Network Assignment-Based Estimation of Origin-Destination Matrices with a Full Travel Demand Model

CENTER FOR TRANSPORTATION RESEARCH

William Alexander, Carlin Liao, Rishabh Thakkar, Karthik Velayutham, Dr. Stephen Boyles

TEXAS
The University of Texas at Austin

Introduction

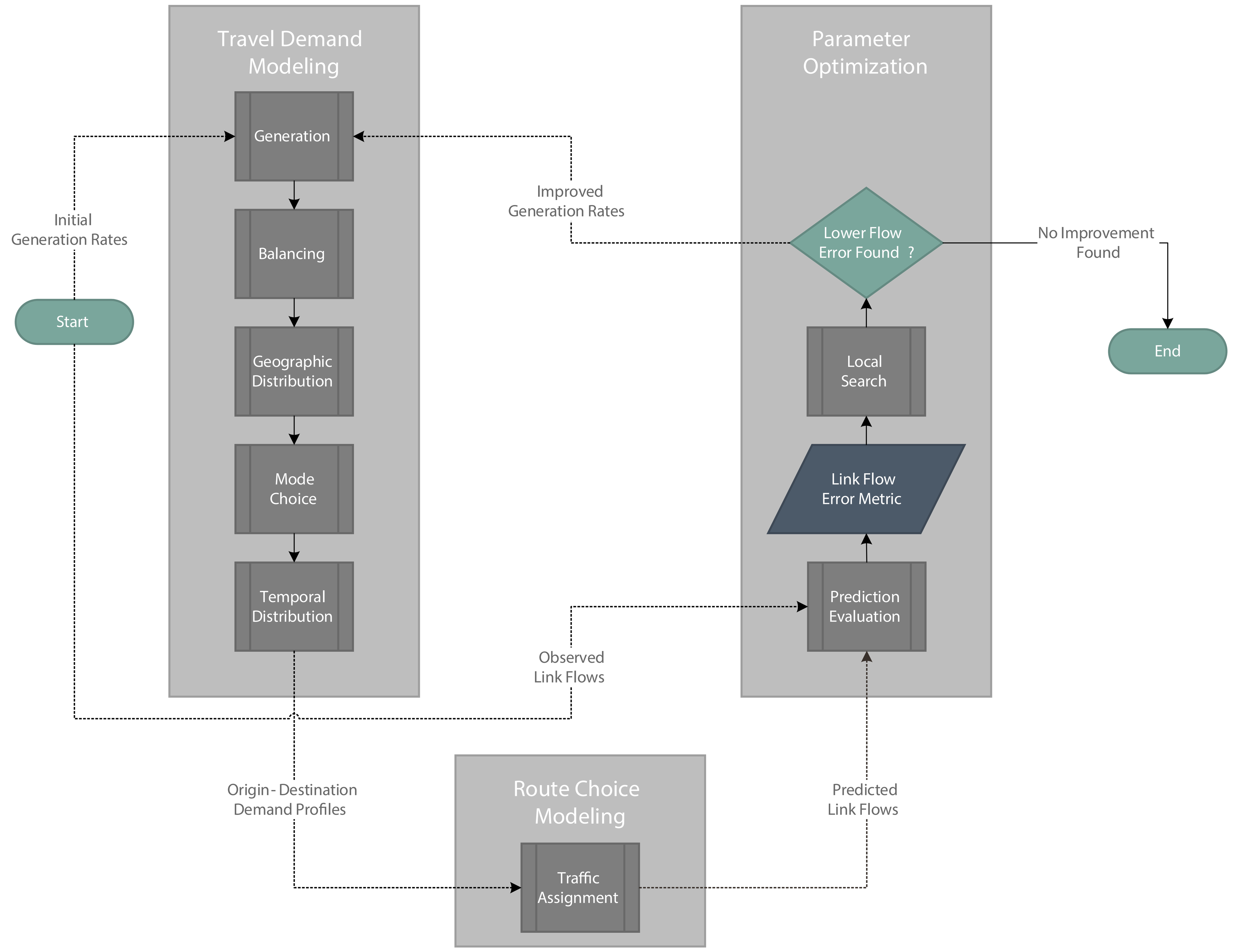
In the network modeling community, traditional techniques for origin-destination matrix estimation focus on adjusting the OD matrix directly, or on tuning the parameters of the distribution model. **These approaches are prone to overfitting and reduce the behavioral interpretation of the OD matrix.**

Instead, **we propose tuning parameters in trip generation**, the earliest stage of the traditional four-step model. Our procedure calibrates an initial estimate of trip generation rates by **using a local search to reduce the error between the network flows** predicted by a demand model and field link flow observations as a proxy for finding the true (but unobservable) trip generation rates.

