NEW APPLICATIONS OF AI TO TRAVEL DEMAND MODELING

Caliper®

Transportation & Mapping Solutions

Maptitude • TransCAD • TransModeler

THE STATISTICAL STATUS QUO

- Travel Models are Statistical Models
  - Travel models are built out of classical statistical models like regression and logit models
  - Travel models use goodness-of-fit statistics like r2 and log likelihood, so they must be statistical
  - Travel models have been around for longer than machine learning / AI – hence, they couldn't use machine learning

THE STATISTICAL STATUS QUO

- Travel Models are Statistical Models
  - Travel models are built out of classical statistical models like regression and logit models
  - Travel models use goodness-of-fit statistics like r2 and log likelihood, so they must be statistical
  - Travel models have been around for longer than machine learning / AI – hence, they couldn't use machine learning
- But…
  - Travel models began as computer algorithms – that later found statistical theory – like a lot of machine learning / artificial intelligence…

WHAT'S THE DIFFERENCE?

- Classical Statistics vs. Machine Learning / AI
  - More Culture than Math
  - Understanding vs. Predicting
  - Single Dataset vs. Multiple Datasets
  - Low Dimensionality vs. High Dimensionality
  - Parametric vs. Nonparametric

WHY AI?

- For Travel Forecasting

WHY AI?

- For Travel Forecasting
  - Prediction, Prediction, Prediction
    - For planning & engineering, not a pure science

WHY AI?

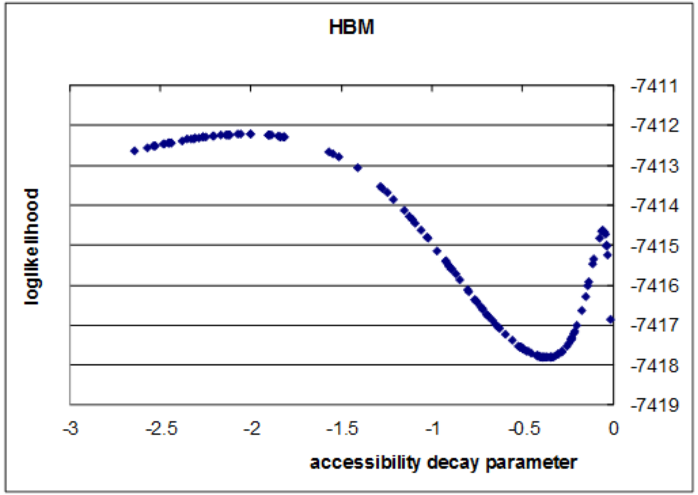
- For Travel Forecasting
  - Prediction, Prediction, Prediction
    - For planning & engineering, not a pure science
  - Multiple Datasets
    - Surveys AND Counts AND Passive Data

WHY AI?

- For Travel Forecasting
  - Prediction, Prediction, Prediction
    - For planning & engineering, not a pure science
  - Multiple Datasets
    - Surveys AND Counts AND Passive Data
  - High Dimensionality
    - Location x Location x Location

WHY AI?

- For Travel Forecasting
  - Prediction, Prediction, Prediction
    - For planning & engineering, not a pure science
  - Multiple Datasets
    - Surveys AND Counts AND Big Data
  - High Dimensionality
    - Location x Location x Location
  - Nonparametric
    - Because of Constraints and Equilibrium
    - Reality of Multiple Optima



KEY INSIGHTS OF ARTIFICIAL INTELLIGENCE

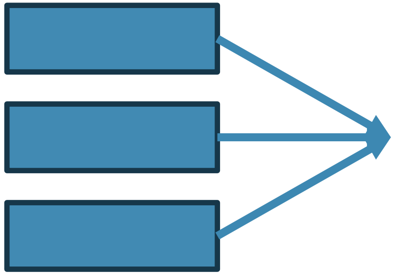
Caliper®

KEY INSIGHTS OF AI

- Ensemble Modeling
- Nonparametric Methods
- Regularization
- Semi-Supervised Learning
- Overparameterization

ENSEMBLE MODELING

- Two models are better than one!
- Models can be combined in parallel or in sequence

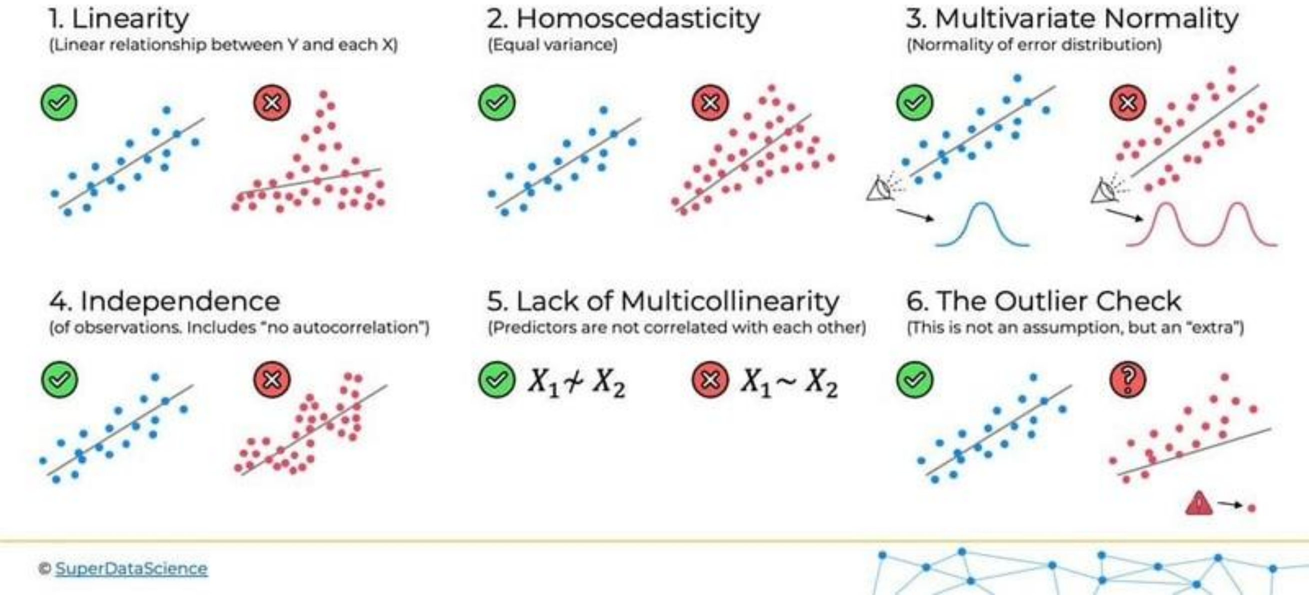


- All models are wrong, but different models are wrong in different ways

- Combining multiple models can use offsetting errors to compensate for the weaknesses of one model with the strengths of another and vice versa

AVOID PARAMETRIC ASSUMPTIONS

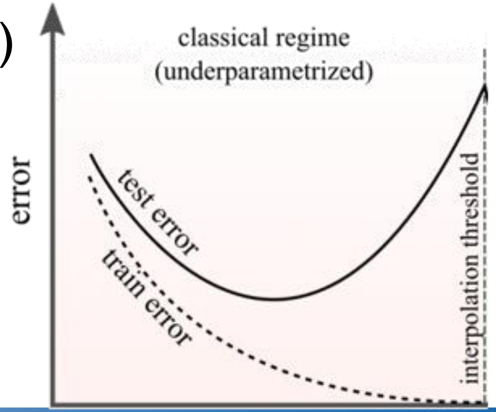
- Turns out those "benign" assumptions often matter



- Specification Error! Omitted Variable Bias

PREDICTION, NOT REPLICATION

- Out-of-Sample or External Validity
  - Does the model generalize?
- Travel Modeling Desperately Needs to Take this In
  - Goal is to predict the future, not replicate the present (base year)
- "Regularization" Methods (How to Avoid Over-fitting)
  - Holdout samples (split data into training and testing sets)
  - k-fold cross-validation
  - Loss (error) function penalties (lasso, ridge, elastic net)
  - Dropout, etc.
  - Early stopping (training, tuning, testing)

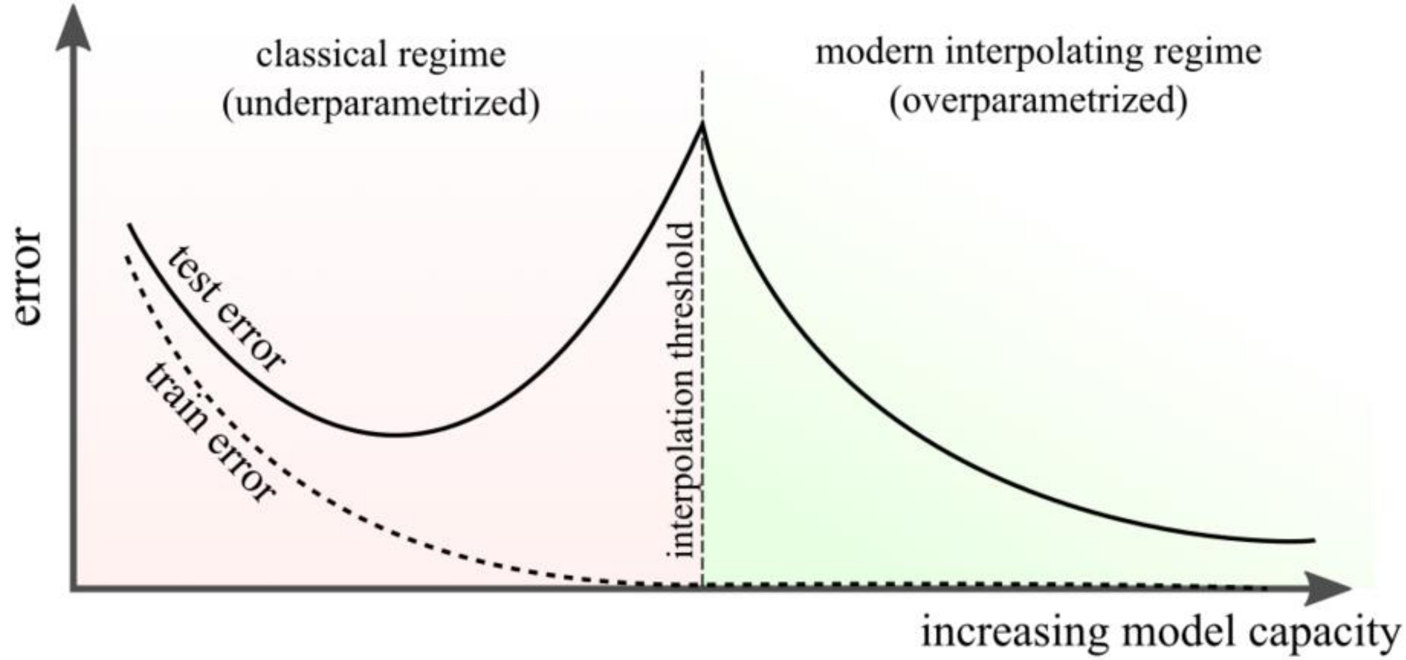


SEMI-SUPERVISED LEARNING

- Reducing specification error by learning not just the parameters of the model (coefficients of the formula) but the structure of the model (the terms and structure of the formula)
- One key aspect of this is dimensionality reduction
  - Principal Component Analysis (PCA) from traditional statistics
  - Discriminant Analysis (LDA, GDA)
  - Filter and Wrapper Methods (correlation, variance filters, RFE, backwards FE)
  - Embeddings (word2vec, POI2vec, etc.)

