Supplemental Information: Testing Pedestrian Flow and Delay Estimation from First-Actuation Latency

David M. Levinson
School of Civil Engineering, The University of Sydney
david.levinson@sydney.edu.au

SI-1. Estimator Details

Pooled censored likelihood

For restrictive pedestrian cycle i, the observed first-actuation latency is $Y_i = B_i - A_i$, where A_i is flashing-don't-walk onset and B_i is the first subsequent actuation. Under a locally homogeneous Poisson process, the waiting time to the first actuation-producing arrival event is exponential. A no-actuation cycle contributes its full restrictive duration C_i as right-censored exposure. Within case j,

$$\hat{\lambda}_j = \frac{K_j}{\sum_{i\in\mathcal{A}_j} Y_i + \sum_{i\in\mathcal{C}_j} C_i}, \qquad \hat{N}_j = T_j\hat{\lambda}_j. \tag{1}$$

The first actuation supplies the cycle's timing observation. The inferred rate represents later arrival events, while calibration accounts for non-actuation and the number of people carried by each event. A compound-Poisson interpretation allows each event to carry a pedestrian-group mark without changing the exponential first-event time.

Posited frequency and delay method

The posited count method uses the reciprocal of mean time from flashing-don't-walk onset to first actuation as pedestrian-group frequency, then expands arrival events to people. With no empty cycles, $K_j/\sum_i Y_i = 1/\overline{Y}_j$. The reported estimator adds the full exposure of cycles with no actuation to the denominator, giving the censored-likelihood form of the same waiting-time model. This right-censoring prevents the rate from being estimated only from cycles in which a call was observed.

The delay method assigns the first inferred arrival the wait from actuation to Walk; the fitted rate supplies the expected number and accumulated wait of later arrivals. Calibration then converts event-equivalent count and delay to observed-person targets.

Timestamp resolution

The MCAV workbook is effectively one-second data. The pooled estimator remains finite when one latency is recorded as zero because other positive or censored exposure remains in the denominator. A fuller interval-censored likelihood would treat integer latency y as $Y \in [y, y+1)$, with contribution

$$P(y \le Y < y+1) = e^{-\lambda y} - e^{-\lambda(y+1)}. \tag{2}$$

This extension is not used in the reported calibration.

Pooled-rate delay

Let R_i be the time from the first actuation to the following Walk. Conditional on the first event, later arrivals in the remaining interval have expected count $\hat{\lambda}_j R_i$ and mean waiting time $R_i/2$. Expected accumulated delay is

$$\hat{D}_i = R_i + \frac{\hat{\lambda}_j R_i^2}{2}. \tag{3}$$

Arrivals during Walk receive zero signal delay under the compliance scenario.

SI-2. MCAV Normalisation and Audit

The source workbook, `11_Intersections_Data.xlsx`, is the Manual Crossing-Actuation Validation dataset. Observations were collected around midday in August 2022; exact dates and times are unavailable. The retained blocks contain 50,635 seconds (14.07 hours) of second-by-second exposure.

The 11 locations are Burwood Road/Railway Parade; Belmore Street/Burwood Road; Beamish Street/Evaline Street; Church Street/Parkes Street; Queens Road/Park Road; Campbell Street/Marsden Street; Argyle Street/Marsden Street; Argyle Street/O'Connell Street; Burrows Avenue/Gleeson Avenue; Brown Street/Liverpool Road; and Station Road/Rawson Street. The final three contribute two observed directions, producing 14 cases.

The normalizer produces a 492-row cycle table, a 14-row case table, seven candidate-estimator comparisons, leave-one-out predictions, case-bootstrap results, paired count and delay factors, and machine-readable plot data.

Table 1: MCAV behavioral totals

Quantity	Count
Observed pedestrian arrivals	1,744
Observed exposure	50,635 seconds
Workbook actuator events	513
Compliant crossers with actuation	478
Compliant crossers without actuation	1,187
Non-compliant crossers with actuation	20
Non-compliant crossers without actuation	43
Actuation not followed by crossing	14
Observed restrictive cycles	492
Cycles with first observed actuation	422

Two of the 1,744 arrivals lack behavioral classification. The per-second classification therefore contains 1,728 classified crossers and 14 actuations not followed by a crossing. An unused workbook comparison sheet reports 479 and 1,188 rather than 478 and 1,187 for its two compliant categories. The analysis uses the per-second columns because those also supply the estimator inputs. Sixty-six initial rows marked `W (Invalid)` precede usable anchors and are excluded.