We establish two blind tests in which a “true” configuration of generation parameters was developed, then a noisy estimate was provided as a starting point for our search algorithm. Then, using the RMSE error metric to determine the closeness of our four-step model’s predicted flows to the “true” observed flows (i.e.. the results from the true parameters’ output route choice), we determine a search direction which guarantees improvement in link flow RMSE.

Using our results, we identify how to set the local search termination criterion for consistent performance across scenarios, and how to adjust the demand model to prioritize reducing error for the trip generation rates of greatest consequence.

Framework



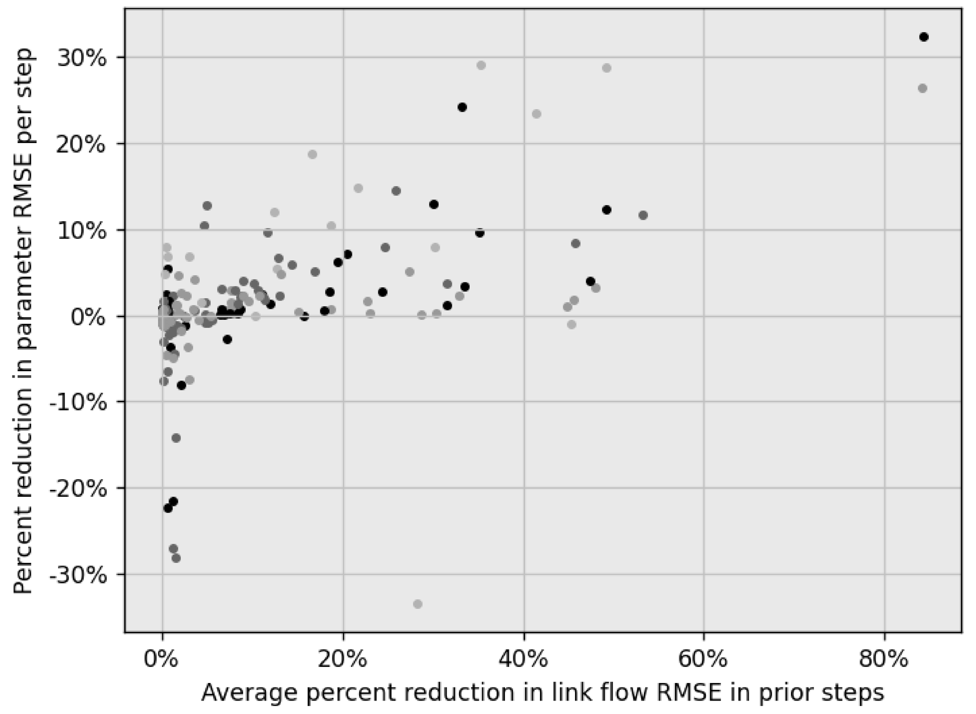
Acknowledgement: This work is supported by the North Central Texas Council of Governments (NCTCOG) through their University Partnership Program. The authors wish to thank Arash Mirzaei and Dr. Hong Zheng of NCTCOG for their guidance throughout this project. We also wish to express gratitude to Dr. Felipe Dias, Dr. Venktesh Pandey, and Priyadarshan Patil for their assistance with this project.

OD Matrix Calculation

The initial steps of the trip model (those preceding traffic assignment) were implemented inside the scope of the free-and-open-source software package *wrap* (Alexander, Patel, et al. 2018). The *wrap* package includes functionality to complete the traditional four steps of the Urban Transportation Modeling System (aka the four-step model), namely trip generation, trip distribution, mode choice, and route choice. *Wrap* outputs in the form of a collection of OD matrices, one for each modeled value of time and vehicle class combination. These OD matrices were loaded into an implementation of Algorithm B, which solved the static traffic assignment problem (Dial, 2006).

Prior research indicates that travelers can be segmented into multiple markets, each with their own values of time which impact travel decision-making. Each market segment’s individual trip purposes are primarily independent four-step models, each with their own trip generation, distribution, and mode choice steps. By developing distinct models for each market, a variety of travel-making behaviors and their interaction effects can be modeled simultaneously.

Using these additional markets poses only a marginal increase in computation time, so we can rapidly prototype new model configurations to develop error values and a search direction for use in our algorithm.