OVERPARAMETERIZATION

- "Impossible" Double Descent



APPLICATIONS TO TRIP GENERATION

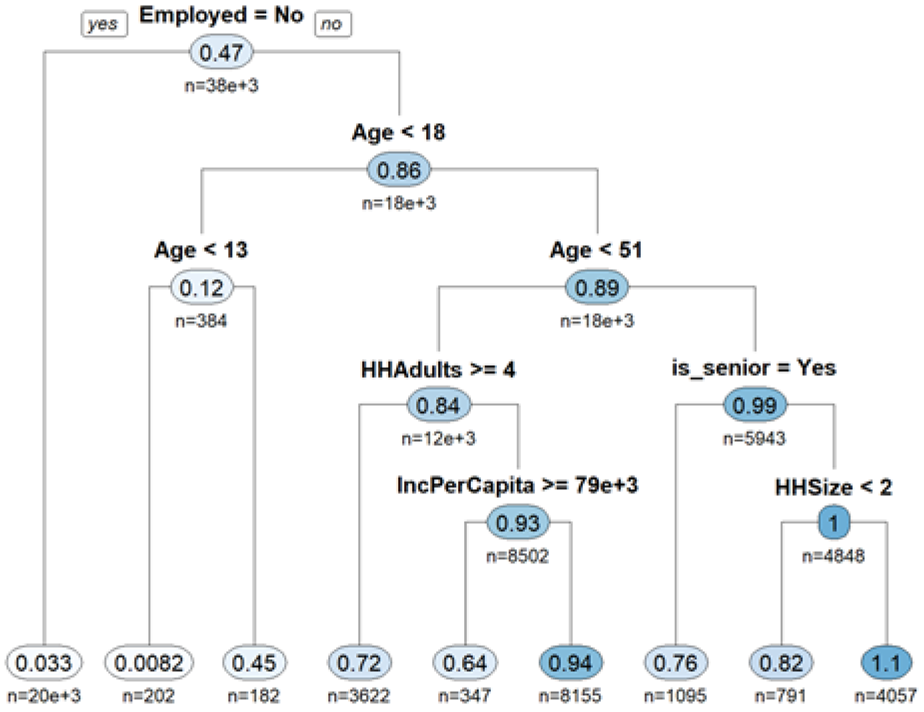
Caliper

TRIP GENERATION USING DECISION TREES

- First tested and implemented for the NC Research Triangle model in 2021
- Since then, implemented for a few MPOs around the country
  - Las Vegas, NV
  - Reno, NV
  - Wichita, KS
- Now, more advanced hybrid regression / multi-class decision tree models being developed for NC statewide model
- Colby Brown has also experimented with using ChatGPT to generate activity patterns

TRIP GENERATION BY DECISION TREES

- The game of 20 Questions
- Advantages of Decision Trees
  - Sensitivity
    - **Age**
    - **Neighborhood / Accessibility**
    - Income
    - Vehicle ownership
    - Household composition
  - Nonlinear effects
  - Full survey support
    - No empty cells like with cross-class



COMPARISON WITH TRADITIONAL MODELS

- Tested classical stats & plain AI methods
  - Cross-classification
  - GLM (up to and including zero-inflated negative binomial)
  - Logit (ordered logit)
  - Extreme Gradient Boosted Decision Trees (XGBoost)

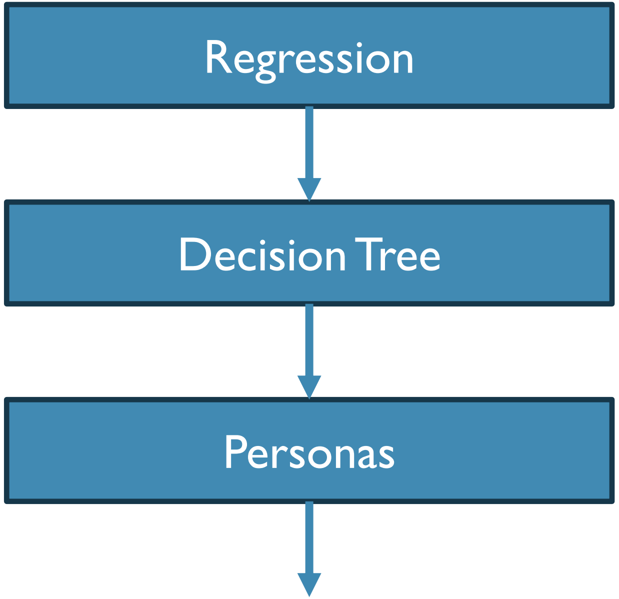
Example: School Trips

Model Type	Pseudo R^2
Logit	0.03
GLM (Regression)	0.22
Cross-Class	0.33
XGBoost	0.60
XAI ANOVA Decision Tree	**0.53**

- Chosen approach: **Explainable Artificial Intelligence (XAI)**
  - ANOVA-based Rationalized Decision Trees
  - **Explainable**, reasonable relationships between trip rates and explanatory variables
  - Confidence that the model is not over-fit to the data

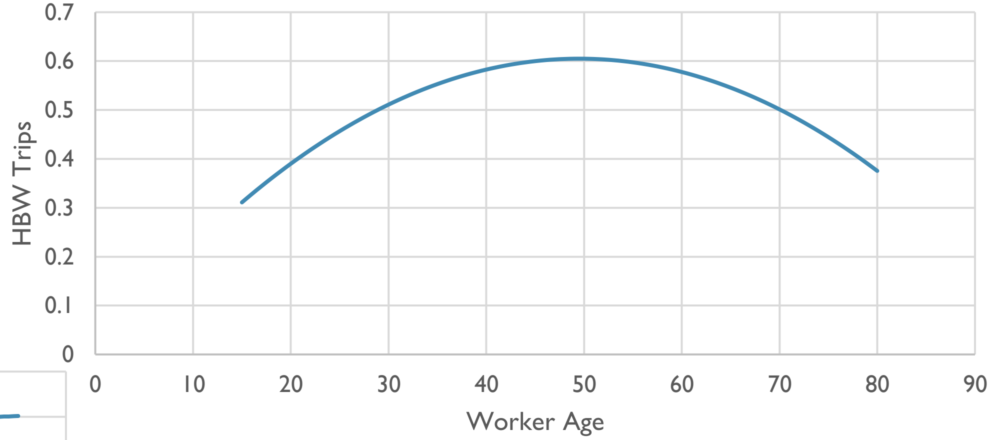
PBHRDC MODELS

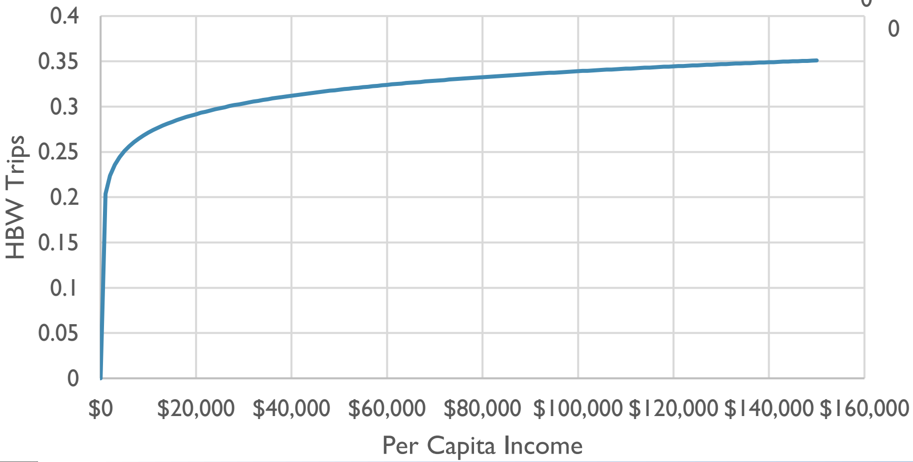
- Currently testing new ensemble methods for NC statewide model
- Form of doubly-boosted model
- Regression on continuous variables expected to affect everyone (income, age, accessibility)
- Decision tree on first residual using categorical variables (gender, employment, marital status, etc.)
- Asserted tree on "personas"



REGRESSION

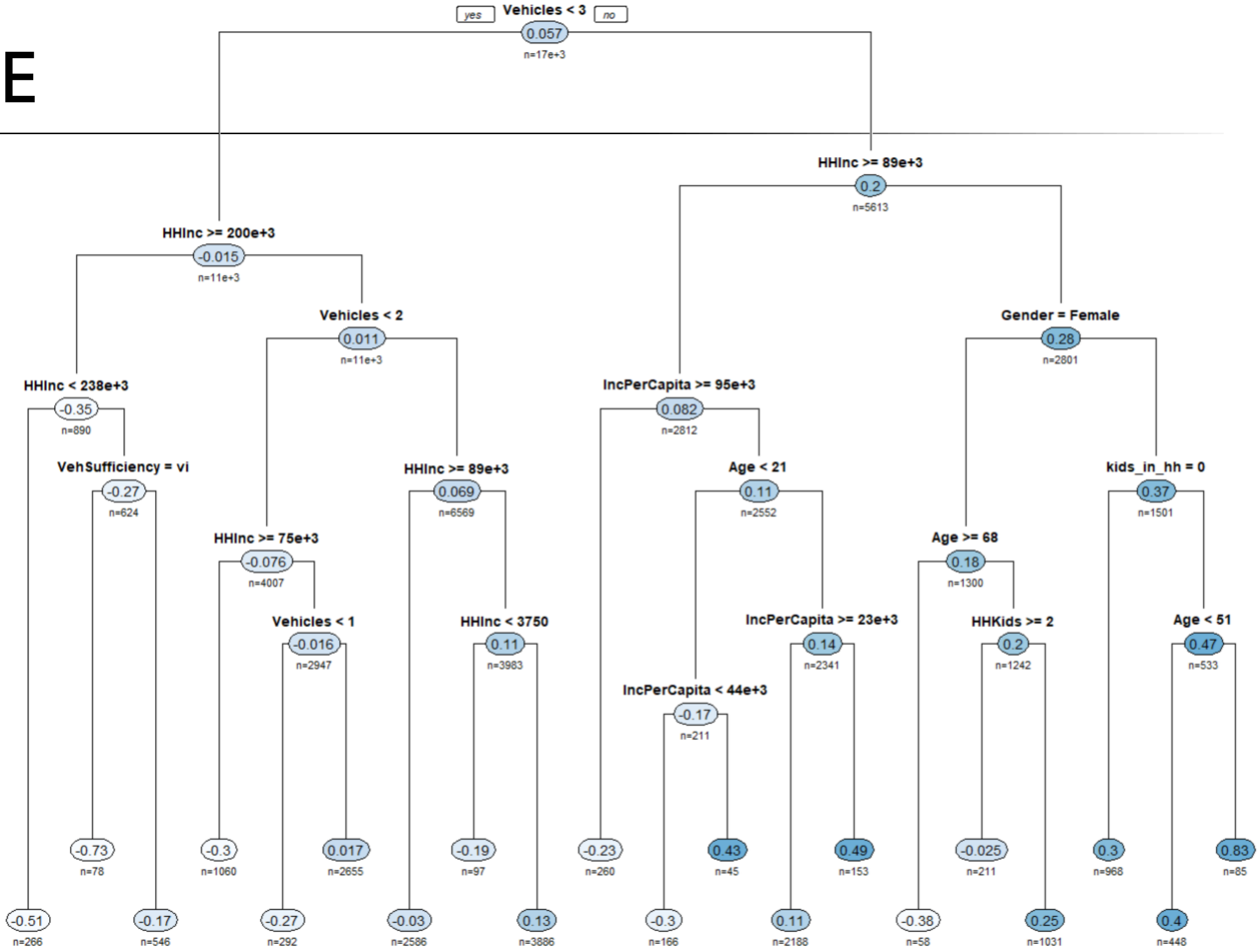
- Age
- Income





DECISION TREE

- ANOVA-based Decision Tree



PERSONA CONSTANTS

- Constant (average second residual) for each "persona"

Multigenerational Family	Senior
	Dad
	Mom
	Adult Child
	Child
Senior(s) & Child(ren)	Senior
	Child
TradFam	Dad
	Mom
	Adult Child
	Child
Working Parents	Dad
	Mom
	Adult Child
	Child
Single Parent	Parent
	Child

Adult Parent(s) Adult Child(ren)	Adult Parent
	Adult Child
Senior Parent(s) Adult Child(ren)	Senior Parent
	Adult Child
Senior Couple	Senior
DINK	Worker
SINK	Worker
	NonWorker
NINK	NonWorker
Other Families	Worker
	Non-Worker
Sinlges	Worker
	Non-Worker
Roommates	Worker
	Non-Worker