SI-3. Candidate Estimators and Cross-Validation

Table 2: MCAV bases and leave-one-case-out pedestrian predictions. The observed total is 1,744 pedestrians.

Model	Base	Base/observed	LOO predicted	LOO MAPE
Recorded actuator events	513.0	29.4%	1,739.1	24.3%
Workbook mean-wait formula	1,183.5	67.9%	1,837.9	30.4%
Pooled first actuation, uncensored only	1,004.7	57.6%	1,740.0	16.9%
Pooled first actuation, censored all time	805.6	46.2%	1,776.7	23.9%
Pooled first actuation, censored restrictive time	745.2	42.7%	1,776.8	24.8%
Cycle plug-in, restrictive exposure	1,067.7	61.2%	1,749.0	15.3%
Cycle plug-in, full exposure	1,164.1	66.7%	1,748.3	14.2%

Table 3: MCAV aggregate-ratio stability. Pooled censored Poisson is the primary estimator; other rows are diagnostics.

Model	Ratio	LOO range	Bootstrap ratio interval	Pooled α interval
Pooled censored Poisson	0.462	0.444–0.478	0.411–0.524	1.91–2.43
Cycle plug-in, restrictive exposure	0.612	0.592–0.622	0.568–0.670	1.49–1.76
Cycle plug-in, full exposure	0.667	0.647–0.677	0.623–0.726	1.38–1.60

The pooled censored model is selected prospectively for cross-dataset robustness. Reciprocal-latency plug-in estimators require a floor because $E[1/Y]$ is infinite for an exponential Y. The selected estimator uses the observed pooled exposure directly and yields similar held-out error in MCAV and video.

For each held-out MCAV case, the unweighted mean alpha is calculated from the other 13 cases and applied to the held-out base. The same procedure is used for the 24 video leg-sessions. A ratio-of-sums alpha fitted to all observations reproduces the pooled total by construction; held-out mean absolute percentage error measures predictive accuracy.

SI-4. MCAV Factor Decomposition and Compliance

The matched pooled-censored count factors have mean 2.202, median 2.240, standard deviation 0.519, interquartile range 1.822–2.565, range 1.351–2.928, and empirical central 95% interval 1.395–2.900. Resampling cases 10,000 times gives 1.938–2.459 for the mean.

The pooled ratio-of-sums separates as

$$2.164757 = 1.440491 \times 1.195339 \times 1.257209. \tag{4}$$

The terms are missing actuation, group arrivals measured as people per occupied one-second interval, and remaining timing/model error. At the observed rates, independent same-second coincidence predicts approximately 1.017 people per occupied second, so roughly 90% of the excess from one to 1.195 is associated with pedestrians arriving in groups.

The replay assuming 100% compliance covers 1,488 restrictive arrivals. Four terminal arrivals lack a following Walk and are assigned zero additional delay. Total compliant delay is 59,782 seconds, or 34.28 seconds over all 1,744 arrivals and 40.18 seconds over covered restrictive arrivals.

Table 4: Paired MCAV calibration factors by intersection-direction case. The delay columns distinguish delay assuming 100% compliance from the observed-compliance scenario.

Intersection-direction case	Count α_N	$\alpha_{D,100}$	$\alpha_{D,\text{obs}}$
Queens Road & Park Road	1.351	1.445	1.381
Argyle Street & O'Connell Street	1.487	1.517	1.483
Brown Street & Liverpool Road (SW to NE)	1.681	1.632	1.573
Station Road & Rawson Street (SW to NE)	1.817	1.854	1.727
Church Street & Parkes Street	1.836	1.886	1.886
Campbell Street & Marsden Street	2.021	1.925	1.863
Burrows Avenue & Gleeson Avenue (NE to SW)	2.107	2.321	2.243
Argyle Street & Marsden Street	2.374	2.096	2.008
Station Road & Rawson Street (NE to SW)	2.444	2.511	2.368
Belmore Street & Burwood Road	2.551	2.396	2.348
Beamish Street & Evaline Street	2.570	2.561	2.359
Brown Street & Liverpool Road (NE to SW)	2.823	2.803	2.729
Burwood Road & Railway Parade	2.844	2.577	2.542
Burrows Avenue & Gleeson Avenue (SW to NE)	2.928	2.942	2.853
Unweighted mean	2.202	2.176	2.097

Against 28,023.35 unexpanded model seconds, $\alpha_{D,100}$ is 2.133 by ratio of totals and 2.176 as the unweighted mean of 14 case factors. The case factors range from 1.445 to 2.942; their empirical central 95% interval is 1.468–2.897 and bootstrap interval for the mean is 1.935–2.414.

