Pedestrian Routing Substrate for Midtown Manhattan

Academic Research Brief for StreetLight Data

Gil Raitses, Researcher, Syracuse University
gilraitses@gmail.com +1 212.365.0255

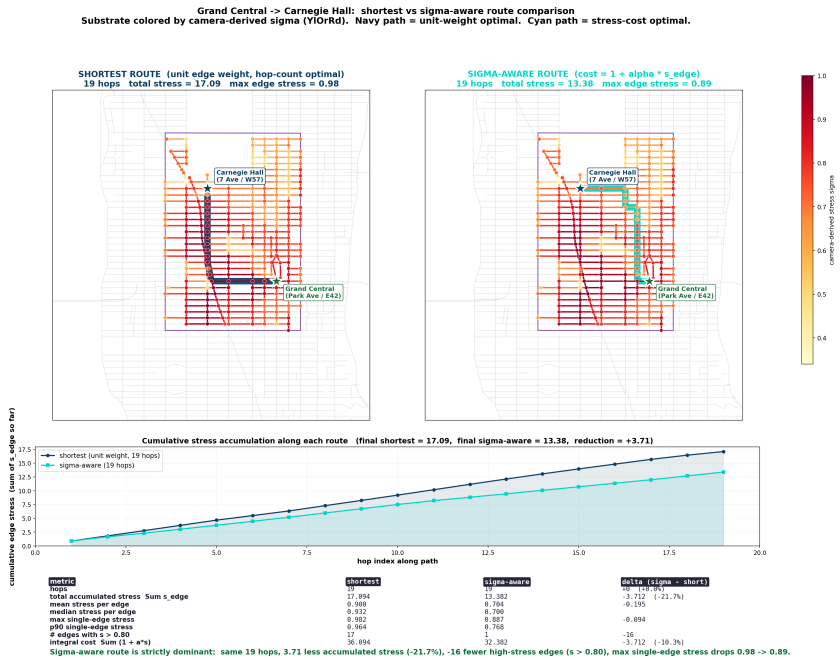
April 24, 2026

Abstract

I am building a graph-based pedestrian routing substrate for the Midtown Manhattan corridor (40th to 59th Streets, Lexington Avenue to 8th Avenue) as part of an academic research project supporting an NSF NRT seed grant proposal. The substrate combines a camera-derived stress field (demand side, from 181 NYC DOT traffic cameras at 30-minute cadence) with a seven-layer cross-street capacity model (supply side, derived from 25 public NYC Open Data sources covering sidewalk geometry, building footprints, scaffolding sheds, station envelopes, open streets, bike lanes, and dining permits). The routing layer combines the two as cost equals stress over capacity, then runs A* pathfinding under either a unit-weight (control) or stress-aware (treatment) cost function.

Validation status: the substrate is internally validated (lived-knowledge alignment at 8 of 12 reference blocks; supply-vs-demand orthogonality at Pearson r = −0.014). **The validation gap is empirical pedestrian volume data at the per-block resolution.** StreetLight Bicycle and Pedestrian Volumes is the closest commercial source for this; public NYC DOT counts are available only at ∼100 fixed locations city-wide, none at the per-block scale the substrate operates on.

Demonstration result



Shortest route (left) vs. stress-aware route (right) between Grand Central and Carnegie Hall. Same nineteen hops; sigma-aware route accumulates 13.38 units of stress versus 17.09 (a 21.7% reduction); sixteen fewer high-stress edges; max single-edge stress drops from 0.98 to 0.89. The contrast generalizes across nine

destinations from Grand Central; recommendations differ from the shortest path in eight of nine cases.

What StreetLight data would unlock

- **Calibrate the stress regression against ground truth.** The camera-derived stress score is currently calibrated against author lived knowledge as the primary anchor. Bicycle and pedestrian volume data at the per-block scale would let the regression be calibrated against empirical foot-traffic.
- **Validate the routing recommendations.** The substrate's stress-aware route claim (22% reduction in accumulated stress) is currently a substrate-level claim. Empirical pedestrian volume data along both candidate routes would let the claim be validated against measured exposure.
- **Time-varying calibration.** The substrate currently varies only in subway egress (driven by MTA hourly ridership). Bicycle and pedestrian counts at peak vs. off-peak hours would let the substrate's other capacity layers (sheds, frontage, station envelopes) be time-varied as well.
- **Publishable research output.** The substrate is the empirical backbone of a graph-theoretic argument about pedestrian routing under second-order constraint structure; a forthcoming paper will cite the data sources used.

Specific data ask

- **Geographic scope:** Midtown Manhattan corridor, 34th to 70th Streets, 3rd Avenue to 8th Avenue (one tier wider than the operational substrate to allow boundary-effect analysis).
- **Temporal scope:** a representative period of approximately one calendar month, ideally with weekday-vs-weekend breakdown and hourly resolution.
- **Spatial resolution:** per-block (not aggregate corridor-level); the substrate operates on 228 cross-street edges and would consume per-edge volume data directly.
- **Modes:** pedestrian volumes primary; bicycle volumes secondary (the substrate has a bike-lane capacity layer that would benefit from validation).
- **Optional second tier:** Manhattan-wide for the longer-term substrate extension (out of scope for the IST 675 deliverable, in scope for the NSF NRT direction).

Academic context and request

The research is hosted at Syracuse University. The NSF NRT seed grant proposal is in preparation; the project advisor is Dr. Jen Schwarz. The work is structured for publication and will cite all data sources used. **I am asking whether StreetLight has a University Research program or academic pricing for a single-geography validation dataset of this kind, and whether a bounded sample or preview extraction for the corridor is available so we can confirm the resolution and granularity match what the substrate consumes.** Funding mechanisms under consideration include an NSF supplement, departmental seed funds, and a paid pilot.

Scheduled call

A scope call is scheduled for Friday afternoon. The agenda items I would like to put on the call up front:

1. Whether StreetLight has a University Research program or academic pricing tier for a single-geography validation dataset.

2. Whether a bounded sample or preview extraction for the Midtown corridor is available before any larger commitment.
3. Available funding mechanisms (NSF supplement, paid pilot, departmental seed).

Supporting materials available on request

- Cross-street capacity model writeup (substrate methodology, ∼30 pages).
- Comprehensive figure walkthrough (∼42 pages, 15 figures with structured discussion).
- Reproducibility documentation: every supply layer in the model traces back to a public NYC Open Data source; build scripts are documented.

Gil Raitses, Researcher, Syracuse University. `gilraitses@gmail.com`*, +1 212.365.0255. Project repository documentation available on request.*