CAREC
Central Asia Regional Economic Cooperation Program

21st Transport Sector Coordinating Committee Meeting

22–23 April 2024 • Almaty, Kazakhstan

21-е заседание Координационного комитета по транспортному сектору

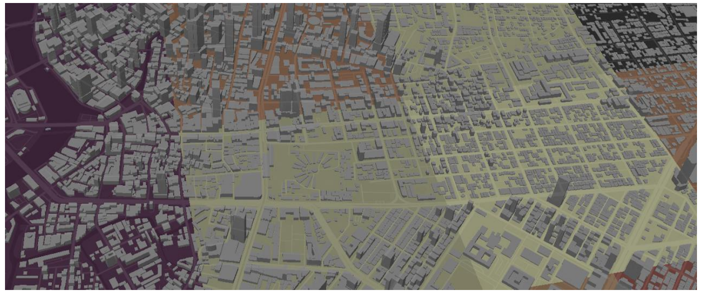
22–23 апреля 2024 года • Алматы, Казахстан

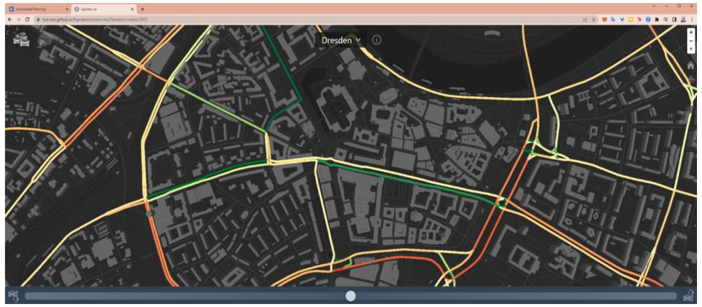


Big data-assisted transport network modeling

Dr. S. Travis Waller
Managing Director
Mobility Thinking Pty Ltd

Automated Traffic Modeling for Rapid Planning





Traditional strategic traffic network models can take many months or years

Data collection, verification of network properties, calibration, and validation

"Big Data" has emerged which can drastically cut the costs and time

A model can be built within weeks

However, data are not models

Even data analytics is not a model

We have implemented new methods to automatically produce "What-if models" capable of **hypothetical planning**

Outcome: *Automate* *traditional models, not* *replace*

*ST Waller, S Chand, A Zlojutro, D Nair, C Niu, J Wang, X Zhang, and VV Dixit **(2021) "Rapidex: A novel tool to estimate origin–destination trips using pervasive traffic data"** Sustainability (Switzerland), vol. 13, pp. 11171 – 11171. https://doi.org/10.3390/su132011171

D Ashmore, ST Waller, K Wijayaratna, and A Tessler **(2022) "Automated Planning For The Strategic Management of Transport Systems In Developing Countries"** Australasian Transport Research Forum Proceedings 28-30 September, Adelaide, Australia. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4191661

S Chand, ST Waller, and D Ashmore **(2022) "Building and Benchmarking Equitable Infrastructure Systems in the Wake of Rapid Urbanisation"** Policy Brief for Task Force 8: Inclusive, Resilient, and Greener Infrastructure Investment and Financing, T20 Summit, Indonesia. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4203715

*ST Waller, M Qurashi, A Sotnikova, L Karva, S Chand **(2023) "Analyzing and modeling network travel patterns during the Ukraine invasion using crowd-sourced pervasive traffic data"** Transportation Research Record, Volume 2677, Issue 10, https://doi.org/10.1177/03611981231161622

R Amrutsamanvar, S Chand, M Qurashi, and ST Waller **(2023) "Rapid Planning: Opportunities with Pervasive Data for Sustainable Mobility"** IEEE Smart Cities Symposium, Prague.

Past Research & Development Teams

2003 - 2011 (Univ. of Texas at Austin)

2011 - 2022 (UNSW, Sydney)

Methods developed across many years and multiple universities.

100+ Researchers, PhDs & Students

More than 40 funding and collaborative partners including:

NVIDIA, U.S. National Science Foundation, Australia Research Council, U.S. Federal Highway Administration, U.S. DOT, TfNSW, Mitsubishi Heavy Industries, Advisian, GoGet Carshare in addition to many other government agencies, software companies, infrastructure firms, advisory firms, banks, insurance companies, startups, etc.

Mobility Thinking Pty Ltd (MOTH)

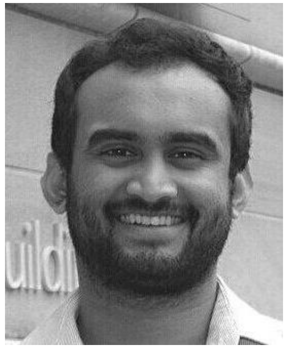
Professor S. Travis Waller
(Managing Director)

Victor Prados-Valerio
(Creative Partner)

Dr. Melissa Duell
(Associate Creative Director)

Dr. David Ashmore
(Creative Partner - MOTH Europe)

Dr. Cecilia da Rocha
(Creative Director)

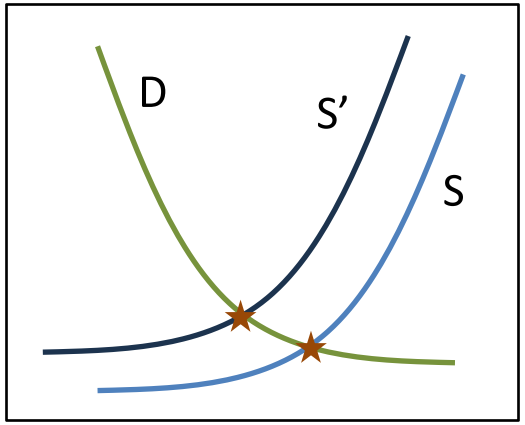


Dr. Kasun Wijayaratna
(Creative Partner)

- Methods transitioned from university for commercialization
- MOTH established in 2018 in Australia

The Need for Planning Models

- Transportation system behaviour
  - Responds non-linearly to changes
  - Is the aggregate response of thousands to millions of individuals making their own self-optimizing decisions
  - Therefore, it is traditionally represented as an equilibrium system
  - Models employ market dynamic explanations



- As a result
  - An underlying equilibrium-based mathematical model has traditionally been necessary for transportation planning and business cases
  - The global universal approach since the 1950s has been the “four step process” for transportation modelling

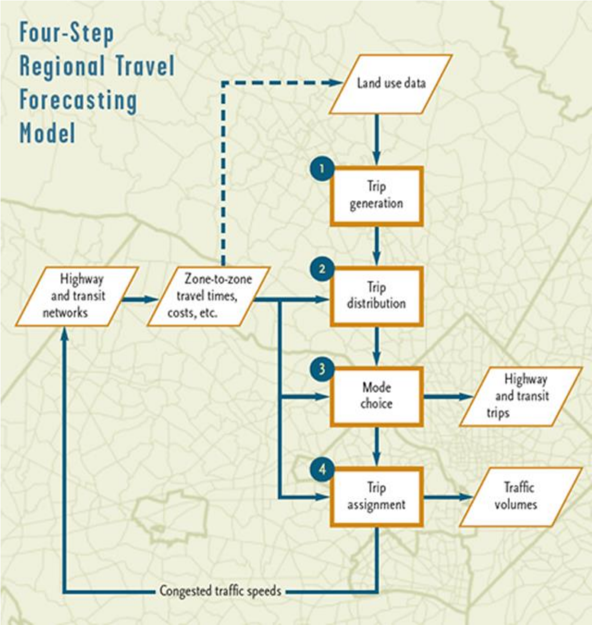
The four-step travel model is a ubiquitous framework for determining transportation forecasts that goes back to the 1950s. It was one of the first travel demand models that sought to link land use and behavior to inform transportation planning. (McNally, 2000)

Traditional Four-step Model for Transportation Planning

The approach is so common, there are Wikipedia pages on each step

- https://en.wikipedia.org/wiki/Trip_generation
- https://en.wikipedia.org/wiki/Trip_distribution
- https://en.wikipedia.org/wiki/Mode_choice (used when multiple modes are in scope)
- https://en.wikipedia.org/wiki/Route_assignment

- Practically, the process often includes
  - Initial step: household travel survey
  - Physical network monitoring (roadway counts, etc)
  - Ongoing network coding and information archiving of infrastructure
  - Ongoing model calibration
- At the final step, traffic assignment, the model estimates or predicts
  - Traffic metrics (volumes, speeds, travel times)
- Because of the need for survey and ongoing monitoring
  - The overall traditional process can consume months or even years