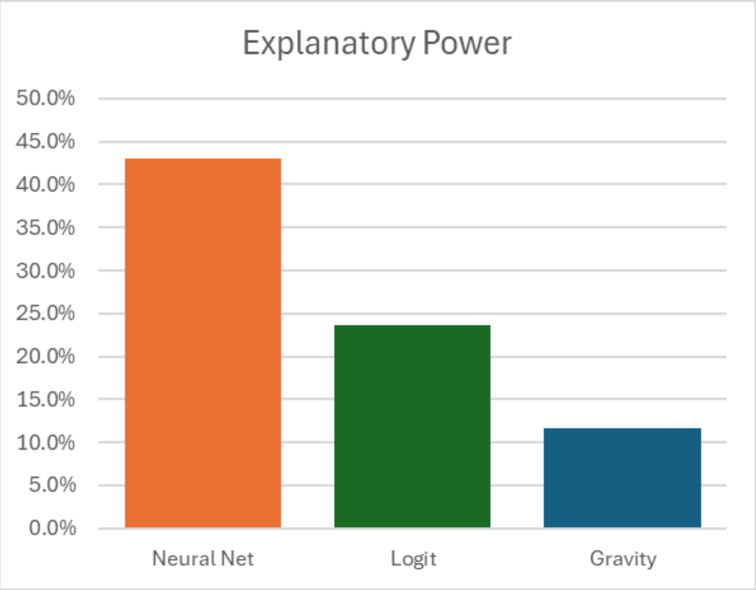
FHWA TMIP PROJECT

OVERVIEW OF PROJECT

- Project to improve travel forecasting through the use of big data and AI
  - Review of literature and practice
  - Testing new methods
  - Implementation pilot projects with case studies
  - "Playbook" for incorporating AI in travel models
  - TMIP webinars to promote Playbook methods
- Current Status: Finalizing Task 2 Report: *Report on Methods and Applications of AI and Big Data to Enhance Travel Forecasting*

PROJECT FOCUS

- Focus on AI
  – References to TMIP resources on big data
- Focus on Practical Improvements for the Near- to Mid-Term
  – Methods to improve/replace individual model components
  – AI-DCMs
  – Primary focus on Destination Choice
    - Largest source of error in existing models – largest opportunity for improvement



CALIPER TEAM LEADERSHIP

Vince Bernardin, PhD
Project Manager

Howard Slavin, PhD
Senior Advisor

Wuping Xin, PhD
Deputy Project Manager

Francisco Pereira, PhD
Senior Advisor

EXPERT PANEL

Kara Kockelman, PhD

Leta Huntsinger, PhD

Mark Bradley

Joshua Auld, PhD

Brian Gregor, PE

Sabya Mishra, PhD

Dan Work, PhD

ARTIFICIAL INTELLIGENCE – DISCRETE CHOICE MODELS

Caliper®

AI-DCM MODELS

- Artificial Intelligence – Discrete Choice Models
- Combine neural networks and logit models
- Attempt to combine the best of both traditional and newer methods
  - Theoretical basis and interpretability of traditional models
  - Explanatory power and accuracy of AI
- Six types proposed so far
  - L-MNL
  - ResLogit
  - TB-ResNet
  - TasteNet
  - RUMnets
  - e-Logit

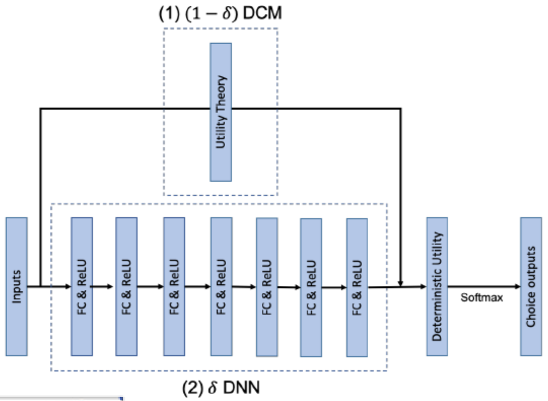
AI-DCM RECOMMENDATIONS

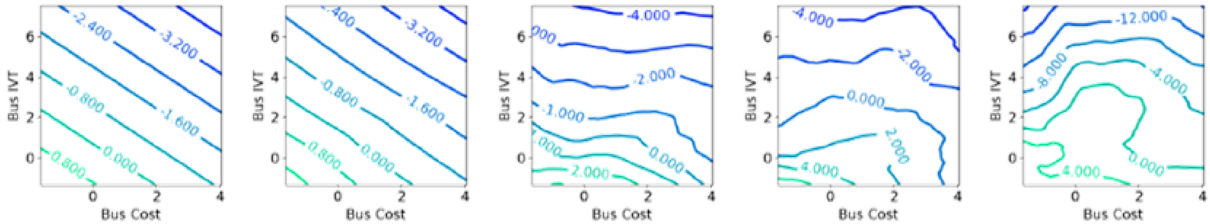
- Goals for destination choice
  - Allow for bounded (imperfect) rationality while avoiding highly irrational behavior
  - Capture cross-effects between alternative destinations
- Recommending TB-ResNets and L-MNL be tested for destination choice
- TasteNet may also offer some improvement for mode choice

TB-RESNETS

- Ensemble of Logit and Deep NN
- Interpretable as a logit or DNN
- Utilities weighted average of logit and DNN
- Weight estimable from data

Fig. 1. Architecture of TB-ResNet. Both DCM and DNN are flexible





(a) MNL (50.6%)
(b) MNL-ResNet ($\delta = 10^{-5}$; 53.1%)
(c) MNL-ResNet ($\delta = 0.008$; 57.0%)
(d) MNL-ResNet ($\delta = 0.05$; 56.1%)
(e) DNN (55.8%)

Fig. 2. Utility functions of MNL-ResNets, MNL, and DNNs. Upper row: visualization of 2D utility functions, and percentages in the parentheses represent the prediction accuracy. Lower row:

L-MNL

- Uses NN to model alternative specific constant (average residual error) similar to boosting
- Or, decomposes systematic utility into traditional theoretical part and data driven part

$$V_{in} = f_i(\mathcal{X}_n; \beta) + r_i(\mathcal{Q}_n, w)$$

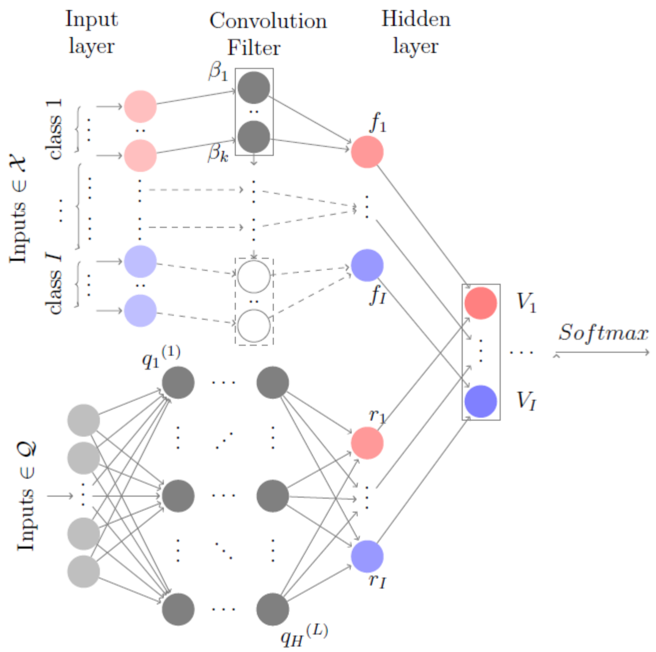


Figure 2: L-MNL model architecture. On the top, we have the I class generalization of a linear-in-parameter MNL model, as depicted in Figure 1. At the bottom, we have a deep neural network (*i.e.*, multilayer and fully connected) that enables us to obtain the representation learning term r_i. The terms from each part are added together defining the new systematic function of Equation (11).

NEURAL NETWORKS FOR DESTINATION CHOICE

REVIEW

- Identified 326 papers from 1993 to present
- Explosion of papers from 2016, peaking in 2020, stabilized around 2018-19 levels
- Needed to prioritize, mostly based on citation rates
- Cursory review of 108 papers and 18 surveys/reviews
- Currently report summarizes of 25 papers
  - Plus brief overview of 15 early papers
  - And appendix with 13 paper summaries
- Conducted formal meta-analysis of 107 published models

CLASSIFYING PAPERS

- **Year of publication**
- **Academic lineage / literature cited** – seven major branches of literature were identified based on literature cited
- **Type/subject domain of journal/conference** – papers were categorized based on whether the journal / conference they appeared in was focused on
  - transportation,
  - geography/GIS,
  - data science, or
  - something else
- **Application (Data Type)**
  - All travel (GPS/LBS trace data, travel surveys)
  - Commuting (Surveys/administrative records on commuting)
  - Transit ODs (Smartcard data)
  - Taxi/TNC/Ride-hailing (Taxi/TNC data)
  - Social Point-of-Interest (Location-based social networking (LBSN) data)
- **Problem formulation**
  - Direct demand
  - Singly constrained
  - Doubly constrained
- **Methodology** – over a dozen neural network methods/architectures were identified

HISTORY

- Papers varied significantly over time and across communities of researchers
- Before 2015, 60% published in transportation or geography journals
- Since 2015, over 80% published in data science journals (6% transp. & geog.)
- Commercial applications for TNCs and location-based marketing
- Development of deep learning

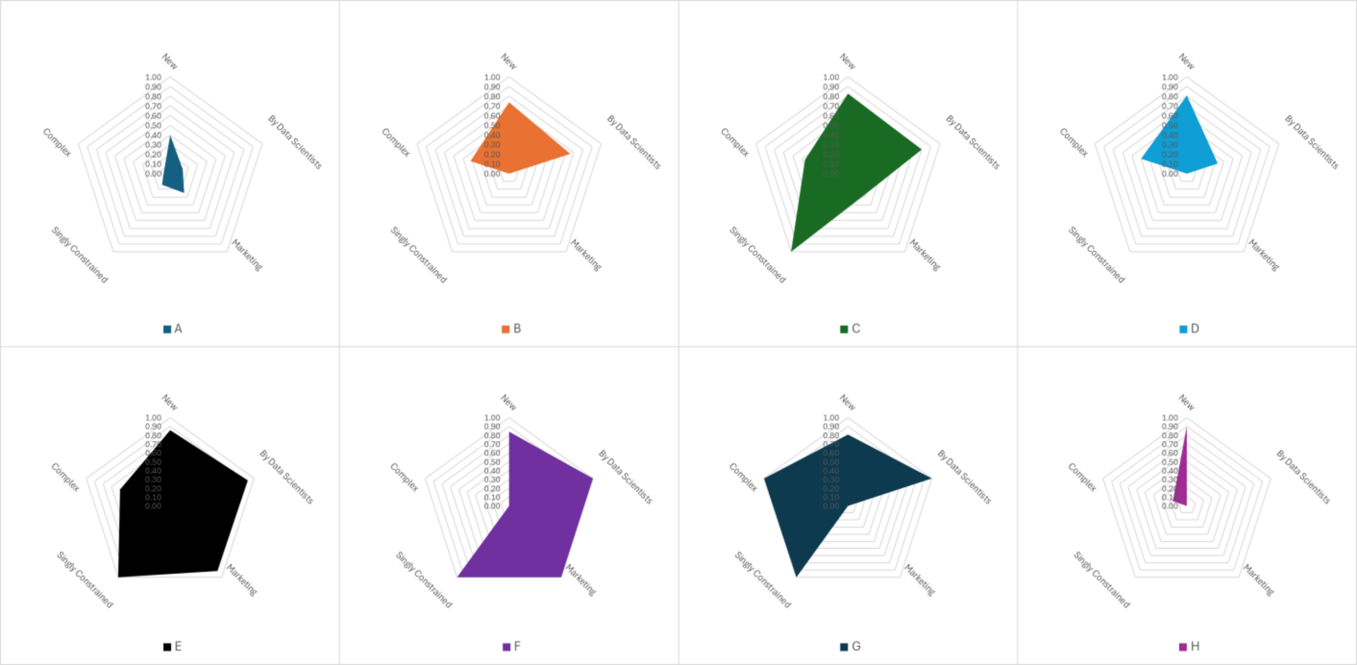
	Transport	GIS / Geography	Data Science	Other	Total
1993		1			1
1994		1			1
1995	1				1
1996		1			1
1997			1		1
1998		1		2	3
1999		1			1
2000	1			3	4
2001				1	1
2002					
2003		1			1
2004	2				2
2005					
2006				1	1
2007			1		1
2008					
2009	2				2
2010	1				1
2011	1				1
2012			1		1
2013					
2014	1				1
2015	1		1		2
2016	2		6		8
2017			14		14
2018	1	1	28	3	33
2019	3	3	37	1	44
2020	9	2	53	2	66
2021	5	1	25	2	33
2022	2	3	29	3	37
2023	2	2	35	4	43
2024	2		19		21
	36	18	250	22	326