Averaged percent reduction in link flow RMSE vs. per-step reduction in parameter RMSE for Sioux Falls trials

Parameter ID	Trips	% Error At Search Termination		% Error At Early Termination		% Error At Start	
		Mean	Median	Mean	Median	Mean	Median
Sioux Falls							
iP1	6,959.82	6.04%	6.46%	5.89%	5.68%	5.17%	4.82%
iP2	9,317.02	8.72%	6.82%	8.05%	6.38%	7.80%	7.09%
iA1	50,449.11	1.46%	1.44%	1.71%	1.58%	4.82%	3.50%
iA2	101,877.16	0.51%	0.53%	1.03%	0.90%	7.05%	4.85%
iiP1	16,912.80	7.61%	6.68%	8.17%	7.27%	8.37%	7.27%
iiP2	9,320.80	7.47%	5.63%	6.03%	3.12%	6.77%	3.98%
iiA1	27,362.40	10.51%	10.47%	4.31%	2.73%	7.99%	5.45%
iiA2	42,026.82	5.54%	4.53%	2.75%	2.42%	4.27%	2.48%
Chicago Sketch							
iiiP1	150,502.33	1.72%	0.55%	3.43%	3.00%	9.75%	10.15%
iiiP2	29,149.52	3.04%	1.37%	10.77%	9.63%	8.74%	6.54%
iiiA1	17,172.37	9.90%	8.30%	9.99%	9.89%	10.08%	8.67%
iiiA2	22,970.61	8.07%	6.67%	12.16%	9.16%	13.75%	12.17%
ivP1	49,432.37	2.24%	0.19%	4.38%	1.54%	8.28%	4.36%
ivP2	229,367.75	0.29%	0.19%	0.58%	0.37%	6.94%	4.64%
ivA1	388,644.23	14.08%	14.00%	7.64%	7.17%	7.72%	7.12%
ivA2	63,052.41	8.56%	8.63%	6.17%	5.97%	6.27%	6.22%

Parameter percent error statistics

Model Parameter Improvement

Given an error function ϵ (we use RMSE), we use L-BFGS-B local search to successively modify an initial t to produce progressively better flow estimates (Byrd 2003). Let $h(t) = \epsilon\left(F(t)\right)$ be our link flow error objective function and t_i be the estimated trip generation rate parameters. Initialize Hessian estimate H_0 as the identity matrix of the same size as the number of rates. Starting with an initial guess t_0 and its corresponding objective function value x_0, each search iteration proceeds as follows:

1. Determine if any entries of t_i are at their boundary, 0. If so, fix their values at 0 and do not include them in the step calculations, effectively reducing the size of our input parameters. Remove their associated dimensions from H_0 and other values to match.
2. Find the descent direction $p_i = -H_i^{-1} g_i$.
3. Find the optimal step size along that direction $\alpha = \underset{\alpha}{\operatorname{argmin}}\, h(t_i + \alpha p_i)$ using backtracking line search.
4. Set our new best guess $t_{i+1} = t_i + \alpha p_i$.
5. Increment i and use finite differencing to calculate the function’s gradient g_i at the new guess. Terminate the search here if the length of the gradient is sufficiently small.
6. Let $s = \alpha p_{i-1}$, $y = g_i - g_{i-1}$, and $U = I - \frac{\vec{y} s^T}{\vec{y}^T s}$. Update our Hessian estimate $H_i = U H_{i-1} U + \frac{s s^T}{\vec{y}^T s}$.