Metropolitan Washington Council of Governments. https://www.mwcog.org/transportation/data-and-tools/modeling/four-step-model/ (Accessed April 2024)

Also, https://www.transitwiki.org/TransitWiki/index.php/Four-step_travel_model (maintained by UCLA and Caltrans)

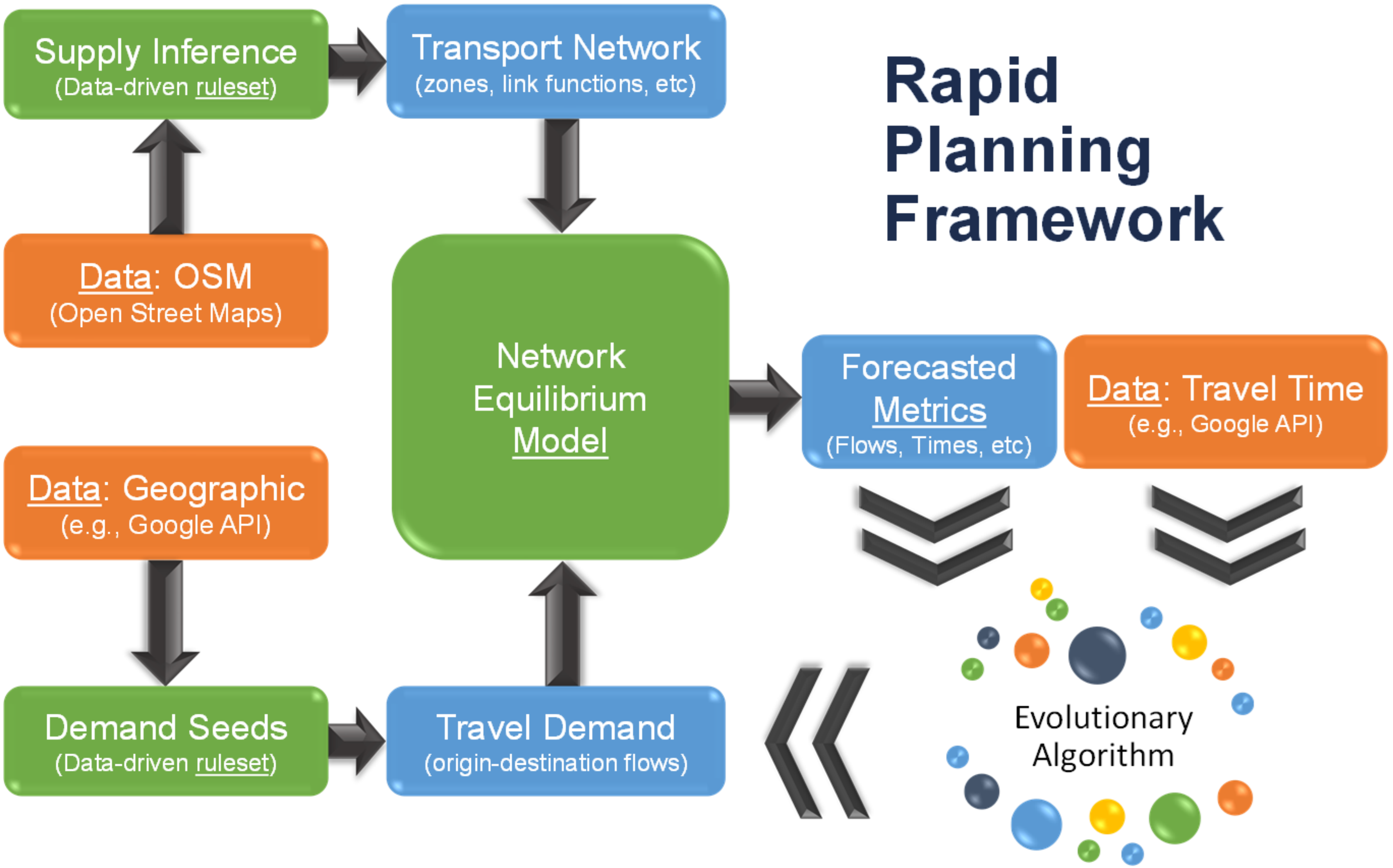
Key Innovation for the Presented Modelling Methodology

- From pervasive data: we begin at the 4th step with traffic metrics, then use machine learning/AI to estimate the travel demand
- The relevant steps are run in reverse (without the need for surveys or ongoing network monitoring)
- Critical: We maintain the traffic assignment and trip modelling steps
  - A key difference from purely data analytic or statistical approaches which do not utilize the step-models at all!

Rapid Transport Planning: Methodological Framework

Waller et al. (2021)

- Use crowd sourced and pervasive data
- Network inference tools to automatically develop planning network from OSM and historic data on transport capacities.
- A Machine Learning, Evolutionary Algorithm, implemented to infer aggregate origin-destination travel demand forecast from observed data.



Over 20 years experience on Evolutionary Algorithms (Machine Learning)

A sampling of peer-reviewed scientific journal publications

Traffic Signal Optimization

Sun D; Benekohal RF; Waller ST (2003) **'Multi-objective traffic signal timing optimization using non-dominated sorting genetic algorithm II'**, Lecture Notes in Computer Science, vol. 2724, pp. 2420 - 2421, http://dx.doi.org/10.1007/3-540-45110-2_143

Sun D; Benekohal RF; Waller ST (2006) **'Bi-level programming formulation and heuristic solution approach for dynamic traffic signal optimization'**, Computer-Aided Civil and Infrastructure Engineering, vol. 21, pp. 321 - 333, http://dx.doi.org/10.1111/j.1467-8667.2006.00439.x

Vending Machine Allocation

Grzybowska H; Kerferd B; Gretton C; Travis Waller S (2020) **'A simulation-optimisation genetic algorithm approach to product allocation in vending machine systems'**, Expert Systems with Applications, vol. 145, http://dx.doi.org/10.1016/j.eswa.2019.113110

Transport Network Design

Jeon, K., J.S. Lee, S. Ukkusuri, and S.T. Waller (2009) ‘**New approach for relaxing computational complexity of discrete network design problem using selectorecombinative genetic algorithm**’ Journal of the Transportation Research Board, Vol 1964, Issue 1, pp. 91-103, 2006. https://doi.org/10.1177/0361198106196400111

Lin DY; Unnikrishnan A; Waller ST (2009) **'A genetic algorithm for bi-level linear programming dynamic network design problem'**, Transportation Letters, vol. 1, pp. 281 - 294, http://dx.doi.org/10.3328/TL.2009.01.04.281-294

Lin DY; Waller ST (2009) **'A quantum-inspired genetic algorithm for dynamic continuous network design problem’**, Tr. Letters, v. 1, pp. 81 - 93, http://dx.doi.org/10.3328/TL.2009.01.01.81-93

Ready-Mixed Concrete Delivery

Maghrebi, M., Periaraj, V., Waller, S. T., & Sammut, C. (2014) ‘**Solving Ready-Mixed Concrete Delivery Problems: Evolutionary Comparison between Column Generation and Robust Genetic Algorithm’** In R. Issa (Ed.), ASCE - Computing in Civil and Building Engineering. Orlando, USA, 23-25 Jun 2014. https://doi.org/10.1061/9780784413616.176

Maghrebi M; Waller ST; Sammut C (2014) **'Sequential Meta-Heuristic Approach for Solving Large-Scale Ready-Mixed Concrete–Dispatching Problems'**, Journal of Computing in Civil Engineering, vol. 30, pp. 04014117 - 04014117, http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000453

Rapid Transport Modelling (including network and trip estimation)

Waller ST; Chand S; Zlojutro A; Nair D; Niu C; Wang J; Zhang X; Dixit VV (2021) **'Rapidex: A novel tool to estimate origin–destination trips using pervasive traffic data'**, Sustainability (Switzerland), vol. 13, pp. 11171 - 11171, http://dx.doi.org/10.3390/su132011171

Waller, Travis and Qurashi, Moeid and Sotnikova, Anna and Karva, Lavina and Chand, Sai (2023) ‘**Analyzing and modeling network travel patterns during the Ukraine invasion using crowd-sourced pervasive traffic data**’ Transportation Research Record: Journal of the Transportation Research Board, Vol 2677, Issue 10, pp. 491-507, 2023. https://doi.org/10.1177/03611981231161622

