-driven and dynamic traffic simulation modelling;
- Estimated day-to-day OD flows for the OD demand prediction module when an incident occurs. We propose several machine learning models for OD demand prediction to reinforce the traffic simulation;
- Deployed traffic micro-simulation modelling according to real-life adaptive signal control by applying the same controllers' logics to simulated vehicles.
- Predicted incident durations using several machine learning models with enhanced features and obtained 5min accuracy in prediction for San Francisco.

Publications

Shafiei, S.; Mihaita, A.S.; Nguyen, H.; Cai, C.; A data-driven and traffic simulation approach to predict the short-term traffic state affected by non-recurrent incidents, Transportation Letters: the International Journal of Transportation Research (IF=1.84), Accepted 8th of April 2020, DOI: 10.1080/19427867.2021.1916284

Mihaita, A.S., Liu, Z., Cai, C., Rizoiu, M.A "Arterial incident duration prediction using a bi-level framework of extreme gradient-tree boosting", ITS World Congress 2019, Singapore, 21-25 Oct 2019, H5=11, Preprint link

Grigorev, Ar., Lee, S., Ch, F., Mihaita, A.S., Modelling impact of traffic disruptions by using machine learning methods, Transport Research Association for NSW Symposium, 18th of November 2020, Presentation link .

UTS

QUESTIONS or Collaboration ideas?

Adriana-simona.Mihaita@uts.edu.au

FM LAB

www.fmlab.org