PROBLEM FORMULATION

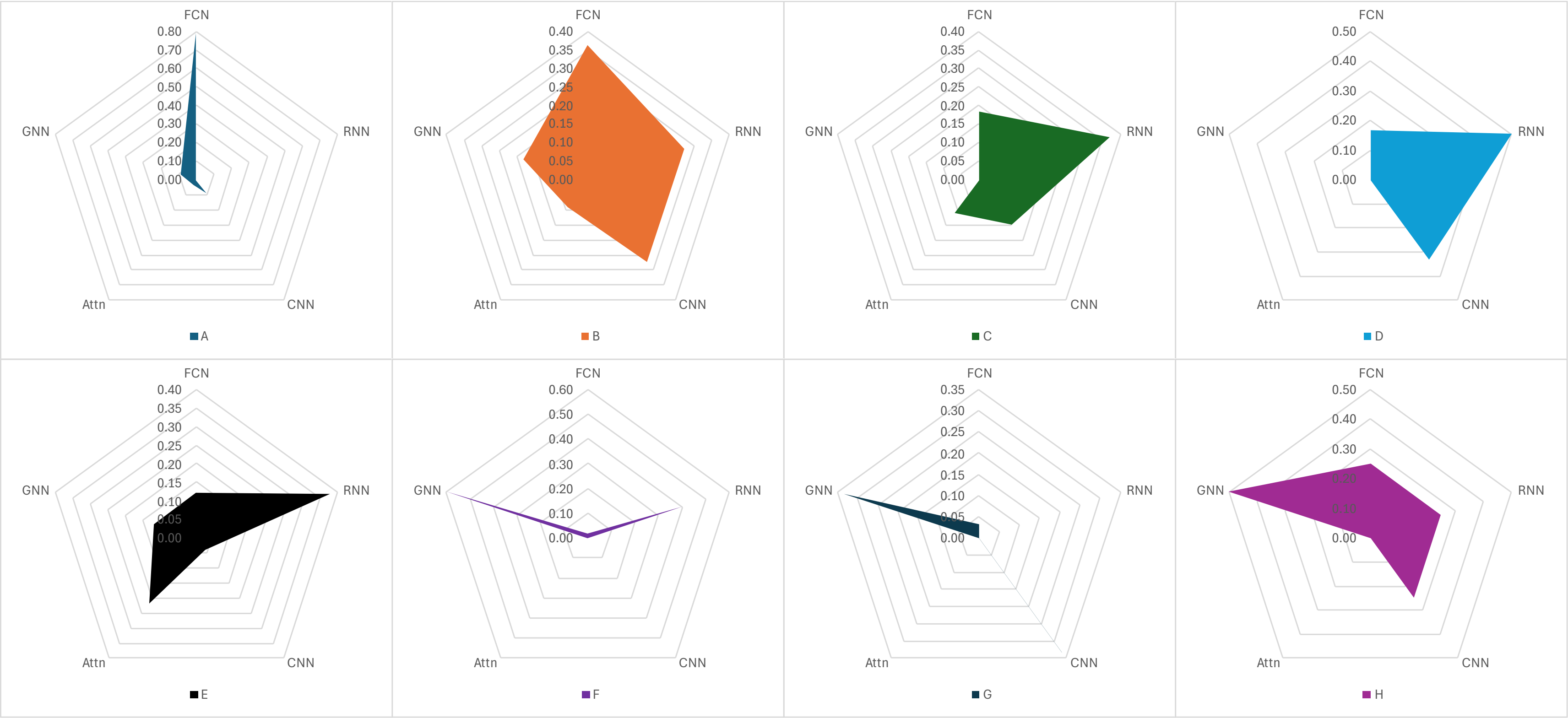
- Degrees of Freedom / Constraints
  - Direct Demand / Unconstrained models try to predict both level and distribution of OD demand
  - Singly-Constrained models try to predict distribution of OD demand given constraint to one marginal (number of trips generated/produced)
  - Doubly-Constrained models try to predict distribution of OD demand given constraint to both marginals
- Extremely uneven coverage in the literature
- Unclear if comparisons are fair

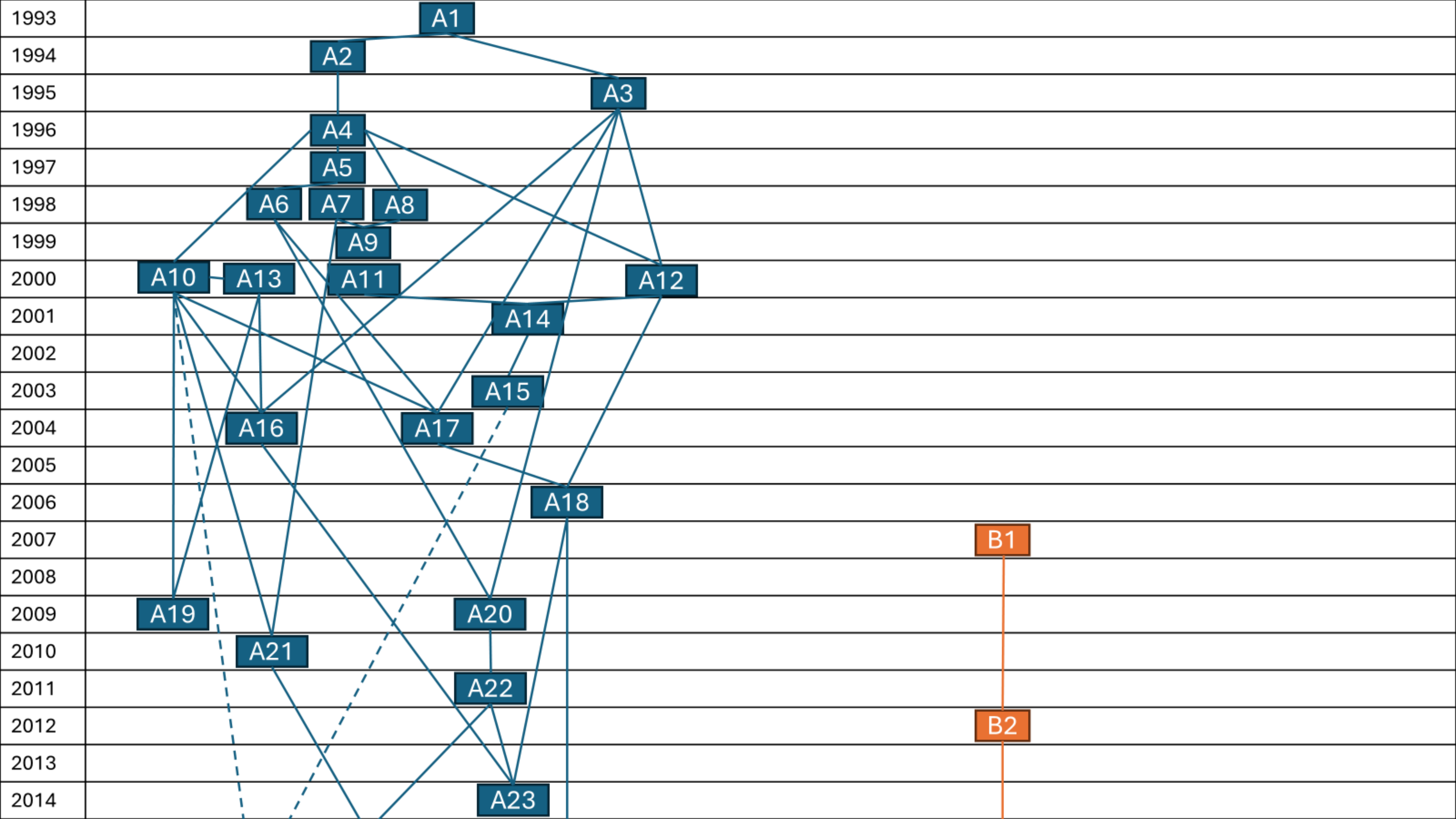
BRANCHES OF THE LITERATURE

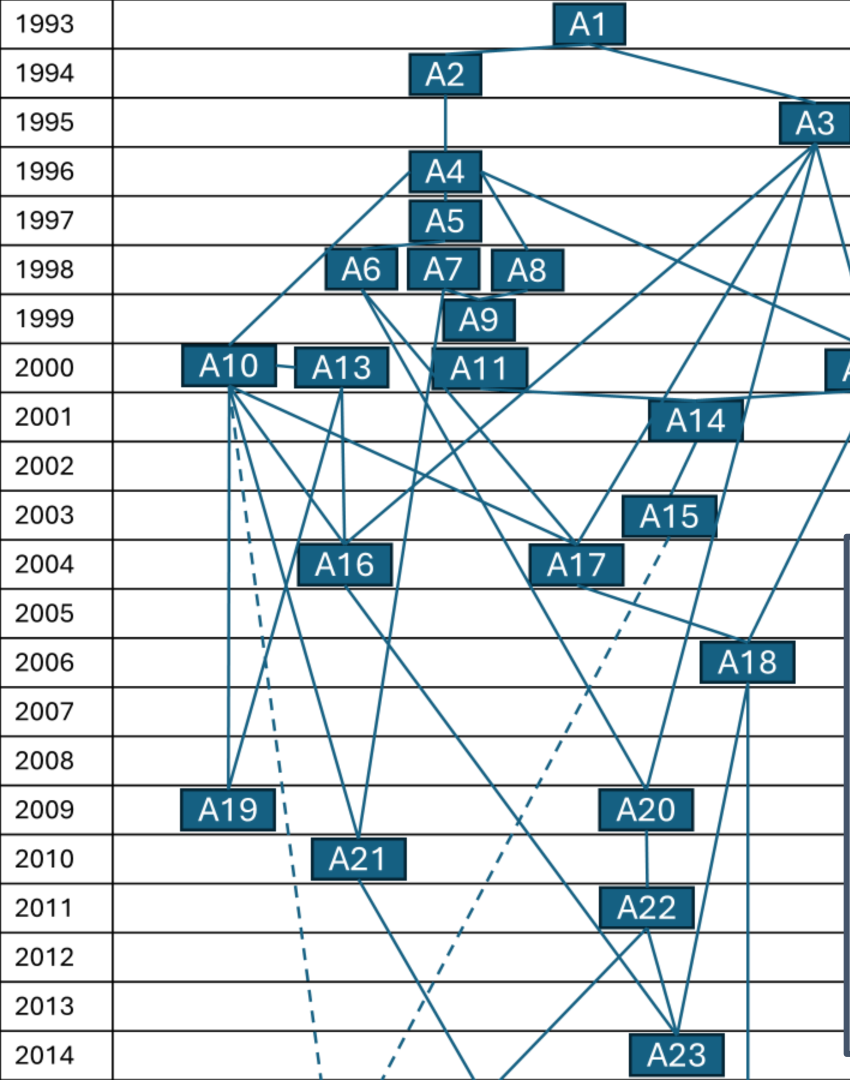
- Eight branches of the literature
  - Based on citations, but vary across many dimensions