Rapid Planning: Introduction of Methodology

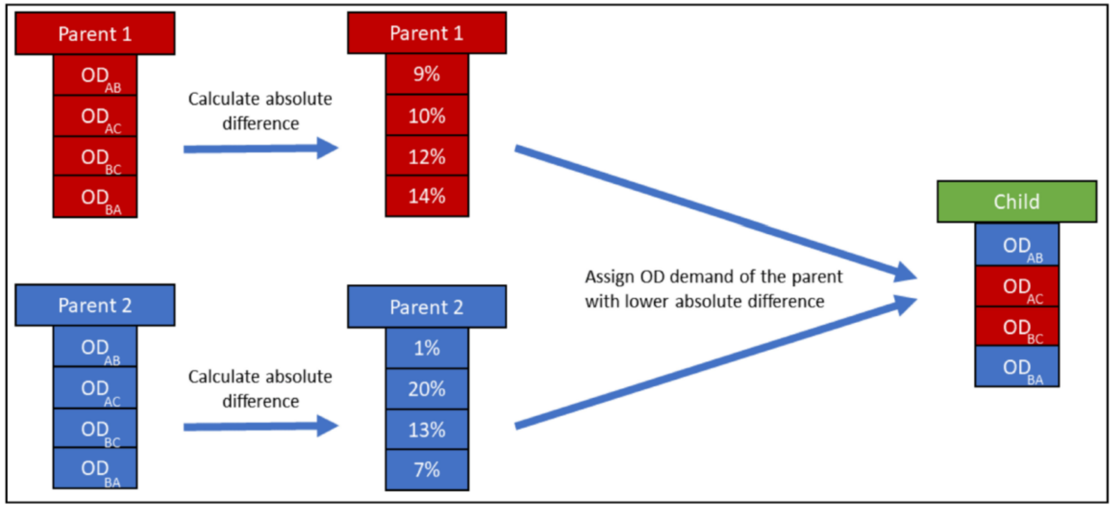
Waller et al. (2021)

Fitness Functions

Acronym	Method Name	Governing Equation	Notation
MAPE-ODTT	Mean absolute percentage error of OD travel times.	$E = \sum_{rs} d_{rs} \cdot \frac{\lvert TT_{rs}^{est} - TT_{rs}^{obs} \rvert}{TT_{rs}^{obs}}$	• E—Error value. • TT_{rs}^{est}—Estimated (from a solution) travel time between OD pair r and s. • TT_{rs}^{obs}—Observed (from any pervasive platform) travel time between OD pair r and s. • N_{OD}— Number of OD pairs. • f_{ij}^{est}—Estimated (from a solution) flow between link i and j. • f_{ij}^{obs}—Observed (from loop detector or other sources) flow between link i and j. • N_f—Number of links in the network where flow values are known. • t_{ij}^{est}—Estimated (from a solution) travel time between link i and j. • t_{ij}^{obs}—Observed (from any pervasive traffic platform) travel time between link i and j. • N_t—Number of links in the network where travel time values are known. • R_i^{est}—Estimated (from a solution) travel time along a user defined route/corridor i. • R_i^{obs}—Observed (from any pervasive platform) travel time along a user defined corridor i. • N_R—Number of user-defined corridors.
RMSE-ODTT	Root mean square error of OD travel times.	$E = \sqrt{\frac{\sum_{rs} \left(TT_{rs}^{est} - TT_{rs}^{obs} \right)^2}{N_{OD}}}$	
MAPE-LF	Mean absolute percentage error of link flows.	$E = \sum_{ij} \frac{\lvert f_{ij}^{est} - f_{ij}^{obs} \rvert}{f_{ij}^{obs}}$	
RMSE-LF	Root mean square error of link flows.	$E = \sqrt{\frac{\sum_{ij} \left(f_{ij}^{est} - f_{ij}^{obs} \right)^2}{N_f}}$	
RMSE-LTT	Root mean square error of link travel times.	$E = \sqrt{\frac{\sum_{ij} \left(t_{ij}^{est} - t_{ij}^{obs} \right)^2}{N_t}}$	
MAPE-LTT	Mean absolute percentage error of link travel time.	$E = \sum_{ij} \frac{\lvert t_{ij}^{est} - t_{ij}^{obs} \rvert}{t_{ij}^{obs}}$	
MAPE-C	Mean absolute percentage error of corridor travel times.	$E = \sum_{i} \frac{\frac{\lvert R_i^{est} - R_i^{obs} \rvert}{R_i^{obs}}}{N_R}$	

Initial Solutions

Acronym	Method Name	Governing Equation	Notation
TFM	Travel time—free flow travel time model.	$d_{rs} = \frac{\frac{TT_{rs}^{obs}}{TT_{rs}^{f}}}{\sum_{rs} \frac{TT_{rs}^{obs}}{TT_{rs}^{f}}} \cdot D$	TT_{rs}^{obs}—Observed (from any pervasive platform) travel time between OD pair r and s. $TT^{f}{}_{rs}$—Observed free-flow travel time between OD pair r and s. k_{rs}—Average shortest distance between the OD pair r and s when the network is empty. G_r—user-defined proportion value of zone r, where $\sum G_r = 1$. A_s—user-defined proportion value of zone s, *where* $\sum A_s = 1$.
FDM	Free flow travel time—distance model.	$d_{rs} = \frac{\frac{TT_{rs}^{f}}{{k_{rs}}^2}}{\sum_{rs} \frac{TT_{rs}^{f}}{{k_{rs}}^2}} \cdot D$	
TDM	Travel time distance model.	$d_{rs} = \frac{\frac{TT_{rs}^{obs}}{{k_{rs}}^2}}{\sum_{rs} \frac{TT_{rs}^{obs}}{{k_{rs}}^2}} \cdot D$	
CGM	Custom gravity model.	$d_{rs} = \frac{\frac{G_r A_s}{{k_{rs}}^2}}{\sum_{rs} \frac{G_r A_s}{{k_{rs}}^2}} \cdot D$	



Rapid Planning: Comparison with Established Models and Surveys

Waller et al. (2021)

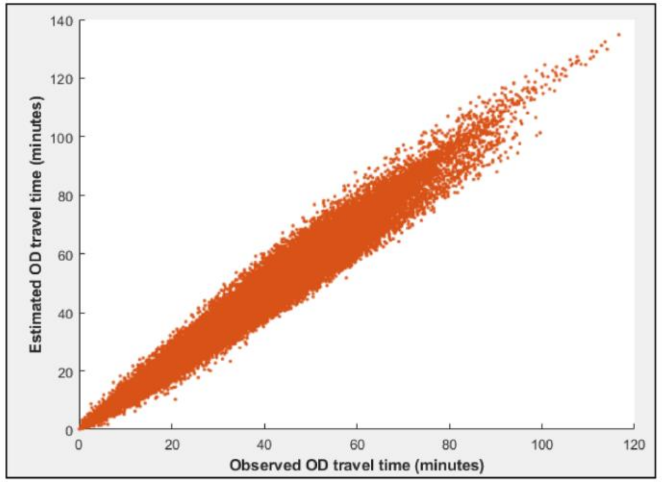
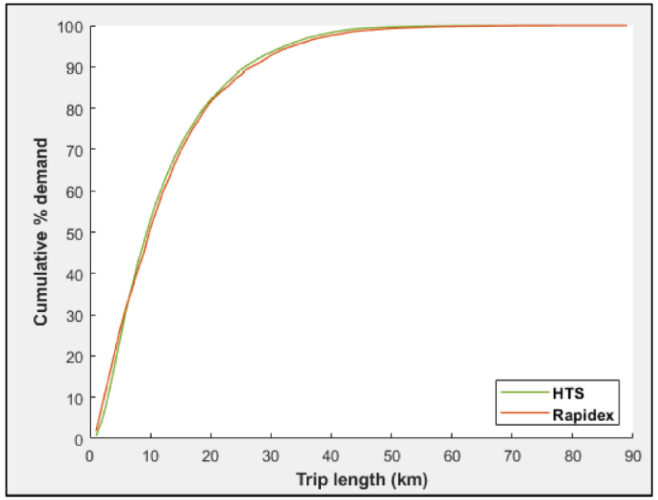


Figure 4. Observed vs. estimated OD travel times.