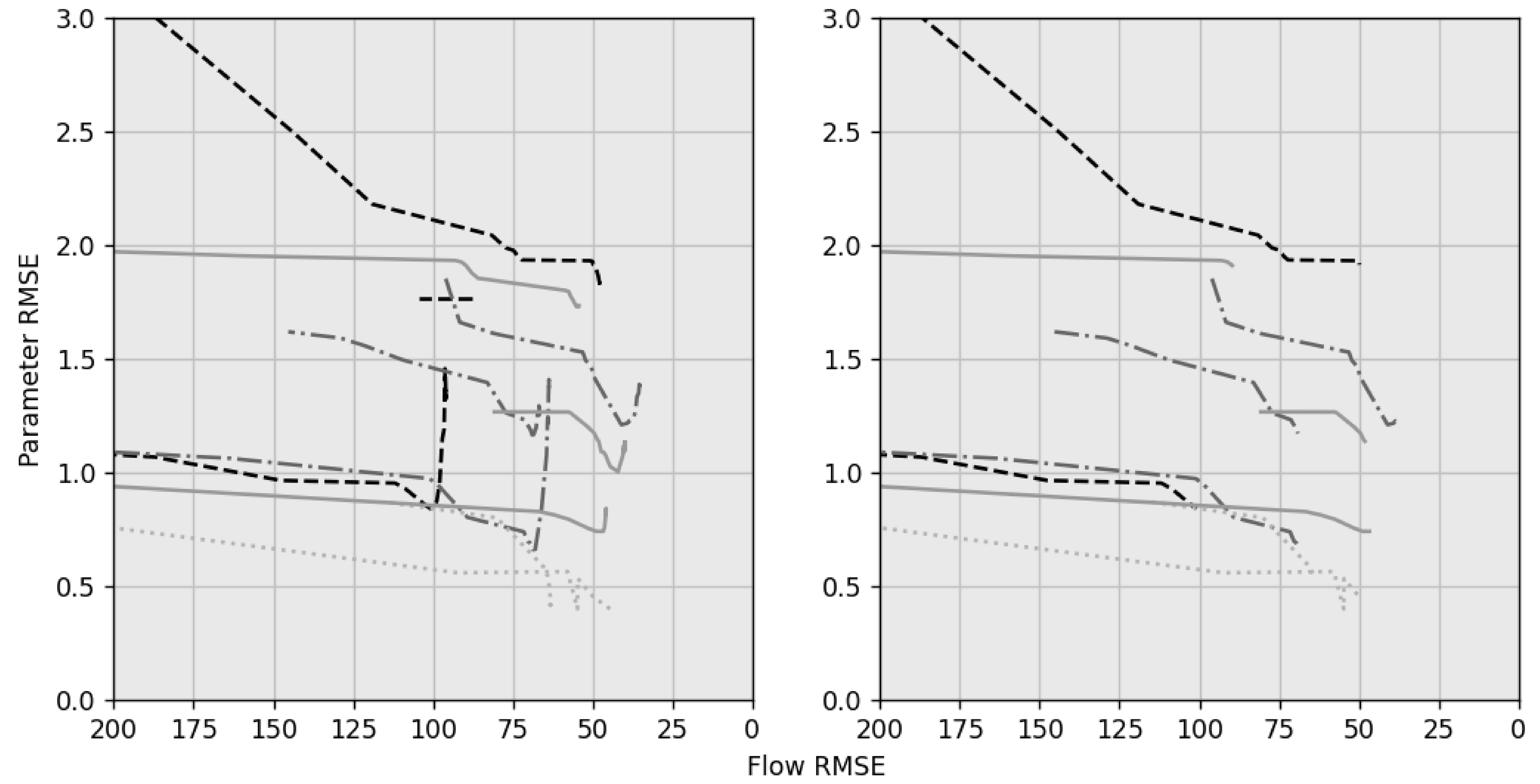
Validation and Evaluation

The framework we propose improves a noisy initial estimate of trip generation rates given only observed link flows. Provided an accurate demand model for a study network, we can tune the generation rates given by a demand survey within their confidence intervals, bringing in flow information from loop detectors and other sensors to further refine our demand model parameters as well as the resulting OD matrix.

Our method was shown to reduce link flow and generation rate parameter RMSE by an average of 81.6% and 35.6%, respectively, in the Sioux Falls test scenario, and by 47.1% and 26.2% in the Chicago scenario (Transportation Networks 2007). This performance is contingent on terminating the search if the average improvement over the last five step iterations falls below 2.5%; failing to do so may lead to overfitting the starting noisy observation and observed link flows, thus nullifying many of the improvements made by our method.

Furthermore, although a typical demand model deciding how trip purposes are balanced is a design decision, our method performs much better on parameters that aren’t obfuscated by scaling during the trip balancing step of the UTMS. We recommend that the trip generation rates a modeler deems more important, e.g.. categories responsible for more trips, should have their rates fixed during the balancing step where possible. Rates the modeler deems as less important should be balanced against those fixed purposes.

Finally, we note that although our method was able to improve on every noisy trip generation rate provided, we found that the quality of the rates output at termination is dependent on the initial estimate, instead of converging to the true parameters. This suggests that further research should explore ways to escape local optima without overfitting.



Link flow and parameter RMSE trends for the final steps of Sioux Falls trials, including the full local search (left) and truncating using the proposed early termination criteria (right)

Conclusions

- A method was developed to fine-tune travel demand model parameters using a local search algorithm, using the RMSE of predicted vs. observed network flows as an objective function
- Blind tests confirmed the efficacy of the method in driving the parameters closer to their ground truth values
- Patterns were determined for identifying the trip generation components with the largest impact on RMSE which should be prioritized in searching for parameter improvements
- A rule of thumb was determined for terminating the local search early to reduce overfitting
- Our methods allow for rapid prototyping of travel demand models to assess their reflection of real-world conditions