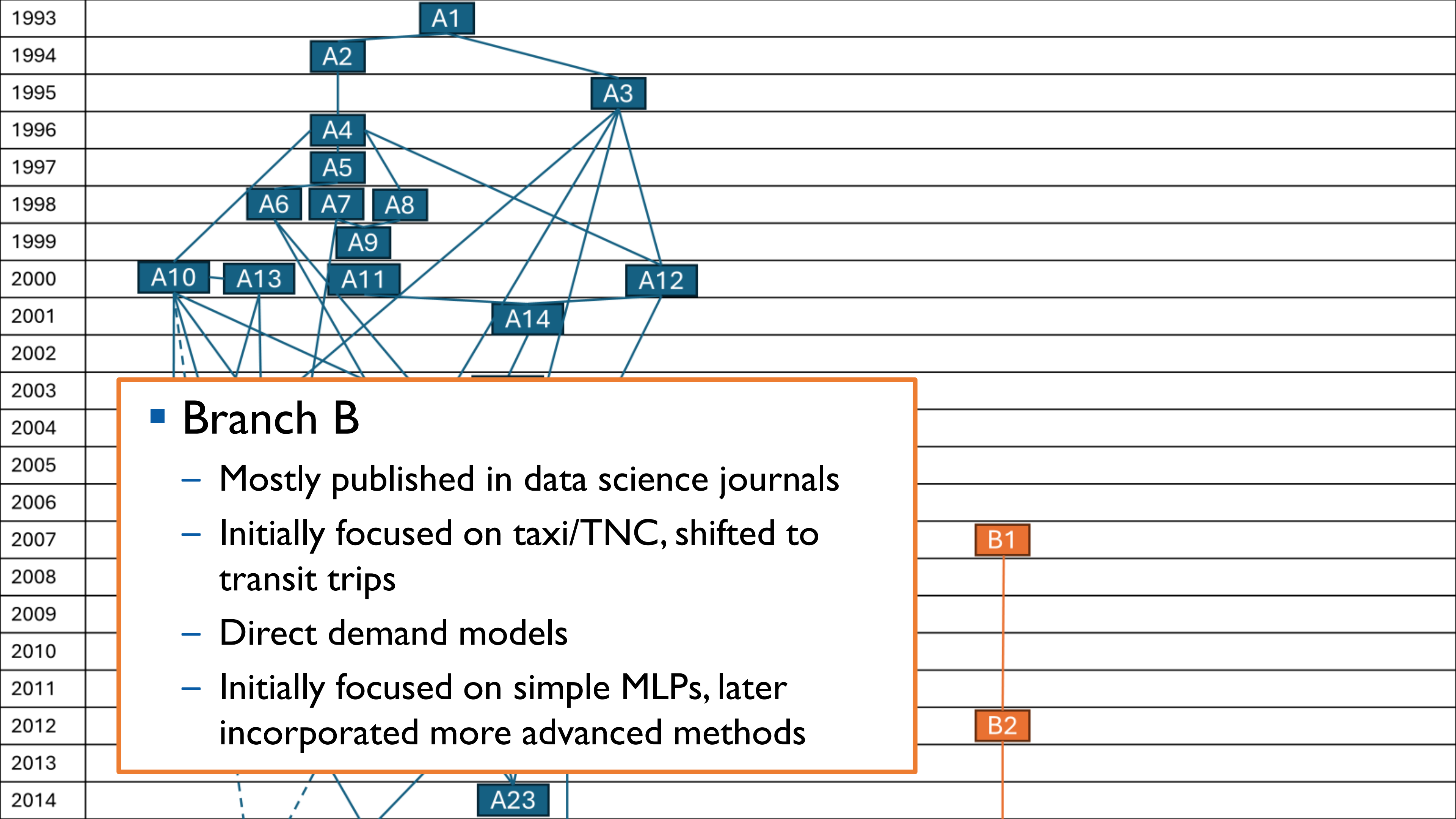
BRANCHES METHODOLOGICAL FOCUS



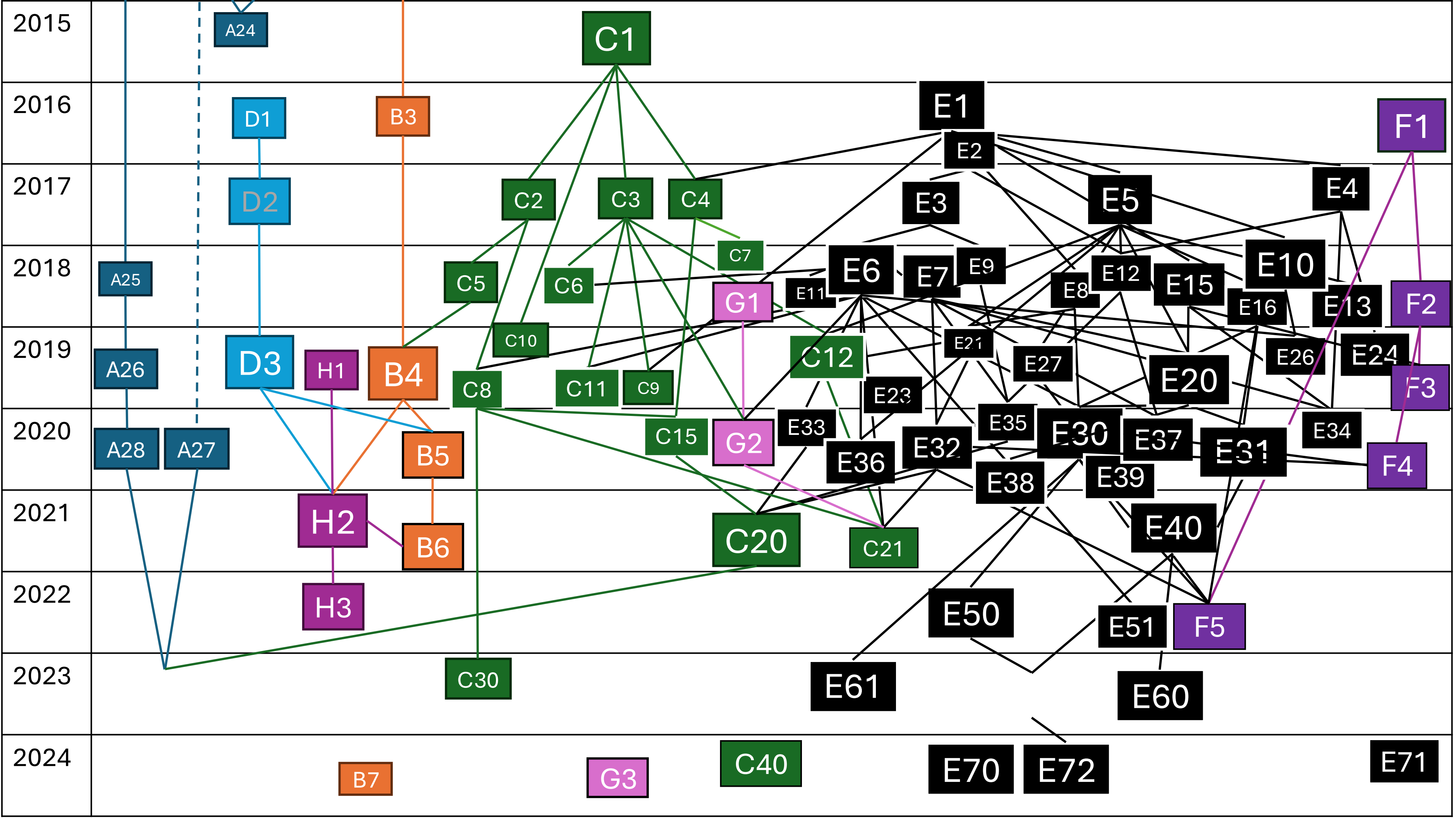

1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
A1
A2
A3
A4
A5
A6
A7
A8
A9
A10
A11
A12
A13
A14
A15
A16
A17
A18
A19
A20
A21
A22
A23
B1
B2



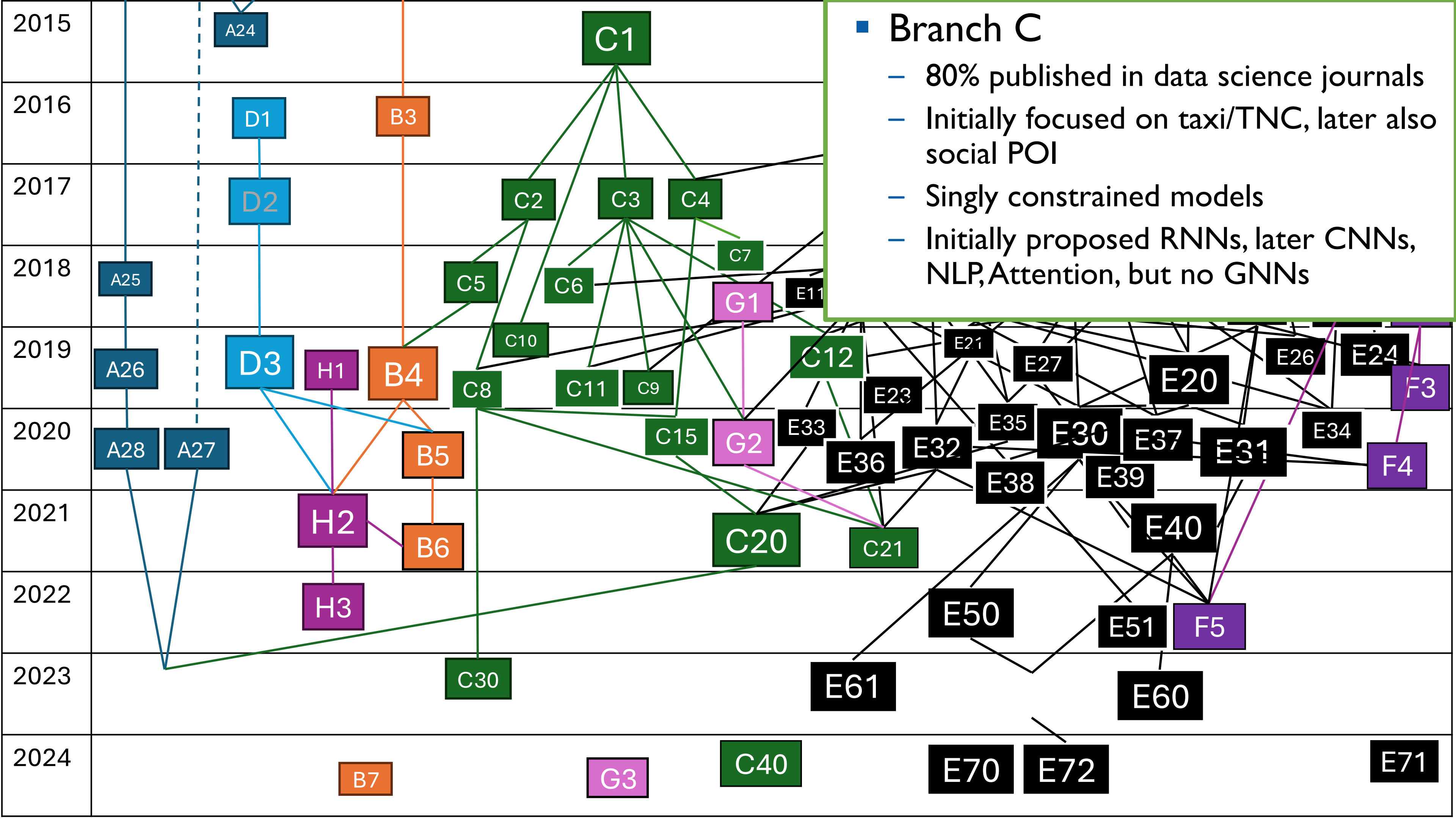
- **Branch A**
  - Mostly published in geography/GIS and transportation journals
  - Initially focused on commuting, later various applications
  - Direct demand models
  - Mostly focused on simple MLPs