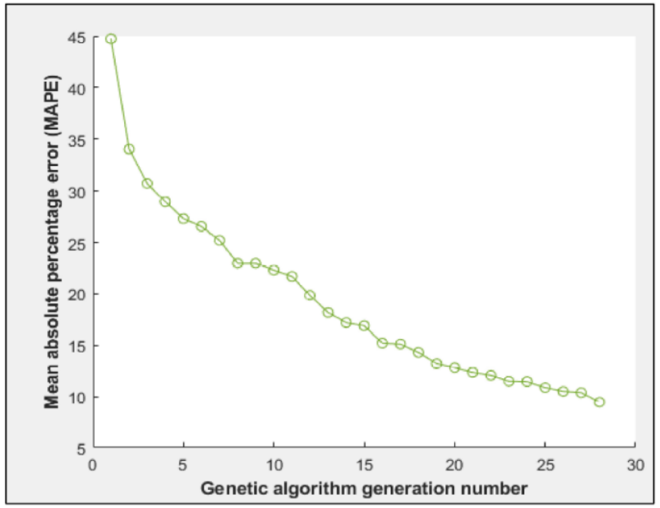
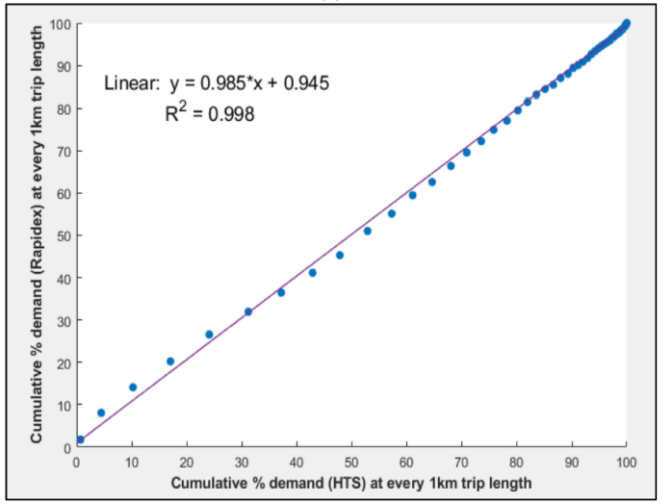


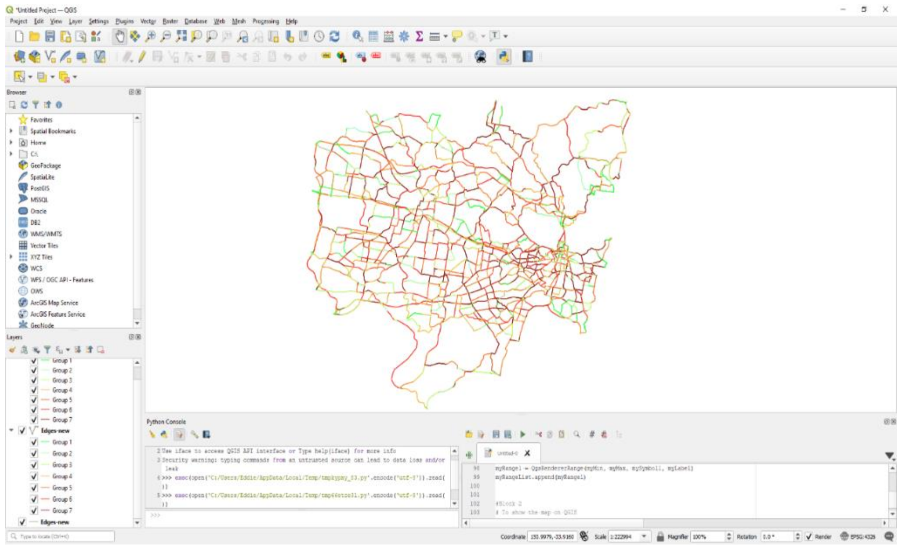
Figure 5. Convergence of the genetic algorithm solution.



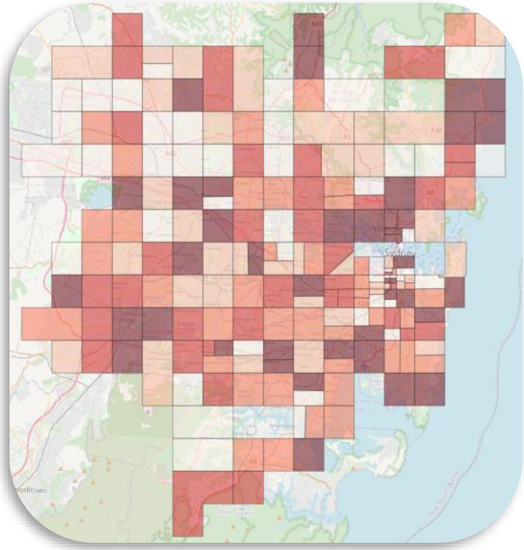
Comparison with

- Observed Data
- Household Travel Survey
- More refined (time-intensive) strategic planning model