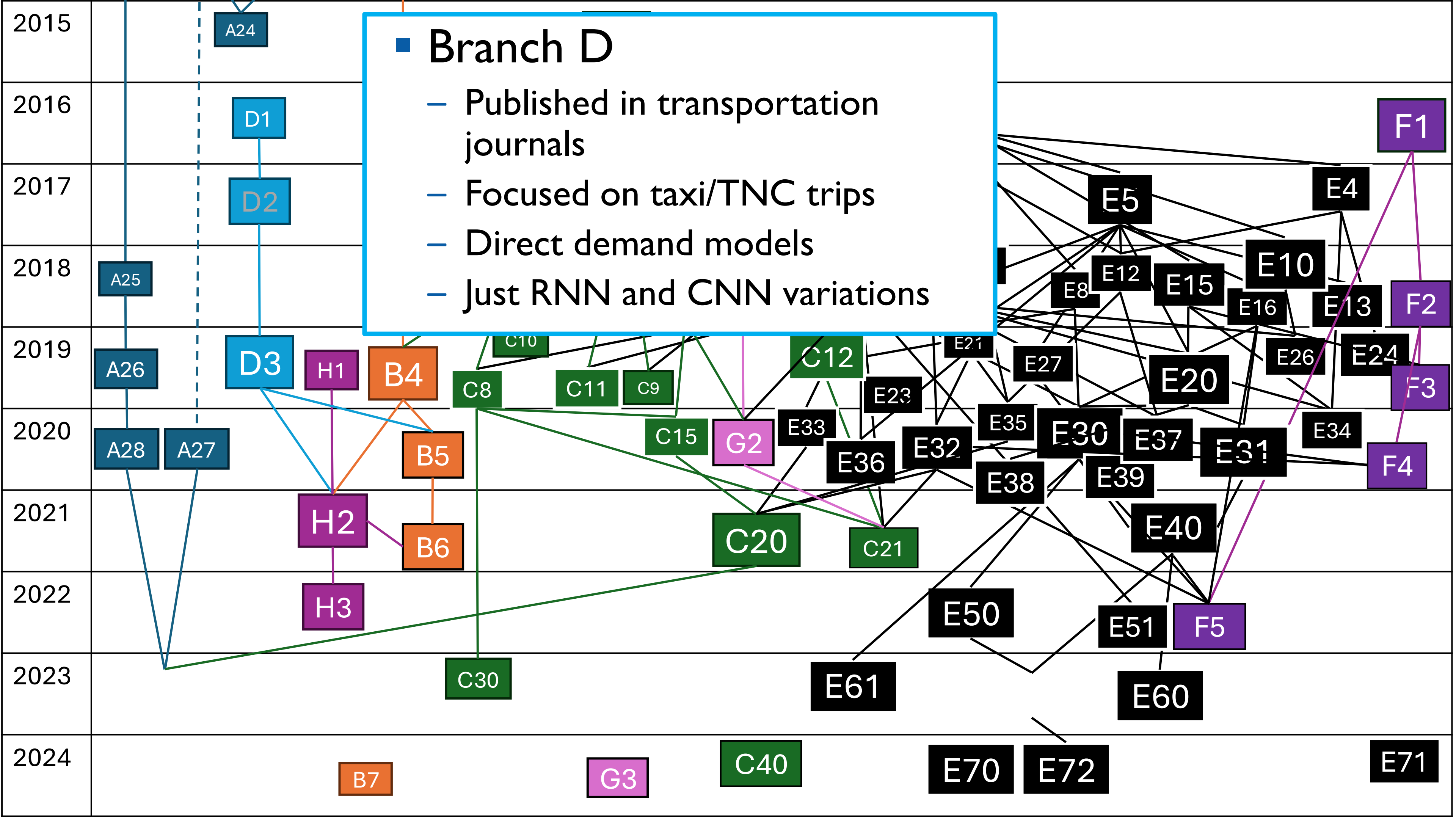
- **Branch B**
  - Mostly published in data science journals
  - Initially focused on taxi/TNC, shifted to transit trips
  - Direct demand models
  - Initially focused on simple MLPs, later incorporated more advanced methods



2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
A24
A25
A26
A27
A28
D1
D2
D3
B3
B4
B5
B6
B7
H1
H2
H3
C1
C2
C3
C4
C5
C6
C7
C8
C9
C10
C11
C12
C15
C20
C21
C30
C40
G1
G2
G3
E1
E2
E3
E4
E5
E6
E7
E8
E9
E10
E11
E12
E13
E15
E16
E20
E21
E23
E24
E26
E27
E30
E31
E32
E33
E34
E35
E36
E37
E38
E39
E40
E50
E51
E60
E61
E70
E71
E72
F1
F2
F3
F4
F5

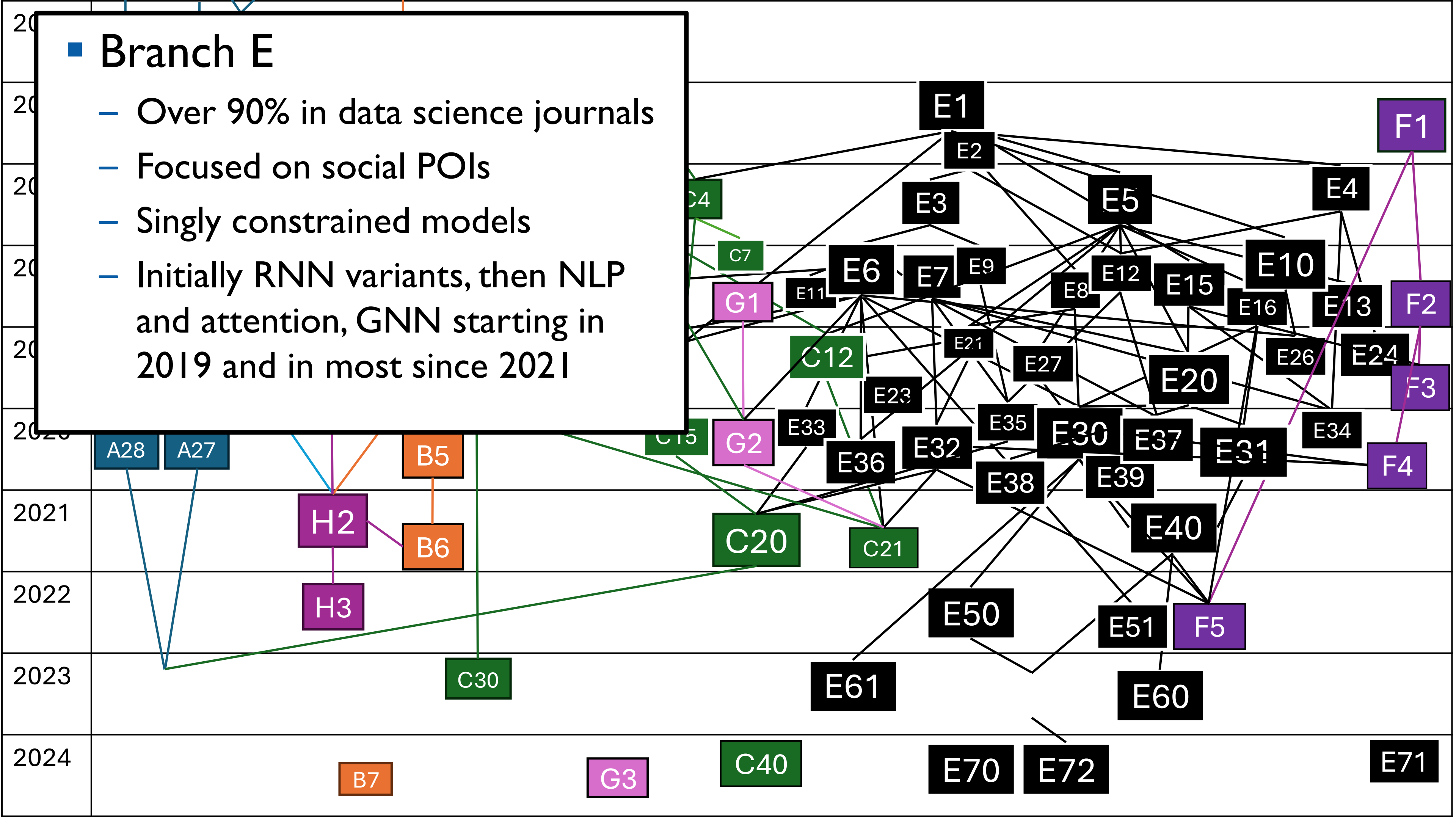


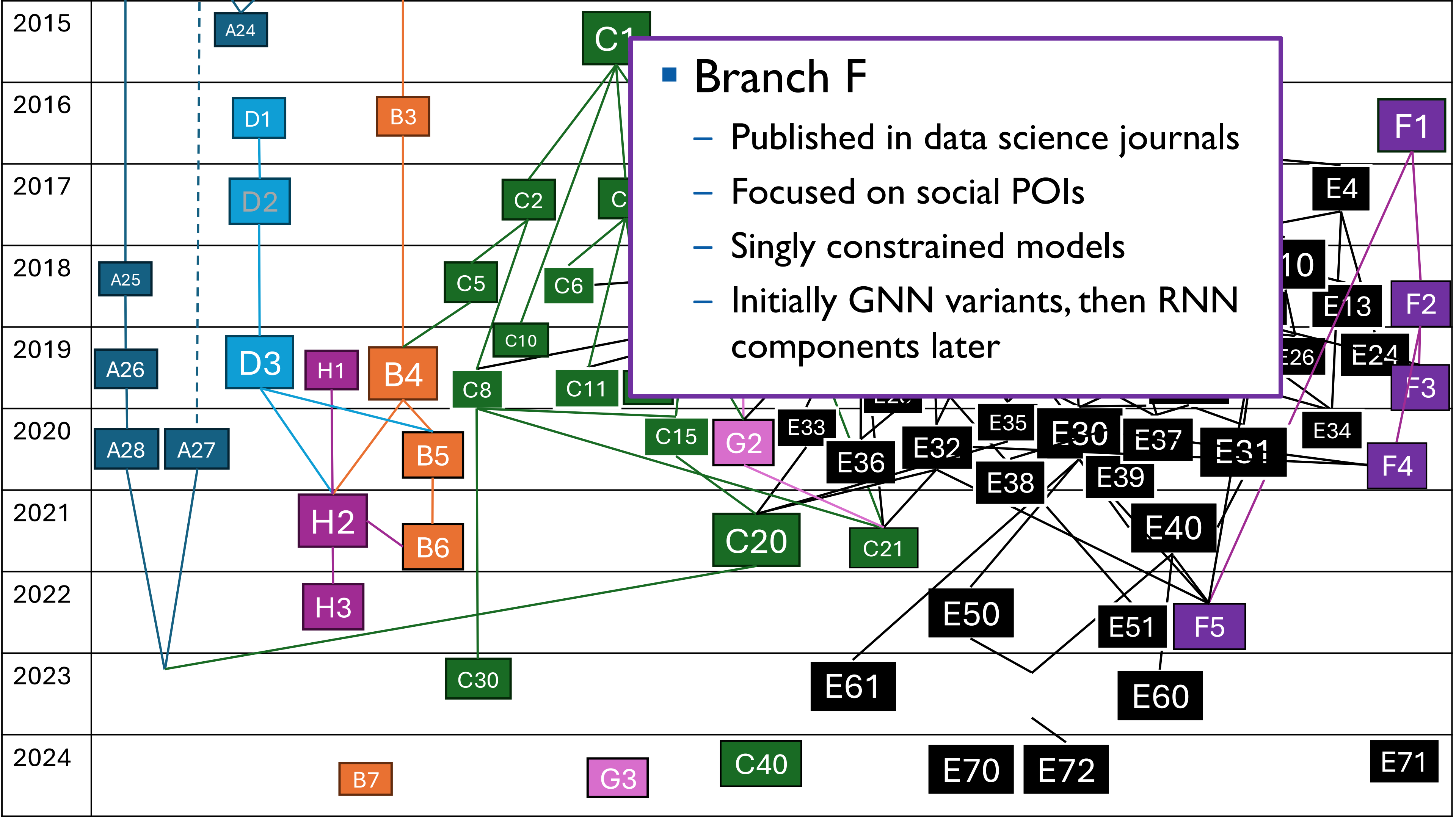
- Branch C
  - 80% published in data science journals
  - Initially focused on taxi/TNC, later also social POI
  - Singly constrained models
  - Initially proposed RNNs, later CNNs, NLP, Attention, but no GNNs