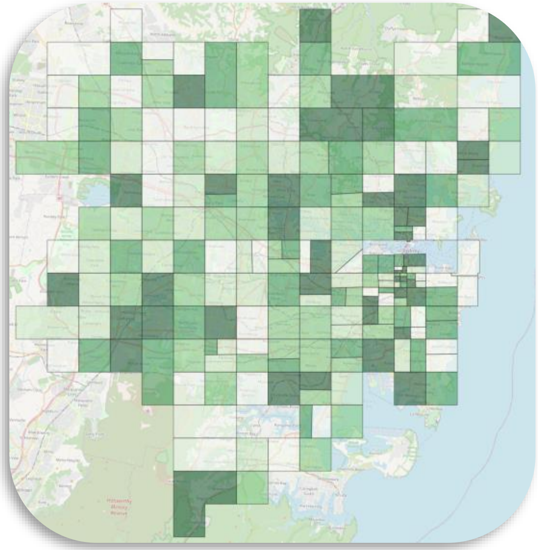
Case Study 1: Sydney Region



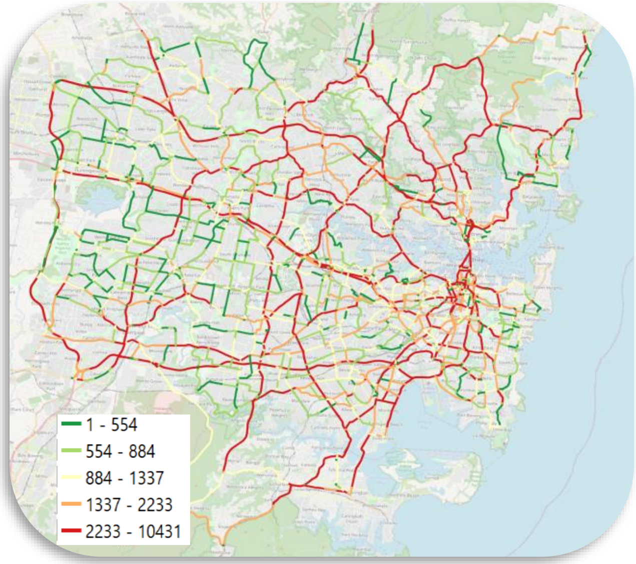
Generations



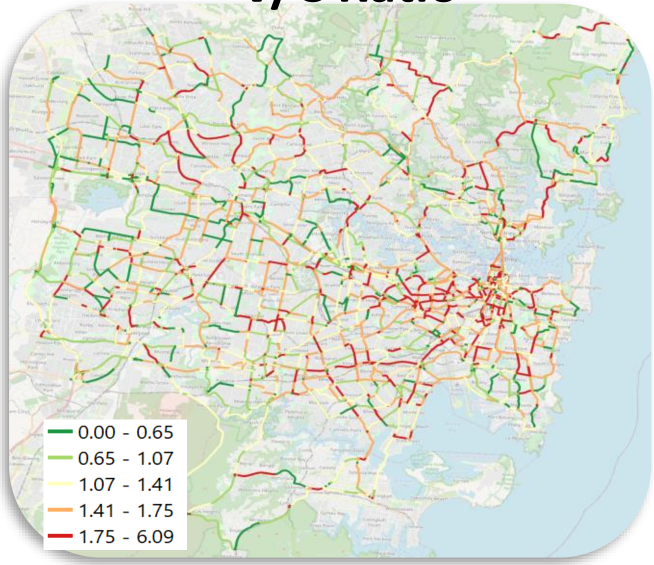
Attractions



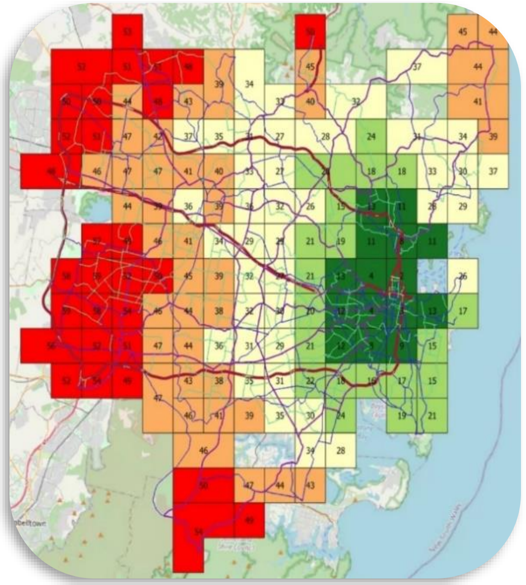
Volume



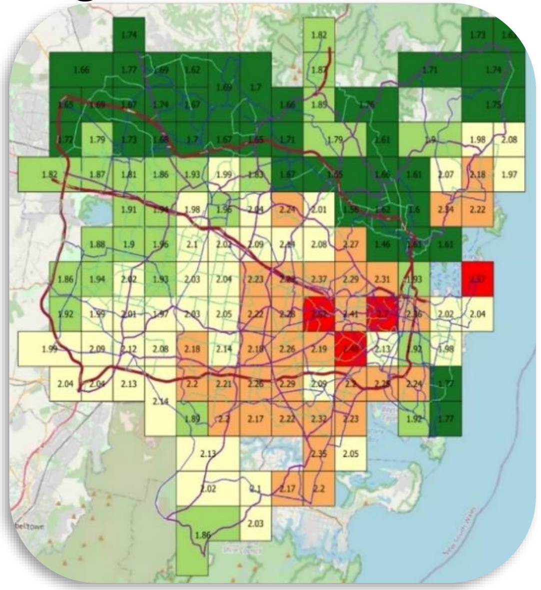
V/C Ratio



Travel time to CBD



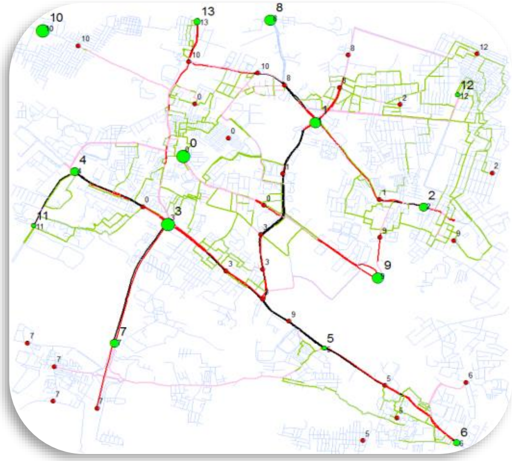
Congestion Index to CBD



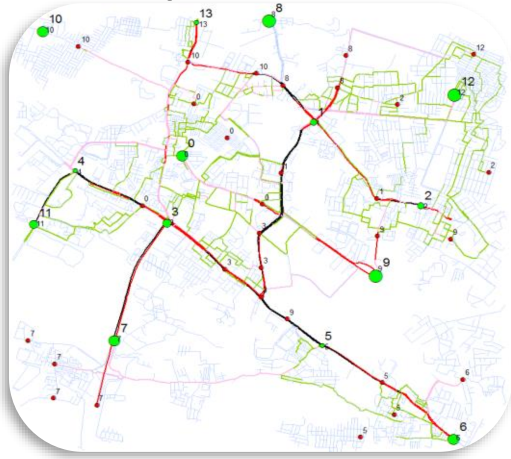
Case Study 2: HiTech City, Hyderabad (India)

Project: needed to establish a model, with no data from agency, to evaluate traffic operational changes related to construction of new metro

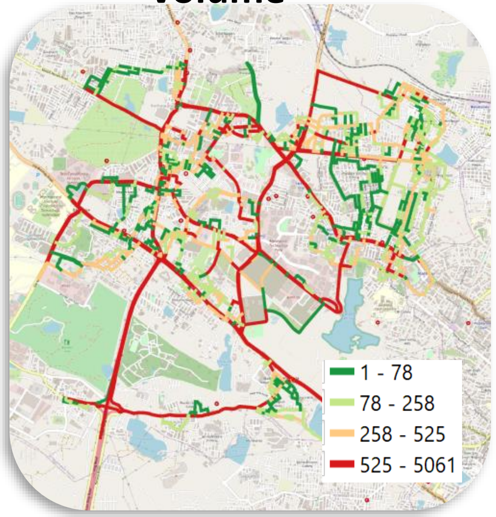
Trip Generations



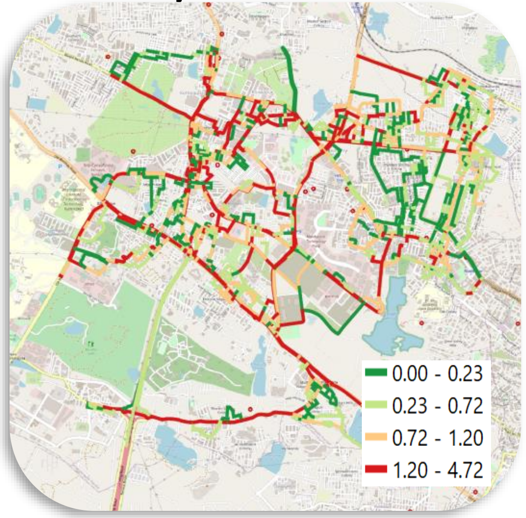
Trip Attractions



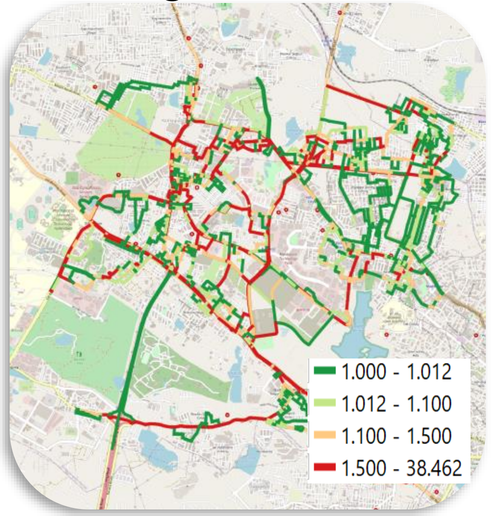
Volume



V/C Ratio



Congestion Index



Models in Ukraine

Waller et al. (2023)

Analysis for 26 February 2022 to 12 April 2022

Focusing on Coefficient of Variance (Std/Mean)

First known paper on travel behavior during human conflict.

Focuses on those who remain in place rather than evacuation/refugee movements.

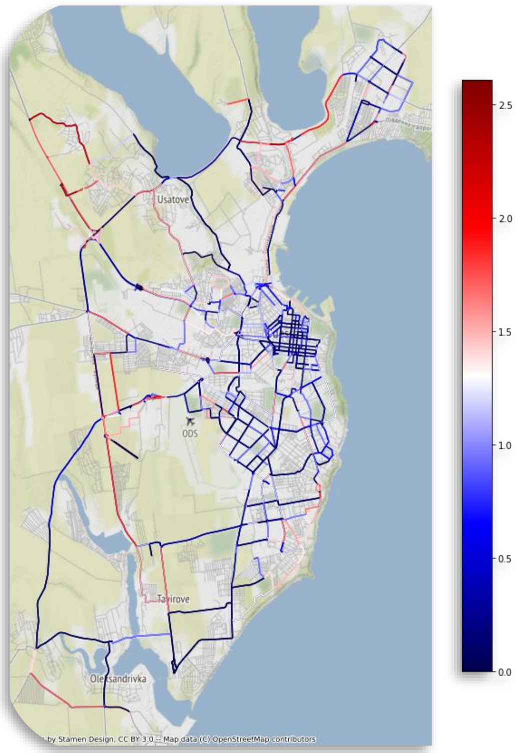
Applications being explored include:

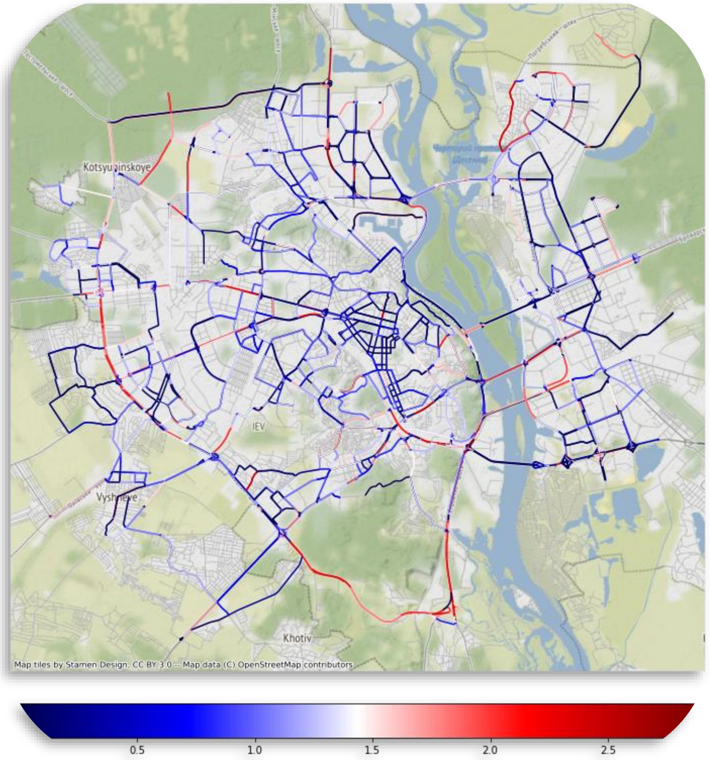
- **Rapid estimation of reconstruction needs**
- **Designing cities that are more resilient to human-conflict**

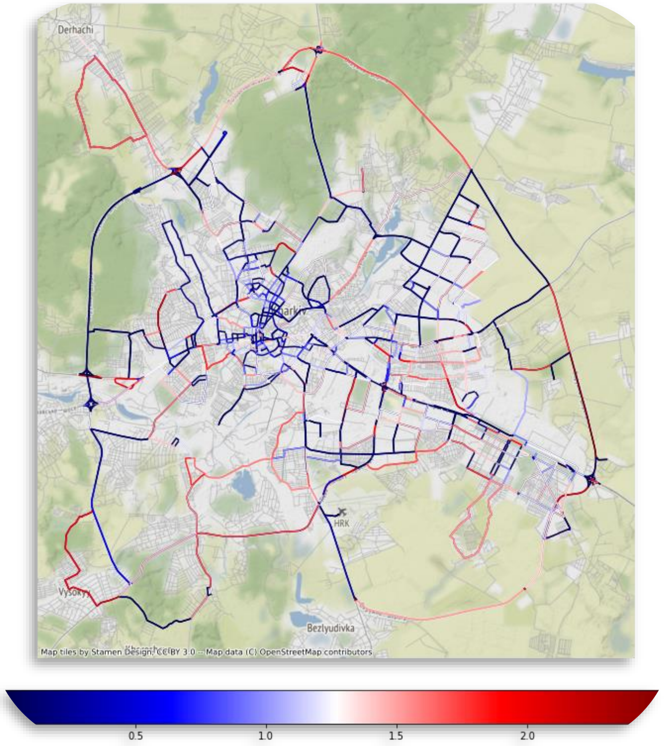
Kyiv
- Links: 4069
- Nodes: 2224

Kharkiv
- Links: 2453
- Nodes: 1017

Odesa
- Links: 1765
- Nodes: 800







Waller, Travis and Qurashi, Moeid and Sotnikova, Anna and Karva, Lavina and Chand, Sai, "**Analyzing and modeling network travel patterns during the Ukraine invasion using crowd-sourced pervasive traffic data**" *Transportation Research Record: Journal of the Transportation Research Board*, Vol 2677, Issue 10, pp. 491-507, 2023.

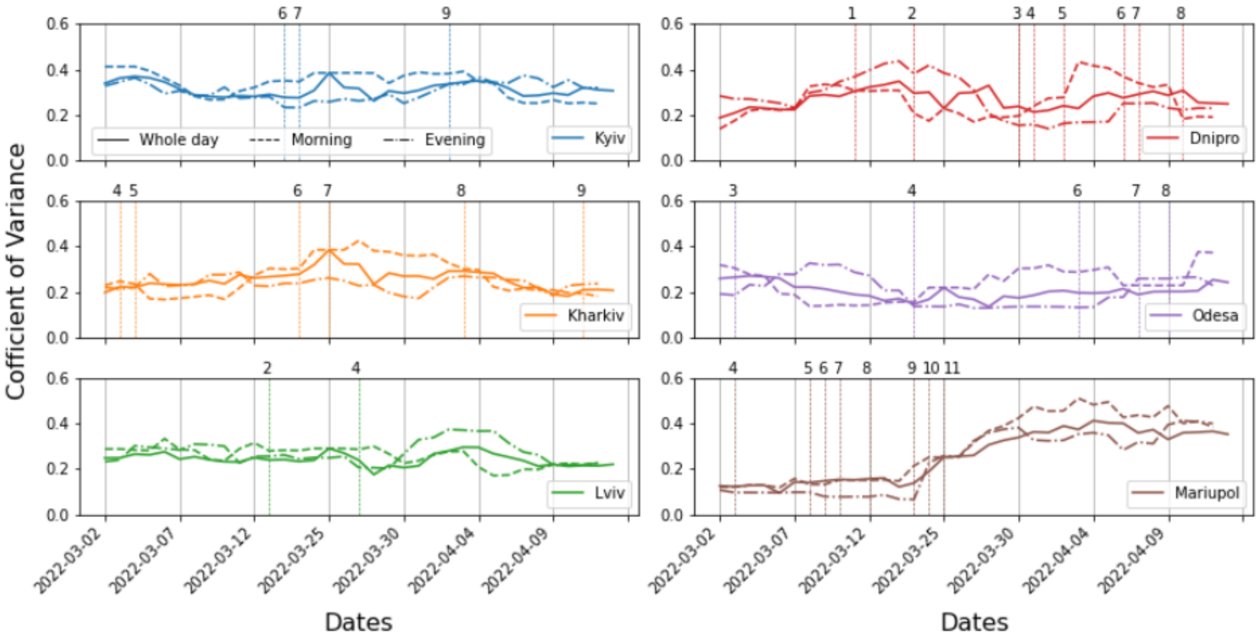
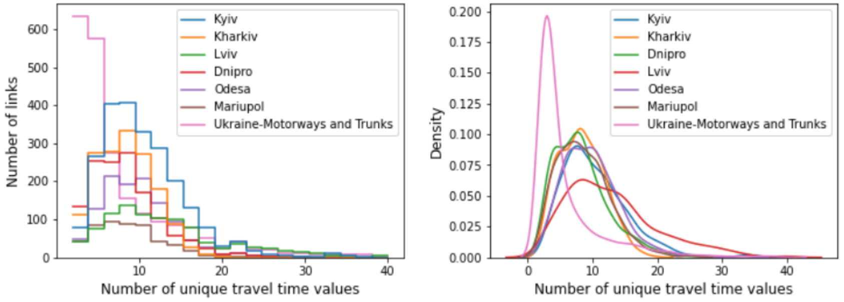


*Figure **3** Network averaged link coefficient of variance for travel times (7-day moving)*

Table 2 Key Statistics from the OD Estimation Analysis

City	***Date***	***% change in average trip length compared to the base case***	***% change in average travel time compared to the base case***	***% change in total demand compared to the base case***
Kyiv	February 28 2022	-	-	-
	March 16 2022	-5.52	-0.28	+3.90
	April 12 2022	+2.74	+1.92	+0.11
Kharkiv	February 28 2022	-	-	-
	March 31 2022	-3.14	+1.55	+6.05
	April 12 2022	+3.40	+11.79	+2.63
Mariupol	February 28 2022	-	-	-
	March 16 2022	+13.11	+28.44	-2.50
	April 12 2022	-6.76	-11.66	+0.58