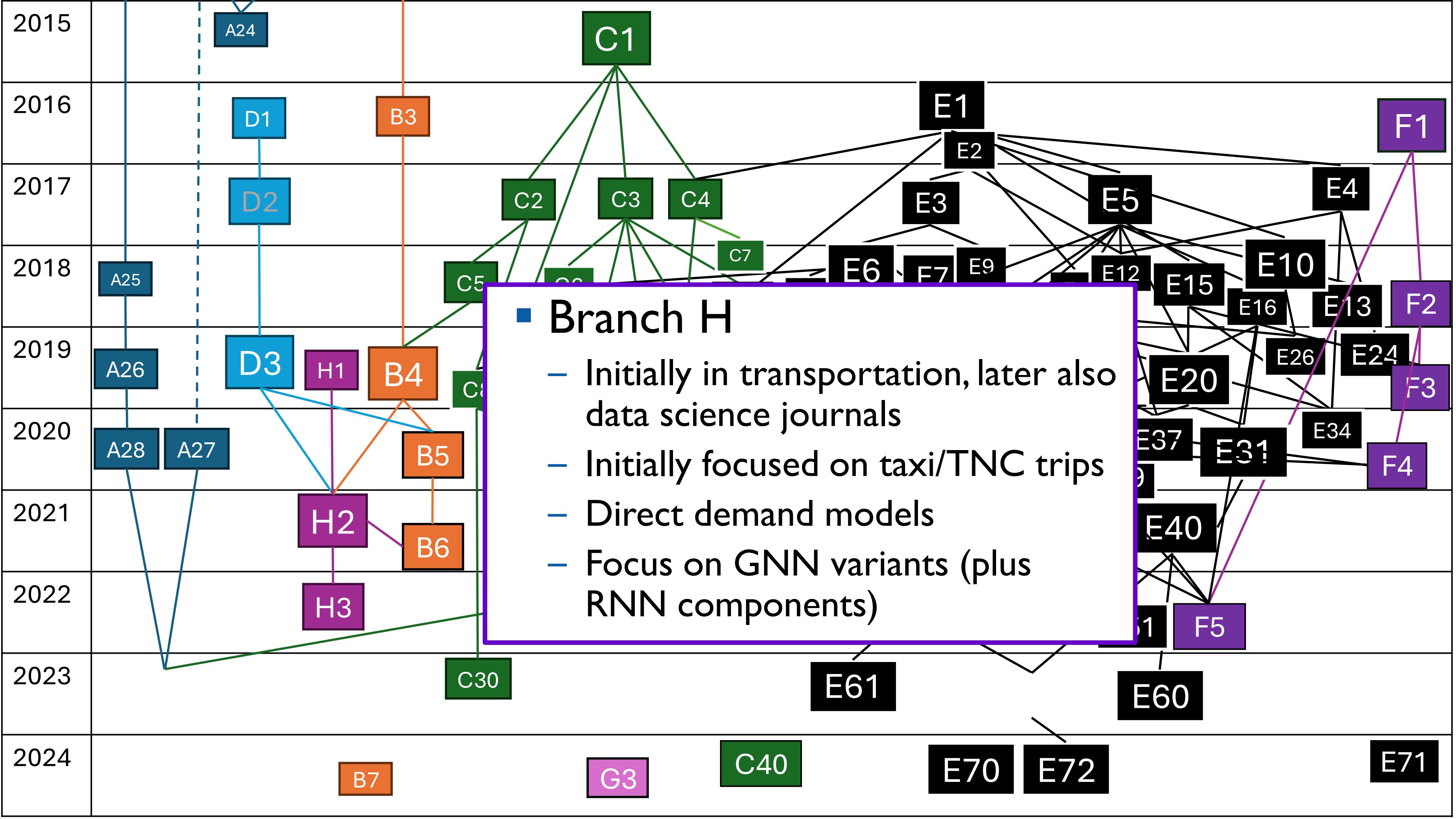


Branch D
Published in transportation journals
Focused on taxi/TNC trips
Direct demand models
Just RNN and CNN variations
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
A24
A25
A26
A28
A27
D1
D2
D3
H1
H2
H3
B4
B5
B6
B7
C8
C10
C11
C9
C12
C15
C20
C21
C30
C40
G2
G3
E4
E5
E8
E10
E12
E13
E15
E16
E20
E21
E23
E24
E26
E27
E30
E31
E32
E33
E34
E35
E36
E37
E38
E39
E40
E50
E51
E60
E61
E70
E71
E72
F1
F2
F3
F4
F5

- **Branch E**
  - Over 90% in data science journals
  - Focused on social POIs
  - Singly constrained models
  - Initially RNN variants, then NLP and attention, GNN starting in 2019 and in most since 2021




Branch F
Published in data science journals
Focused on social POIs
Singly constrained models
Initially GNN variants, then RNN components later
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
A24
A25
A26
A27
A28
D1
D2
D3
B3
B4
B5
B6
B7
H1
H2
H3
C1
C2
C5
C6
C8
C10
C11
C15
C20
C21
C30
C40
G2
G3
E4
E13
E24
E30
E31
E32
E33
E34
E35
E36
E37
E38
E39
E40
E50
E51
E60
E61
E70
E71
E72
F1
F2
F3
F4
F5



Branch H

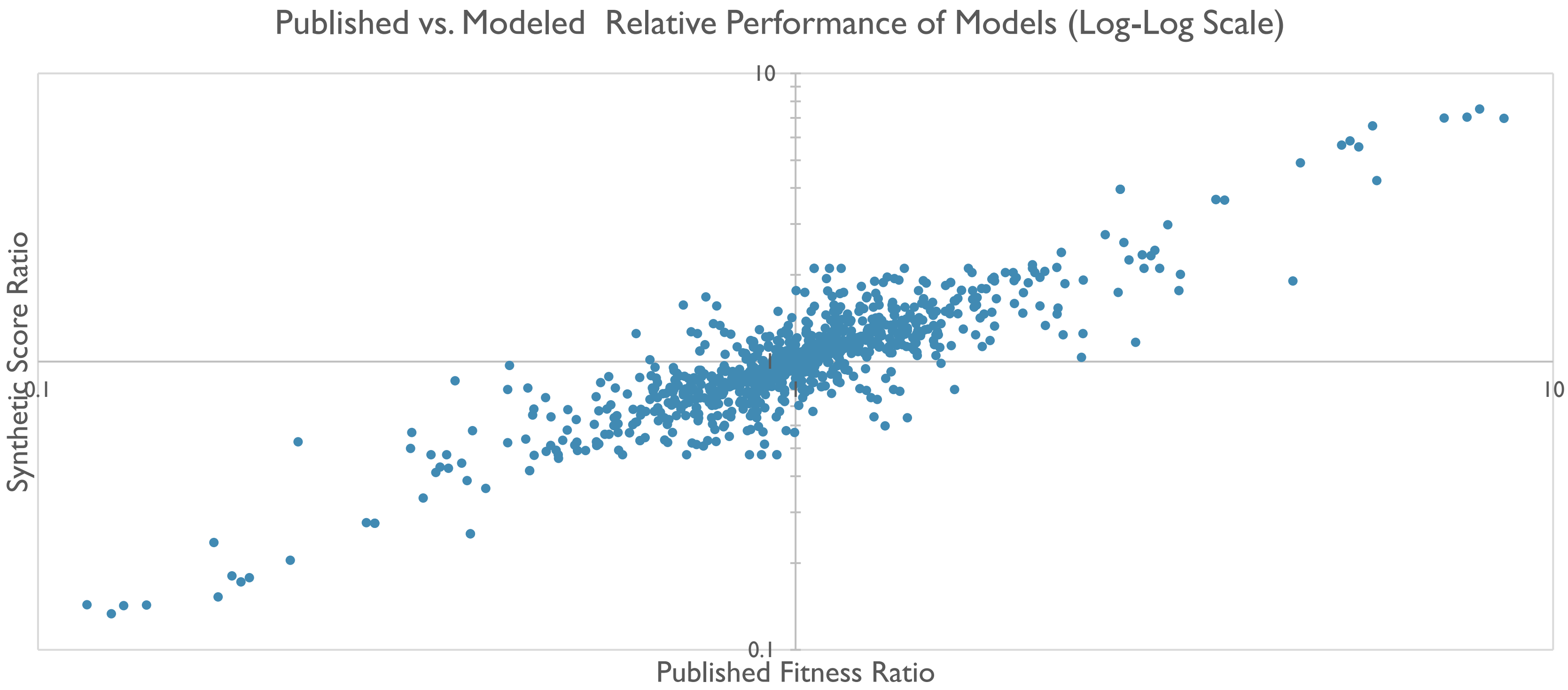
- Initially in transportation, later also data science journals
- Initially focused on taxi/TNC trips
- Direct demand models
- Focus on GNN variants (plus RNN components)

META-ANALYSIS

- Estimated scores for 107 models
- Based on 472 comparisons in 51 papers using 112 datasets
- Initial score calculated as normalized average of ratio of model's goodness-of-fit to other models
- Final score by minimizing squared error of relative comparisons
- Final modeled scores achieved a $r^2 = 0.848$
- No accounting for authorship bias?

SYNTHETIC SCORE RATIOS VS. PUBLISHED

SYNTHETIC SCORE RATIOS VS. PUBLISHED



MAJOR PAPERS WITH META-ANALYSIS SCORES (1)

ID	Paper	Model	Final Score	Initial Score	Year	Cites	Cites / Year	Journal Type	Metric(s)	Constraint	Application
C1	De Brébisson et al. (2015)	RNN	0.25	0.46	2015	254	28.2	Data Science	Distance	1	Taxi/TNC
E1	Liu et al. (2016)	ST-RNN	0.28	0.44	2016	1049	131.1	Data Science	k-Recall, k-F1, MAPE, AUC	1	Taxi/TNC
F1	Xie et al. (2016)	GE	0.04	0.50	2016	440	55.0	Data Science	k-Accuracy	1	All Check-ins
B3	Zhao et al. (2016)	MLP	0.51	0.43	2016	170	21.3	Data Science	sMAPE	0	Taxi/TNC
D1	Toque et al. (2016)	LSTM	0.24	0.41	2016	145	18.1	Transport	MSE	0	Transit OD
E2	Al-Molegi et al. (2016)	STF-RNN			2016	90	11.3	Data Science	k-Recall	1	All Travel
E5	Yang et al. (2017a)	PACE	0.30	0.40	2017	366	52.3	Data Science	k_Hit, k-Prec., k-Rec., k-nDCG, k-MAP	1	All Check-ins
X1	Yin et al. (2017)	SH-CDL	0.41	0.68	2017	300	42.9	Data Science	k-Accuracy, MAE	1	All Check-ins
E3	Yao et al. (2017)	SERM	0.30	0.46	2017	225	32.1	Data Science	k-Hit	1	All Check-ins
E4	Yang et al. (2017b)	JNTM	0.33	0.55	2017	184	26.3	Data Science	k-Recall	1	All Check-ins
C3	Wu et al. (2017)	RNN	0.52	0.46	2017	168	24.0	Data Science	LL, Accuracy	1	Taxi/TNC
C2	Lv et al. (2017)	T-CONV	0.29	0.48	2017	117	16.7	Data Science	Distance	1	Taxi/TNC
E6	Feng et al. (2018)	DeepMove	0.42	0.53	2018	684	114.0	Data Science	Accuracy	1	All Check-ins
E10	Ying et al. (2018)	SHAN			2018	418	69.7	Data Science	k-Recall, AUC	1	All Check-ins
E7	Kong and Wu (2018)	HST-LSTM	0.30	0.54	2018	266	44.3	Data Science	k-Accuracy	1	All Travel
E15	Chang et al. (2018)	CAPE	0.31	0.52	2018	215	35.8	Data Science	k-Recall, MRR	1	All Check-ins
X4	Chu et al. (2018)	MultiConvLSTM	0.34	0.50	2018	174	29.0	Transport	RMSE, sMAPE	0	Taxi/TNC
F2	Wang et al. (2018a)	GeoIE	0.35	0.48	2018	161	26.8	Data Science	k-Recall, k-Precision	1	All Check-ins
E13	Ma et al. (2018)	SAE-NAD			2018	151	25.2	Data Science	k-Precision, k-Recall, k-MAP	1	All Check-ins
G1	Ouyang et al. (2018)	NPGN	1.00	1.00	2018	146	24.3	Data Science	JSD	1	All Travel
E8	Manotumruska et al. (2018)	CARA	0.30	0.46	2018	144	24.0	Data Science	k-Accuracy, NDCG@10	1	All Check-ins
E12	Zhao et al. (2018b)	ST-LSTM	0.36	0.55	2018	71	11.8	Data Science	k-Accuracy, MAP	1	All Check-ins
A25	Pourebrahim et al. (2018)	MLP	0.51	0.43	2018	47	7.8	Data Science	RMSE		All Check-ins
E11	Atlaf et al. (2018)	STA-GRU	0.37	0.53	2018	32	5.3	Data Science	k-Recall, k-NDCG, AUC, MRR	1	All Check-ins
B4	Wang et al. (2019)	GEML	0.29	0.70	2019	275	55.0	Data Science	RMSE, sMAPE	0	Taxi/TNC
C8	Rossi et al. (2019)	NLP-LSTM	0.29	0.49	2019	114	22.8	Transport	Distance	1	Taxi/TNC
D3	Liu et al. (2019a)	ConvLSTM	0.31	0.47	2019	263	52.6	Transport	RMSE, MAPE	0	Taxi/TNC
E20	Zhao et al. (2019a)	STGCN	0.32	0.48	2019	484	96.8	Data Science	k-Accuracy, MAP	1	All Check-ins
X5	Fang et al. (2019)	GSTNet	0.33	0.51	2019	191	38.2	Data Science	MAE, SMAPE		Transit OD
E24	Huang et al. (2019)	ATST-LSTM	0.46	0.60	2019	180	36.0	Data Science	k-Precision, k-Recall, k-F1	1	All Check-ins