(a) Histograms of unique link travel times

(b) Probability density functions for unique link travel times

Figure 1 TomTom data reliability plots

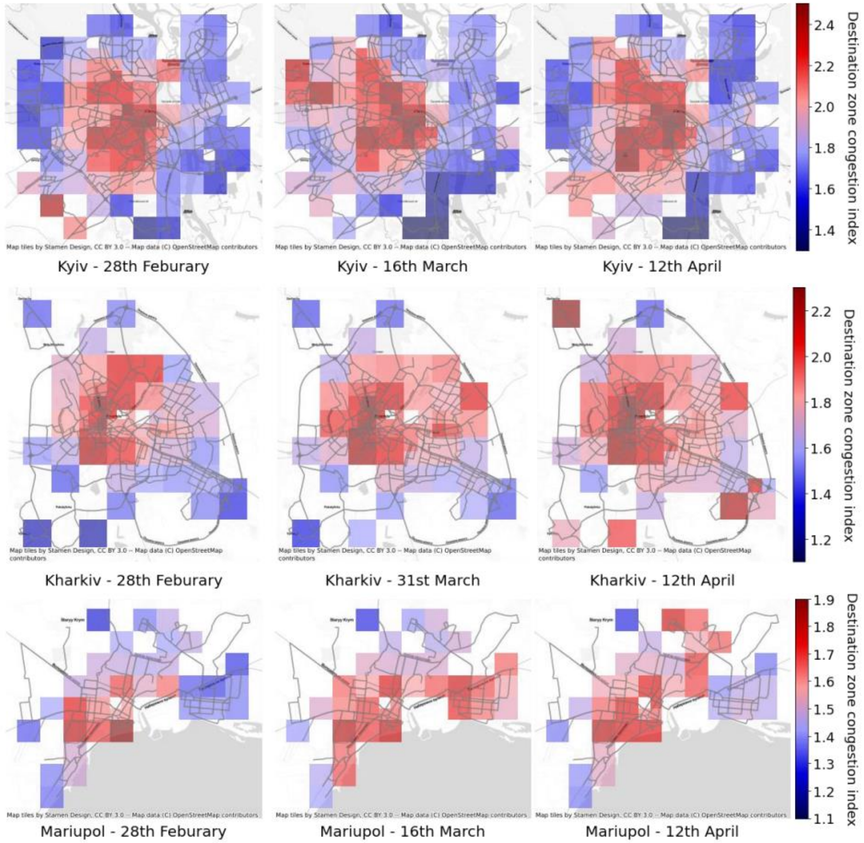


Figure 5 Hue maps of average congestion index for zones as the destination

Rapid Planning Model: Armenia

Links: 3,677

Nodes: 1,962

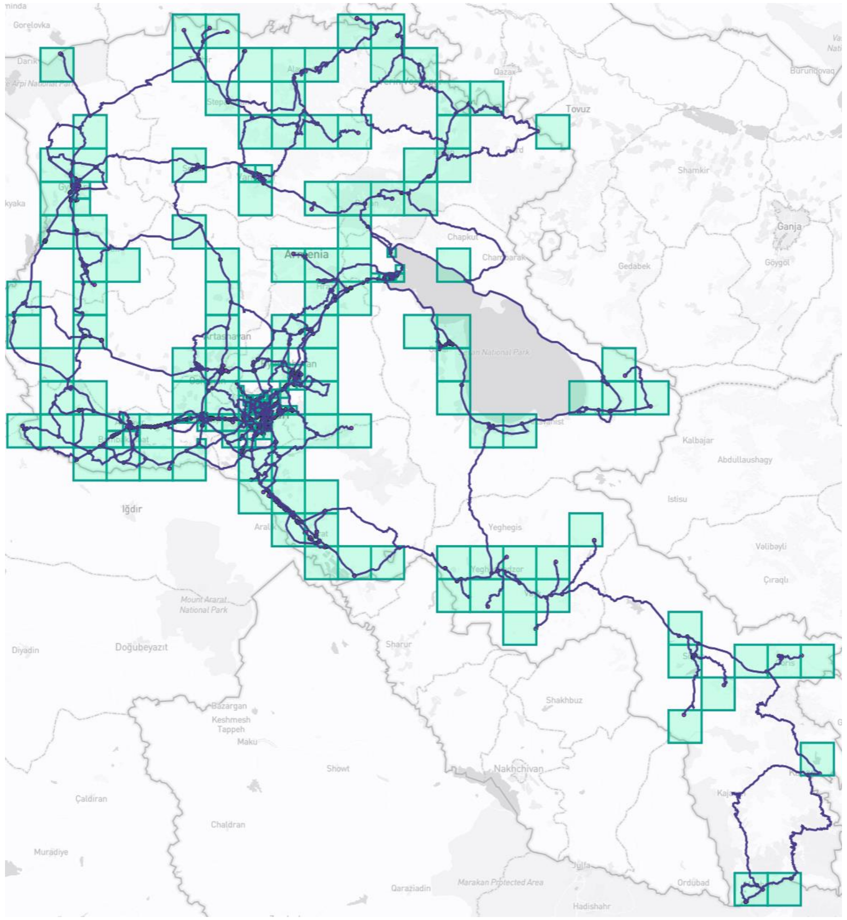
Zones: 175

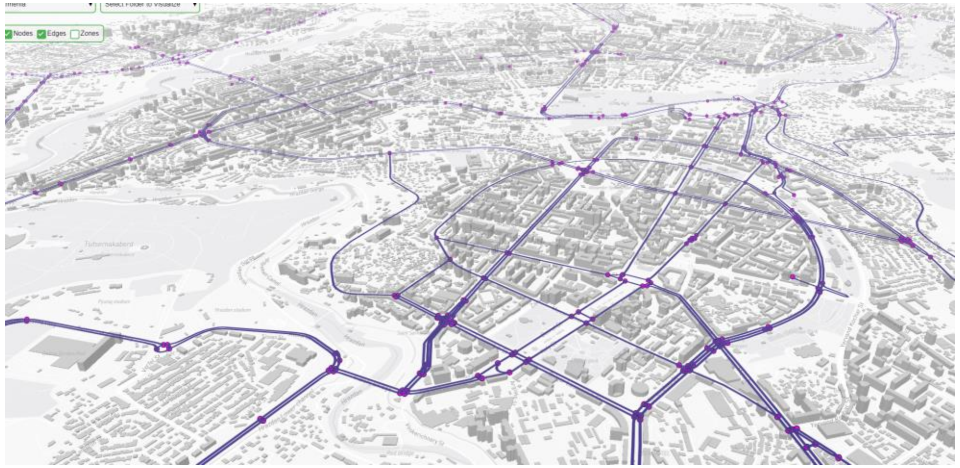
Avg Travel Time: 37 min

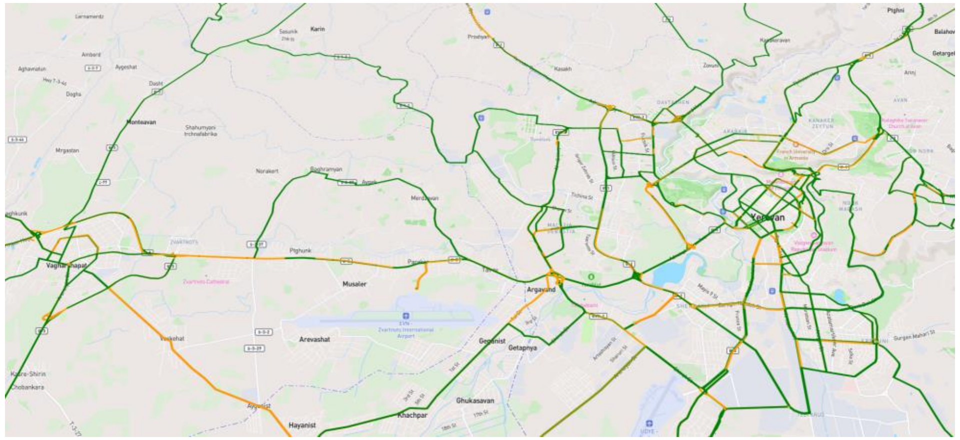
Avg Distance: 30.57 km

Modelled:

- Traffic route assignment
- Volume/Capacity
- Travel Time
- Speed
- Congestion





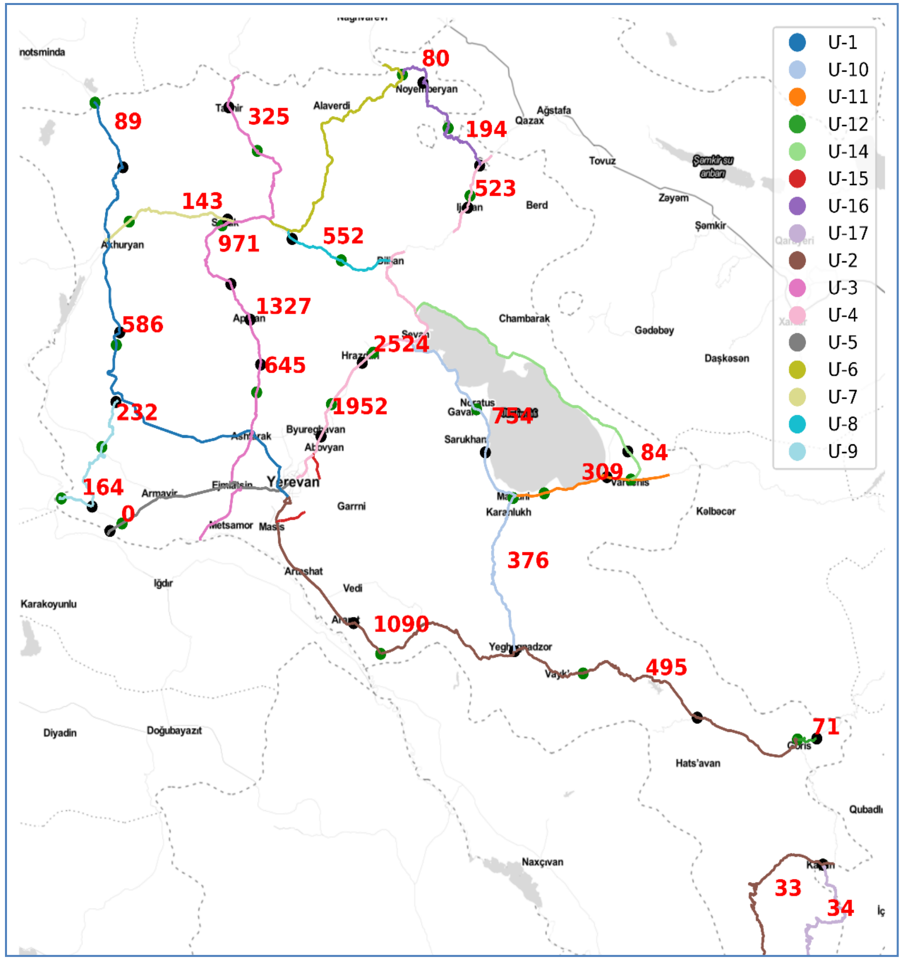


Rapid Planning Model Comparison with Reported Daily Flows

*Reported data is from 2019 unless noted otherwise due to report data omission

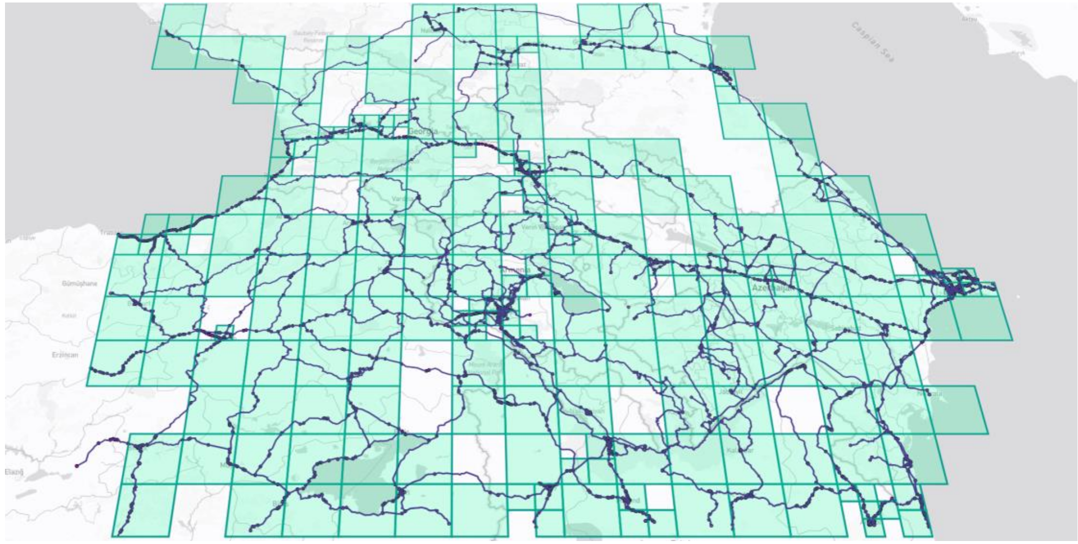
Road type	RPModel Estimated AADT	Reported AADT	RPModel Lengths	Reported Lengths
Interstates	3,612 vpd	3,600 vpd	1,798 km	1,724 km
Republican	1,107 vpd	1,078 vpd	1,452 km	1,968 km

Road No.	Name	Reported AADT 2019 Average (vpd)	Rapid Planning Modelled AADTs Monday (12-12-2022 Snapshot in 9-10am) **Throughput** flow along roadway (AADT vpd)
M-1	Yerevan-Gyumri- Georgia border	24,551	23,484
M-3:	Margara-Vanadzor-Tashir-Georgian border:	6,294	8,226
M-4:	Yerevan-San-Ijan-Adr:	19,512	25,932
M-5:	Yerevan-Armavir-Turkey border:	20,390	22,292
M-8:	Vanadzor-Dilijan	1,415 (2018)	3,423
M-10:	Saint-Martuni-Getap	5,117	5,756



South Caucasus Model

- Coverage including
  - Armenia, Azerbaijan and Georgia
  - with parts of Iran, Turkey, and Russia.
- Two network versions were modelled
  - First network
    - 20,274 links
    - Total length of 39,392 km
    - 221 traffic analysis zones
  - Second streamlined network
    - 6,839 links
    - Total length of 12,542 km
    - 119 traffic analysis zones



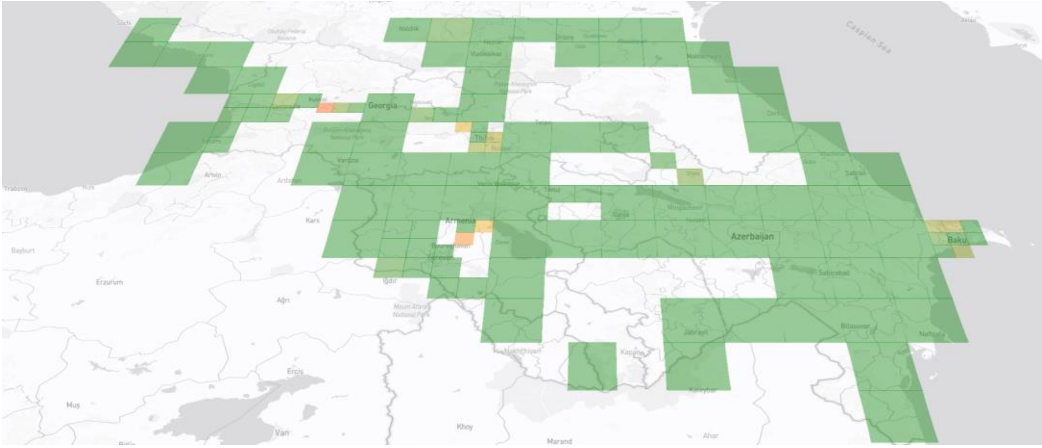
We would like to acknowledge collaboration with NVIDIA

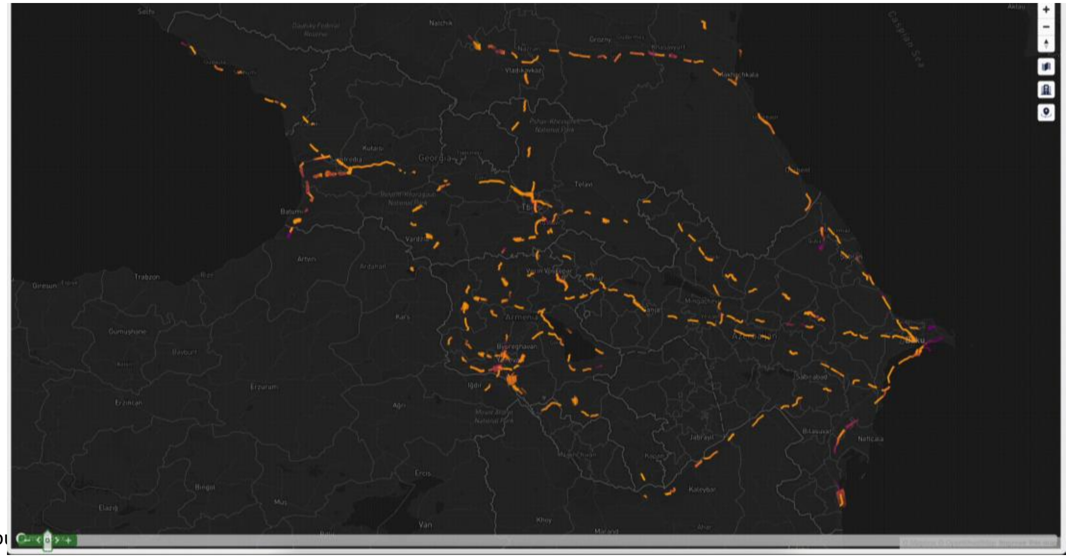
South Caucasus Model

- Base Case
  - 63,357,589 total Vehicle Kilometers Traveled (VKT)

- Comparison
  - Travel times collected on ***all*** links
    - RMSE 16.19 seconds
  - 76 specific link counts were also provided to support direct comparison by the broader team

- Under all borders fully operational “What if” scenario
  - 62,621,005 VKT
  - 736,586 (1.16%) reduction
    - Note, based on scope of work did not include induced future demand





Summary

Through Big Data and Machine Learning

New data sources are increasingly available

- Modelling can be dramatically accelerated
- But, for planning, models must maintain critical properties
  - Analytics is distinct from planning

Always happy for any further questions

- travis@mobilitythinking.com