MAJOR PAPERS WITH META-ANALYSIS SCORES (2)

ID	Paper	Model	Final Score	Initial Score	Year	Cites	Cites / Year	Journal Type	Metric(s)	Constraint	Application
F3	Qian et al. (2019)	STA			2019	157	31.4	Data Science		1	All Check-ins
C12	Gao et al. (2019a)	VANext	0.43	0.55	2019	126	25.2	Data Science	k-Top	1	All Check-ins
E23	Zhuo et al. (2019b)	APOIR	0.26	0.54	2019	124	24.8	Data Science	k-Precision, K-Recall, k-MAP, k-DCG	1	All Check-ins
E26	Wu et al. (2019)	LSPL			2019	82	16.4	Data Science	k-Precision, k-MAP	1	All Check-ins
C11	Zhang et al. (2019)	ELM	0.33	0.47	2019	80	16.0	Transport	RMSE, MAE	1	Taxi/TNC
A26	Pourebrahim et al. (2019)	MLP	0.51	0.43	2019	60	12.0	Data Science	MSE, R2	0	All Check-ins
E30	Sun et al. (2020)	LSTPM	0.37	0.52	2020	323	80.8	Data Science	k-Recall, k-NDCG	1	All Check-ins
E31	Lian et al. (2020)	GeoSAN	0.46	0.55	2020	227	56.8	Data Science	k-HR, k-NDCG	1	All Check-ins
E36	Wu et al. (2020a)	PLSPL	0.39	0.48	2020	170	42.5	Data Science	k-Precision, k-MAP	1	All Check-ins
E32	Yang et al. (2020)	Flashback	0.48	0.60	2020	160	40.0	Data Science	k-Accuracy, MRR	1	All Check-ins
B5	Liu et al. (2020)	PVCGN	0.27	0.49	2020	140	35.0	Transport	RMSE, MAE, MAPE	0	Transit OD
E37	Yu et al. (2020)	CatDM			2020	137	34.3	Data Science		1	All Check-ins
E38	Zhao et al. (2020a)	ASPPA	0.31	0.53	2020	124	31.0	Data Science	k-HR, MRR	1	All Check-ins
E39	Lim et al. (2020)	STP-UDGAT	0.33	0.52	2020	121	30.3	Data Science	k-Accuracy, MAP	1	All Check-ins
F4	Feng et al. (2020b)	HME	0.27	0.50	2020	121	30.3	Data Science	k-Recall, k-Precision, MAP	1	All Check-ins
X6	Shi et al. (2020)	MPGCN	0.29	0.48	2020	92	23.0	Data Science	RMSE	0	Taxi/TNC
A27	Yao et al. (2020)	SI-GCN			2020	87	21.8	Transport		0	Taxi/TNC
E33	Chen et al. (2020)	DeepJMT	0.45	0.61	2020	71	17.8	Data Science	k-Hit, MAP	1	All Check-ins
A28	Liu et al. (2020c)	GMEL	0.19	0.51	2020	62	15.5	Data Science	RMSE, MAE, DSC	0	Commuting
E40	Luo et al. (2021)	STAN	0.44	0.54	2021	275	91.7	Data Science	k-Recall	1	All Check-ins
H2	Ke et al. (2021)	ST-ED-RMGC	0.30	0.54	2021	170	56.7	Transport	RMSE, MAE, MAPE	0	Taxi/TNC
C20	Simini et al. (2021)	DeepGravity	0.46	0.68	2021	162	54.0	Other	RMSE, DSC, JSD	1	Commuting
B6	Zhang et al. (2021)	CAS-CNN	0.32	0.48	2021	126	42.0	Transport	RMSE, MAE, WMAPE	0	Transit OD
E50	Yang et al. (2022)	GETNext	0.46	0.54	2022	122	61.0	Data Science	k-Accuracy, MRR	1	All Check-ins
H3	Zhang et al. (2022a)	DNEAT			2022	69	34.5	Transport		0	Taxi/TNC
F5	Wang et al. (2022a)	GSTN	0.36	0.46	2022	68	34.0	Data Science			All Check-ins
E60	Long et al. (2023)	DCLR	0.46	0.56	2023	57	57.0	Data Science	k-HR, k-NDCG	1	All Check-ins
E61	Qin et al. (2023)	DisenPOI	0.37	0.47	2023	50	50.0	Data Science	AUC, Logloss	1	All Check-ins
E62	Yan et al. (2023)	STHGCN	0.56	0.60	2023	39	39.0	Data Science	k-Accuracy, MRR	1	All Check-ins
C30	Wang et al. (2023a)	LLMob	0.39	0.48	2023	37	37.0	Data Science	k-Accuracy, W F1, k-nDCG	1	All Travel
A30	Yin et al. (2023)	ConvGCN-RF	0.26	0.58	2023	33	33.0	GIS/Geography	RMSE, MAPE, DSC	0	Commuting
E70	Feng et al. (2024)	LLMove	0.60	0.55	2024	10	20.0	Data Science	k-Accuracy, MRR	1	All Check-ins
E72	Li et al. (2024)	LLM4POI	0.65	0.64	2024	11	22.0	Data Science	k-Accuracy	1	All Check-ins
E71	Fu et al. (2024)	SLS-REC			2024	15	30.0	Data Science		1	All Check-ins

META-ANALYSIS RESULTS

- Best methods
  - GAI
    - GAN
    - LLM
  - SSL
  - CNN
    - CNN
    - GCN

Method	Avg. Score
GAI	0.580
SSL	0.455
CNN	0.413
Attention	0.398
FCN	0.397
GNN	0.385
RNN	0.383
NLP	0.328

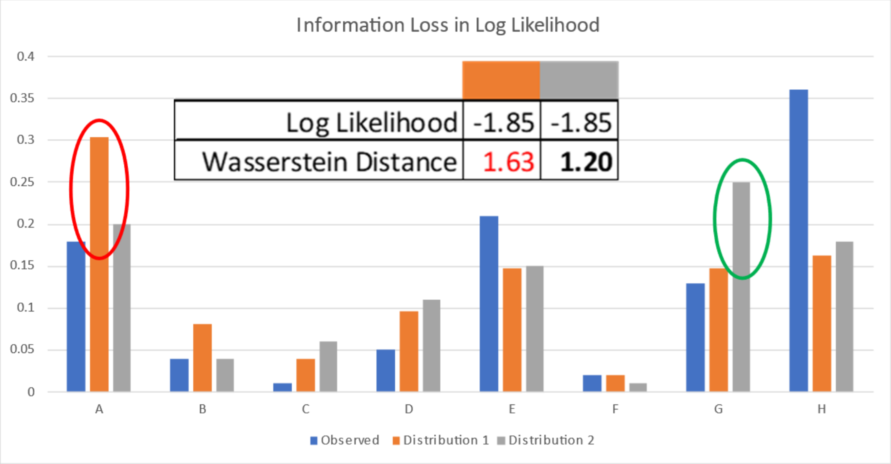
- Recommended models for testing in AI-DCMs
  - DeepGravity (MLP), reference
  - TrajGAN (GAN), highest score
  - STHGCN (SSL GCN), #4 highest score, highest non-GAI, score based on 8 comparisons

- GAI & SSL small sample size
- MLP (FCN) not bad
- Other methods not significant

OTHER REPORT RECOMMENDATIONS

PERFORMANCE MEASUREMENT

- Importance of Out-of-Sample (Holdout Sample) Validation
  - Standard practice of good data science
  - Extremely rare in travel forecasting practice
  - Key opportunity to improve the practice
- Choice of Metric
  - Huge variety of error / goodness-of-fit metrics
  - Minimum Wasserstein distance
    - Powerful in computer vision, with CNNs
    - Gives credit for getting close



SOFTWARE

- Language: Python
  - Compatible with TransCAD, OpenPaths, VISUM
  - Most widely used language for data science
- Data Science Libraries: PyTorch vs. Keras
  - Scikit Learn – many ML/AI methods, but limited DNN
  - Tensorflow – powerful, heavy-duty, complex, difficult to learn
  - Keras – wrapper for Tensorflow
  - PyTorch – more complex than Keras, simpler than Tensorflow

DATA

- Proposed a taxonomy of variables and data sources
- Interested in expert panel feedback on the use / inclusion of variables and data sets as well as derived variables

VARIABLES AND DATA SOURCES

	Use	Household Travel Survey	Census Commute Flows	Aggregate Big Data	Disaggregate Big Data	Censu GIS Data	Confidential QCEW Data	Commercial Establishment Data	Open POI Data	Commercial Visitation Big Data	Parcel Data	Open GIS Data	Commercial Travel Time Data
Acquisition Cost		$$$$		$$	$$			$$		$$$			$$
Processing Cost		$$	$	$	$$	$	$$	$$	$	$	$	$$	$
Choice Observations													
Individual Choices	Common	X			X								
Aggregate Choices	Common		X	X						X			
Choice-Maker / Context Variables													
Primary Variables													
Income	Common	X				X							
Auto ownership	Common	X				X							
Age	Uncommon	X				X							
Gender	Uncommon	X				X							
Family/household members	Uncommon	X				X							
Employment status	Uncommon	X				X							
Race	Rare	X				X							
Time-of-day	Uncommon	X											
[Home location variables]	Rare												
Derived variables													
Latent class	Rare												

	Use	Household Travel Survey	Census Commute Flows	Aggregate Big Data	Disaggregate Big Data	Censu GIS Data	Confidential QCEW Data	Commercial Establishment Data	Open POI Data	Commercial Visitation Big Data	Parcel Data	Open GIS Data	Commercial Travel Time Data
Acquisition Cost		$$$$		$$	$$			$$		$$$			$$
Processing Cost		$$	$	$	$$	$	$$	$$	$	$	$	$$	$
Choice Alternatives (Location) Variables													
First Order (Single spatial index, location attributes)													
Primary Variables													
Employment by Industry by Zone	Standard						X	X					
Demographics by Zone	Standard					X							
Zoning category	Uncommon										X		
Square footage	Uncommon										X	X	
Park area	Uncommon											X	
Cemetery area	Uncommon											X	
Water area & boundaries	Uncommon											X	
State / County / City	Uncommon											X	
Railroad	Uncommon											X	
Land cover	Rare											X	
Establishments by Industry by Zone	Rare						X	X					
Category	Rare								X	X			
Industry	Rare						X	X					
Employees	Rare						X	X					
Footfall / crowd flow	Rare									X			
with sales	Rare							X					
credit score	Rare							X					
Derived variables													
Accessibilities	Uncommon												
Land use diversities	Uncommon												
Densities	Uncommon												
Terrain	Rare												
Second Order (Two spatial indices; location pair attributes)													
Primary Variables													
Travel time & network distance	Standard	X			X								X
Derived variables													
Boundary Crossings	Uncommon												
Similarity / Dissimilarity	Rare												

OTHER AI APPLICATIONS IN TRAVEL FORECASTING

Caliper

TRIP / ACTIVITY GENERATION

- As with destination choice, most of the literature is in data science journals where the problem is most commonly called "crowd flow prediction"
- Some limited transportation literature
  - Decision trees outperform traditional statistical models
- Most applications in the practice of any ML/AI methods
  - May want to do a webinar in Task 5 on this

MODE CHOICE

- Most extensive literature on AI/ML in transportation journals
  - Perhaps 1,000 papers
- Most compare AI/ML method to logit
  - Most comparisons lack a well-calibrated logit model
  - Most tend to collapse transit sub-modes
- Most seemingly valid comparisons indicate AI/ML offer some, but modest improvements over logit
  - Logit mode choice models generally perform well, hence limited room for improvement

VOLUME PREDICTION

- Generally focused on short-term forecasting
- Generally rely on significant historical data
- Closely related to speed prediction
- Many similar methods as in destination choice

CONTACTS

Vince Bernardin, PhD | Vice-President

vince@caliper.com | +1 812-459-3500

Caliper