The workbook records compliance classes as case or cycle totals rather than person-linked crossing-start times. The observed-compliance scenario therefore multiplies each case's D_{100} by its observed compliant share. It assigns classified noncompliers the same counterfactual wait as other arrivals and zero signal delay when crossing. This gives 57,586.88 seconds (15.996 hours), 33.02 seconds over all pedestrians, $\overline{\alpha}_{D,\text{obs}} = 2.097$, and a pooled ratio of 2.055. The scenario is 0.610 pedestrian-hours (3.7%) below the delay assuming 100% compliance.

SI-5. Video-Coded Peer Reconstruction

The source is associated with Choy and Levinson [2026]. The primary peer analysis uses an author-supplied upstream package dated 20 July 2026. Each of six session folders contains pedestrian-level `PED.csv`, four directional `SIGNAL.csv` files, and a processed single-stage table. The 18 retained files are checksum-verified.

Codes such as `3AM` are site-period identifiers: sites 3, 4, and 5 are Redfern Street/Pitt Street, Campbell Street/Riley Street, and William Henry Street/Harris Street; AM and PM denote 08:00–09:00 and 12:00–13:00. One pedestrian has no physical leg and two mapped pedestrians arrive before the first known signal state. Exact replay otherwise gives 372 Walk, 367 flashing-red, and 1,262 red arrivals.

The field `Accuator` records binary actuator use; a separate press timestamp is absent. A positive observation is placed at arrival. An eligible user arrives during flashing or solid red.

The six behavioral factors have mean 2.642 and range 2.135–3.362. Their pooled factor decomposes as $2.640 = 1.228 \times 2.149$, where 1.228 is all known arrivals divided by restrictive arrivals and 2.149 is restrictive arrivals divided by eligible actuator users. This behavioral identity describes coverage by signal state and actuator-user status. It complements the estimator-matched decomposition below, which instead traces the pooled-censored model base.

Table 5: Video complete-event behavioral expansion factors

Site, time	Observed	Known	Restrictive	Eligible	α_v
Redfern/Pitt, 08–09	442	442	358	151	2.927
Redfern/Pitt, 12–13	354	353	279	105	3.362
Campbell/Riley, 08–09	317	317	258	130	2.438
Campbell/Riley, 12–13	366	364	292	133	2.737
William Henry/Harris, 08–09	288	288	240	128	2.250
William Henry/Harris, 12–13	237	237	202	111	2.135
All sessions	2,004	2,001	1,629	758	2.640

The four directional signal streams expose all restrictive opportunities, including empty cycles. The earliest restrictive actuator user is the first event under the press-at-arrival assumption.

Table 6: Estimator-matched count decomposition. The overall adjustment is the product of the final three factor columns.

Data	Observed base	Universal base	Occupied s	People	Missing actuation	Group arrivals	Remaining	α_N
MCAV	805.6	1,160.5	1,459	1,744	1.440	1.195	1.257	2.165
Video	937.3	1,584.0	1,701	2,003	1.690	1.178	1.074	2.137

For each video leg-session, the universal-actuation replay moves the first event to the first observed restrictive-phase arrival and retains a no-arrival cycle as censored exposure through the next Walk. Across 24 cases, 842 of 1,462 restrictive cycles contain an observed arrival. The universal-actuation base is 1,584.0 events, compared with the observed-actuation base of 937.3. The 2,003 mapped pedestrians occupy 1,701 distinct one-second arrival bins. Thus the exact video identity is

$$2.136936 = 1.689921 \times 1.177543 \times 1.073862. \tag{5}$$

The second factor, people per occupied second, is close to MCAV's and primarily represents group arrivals. A larger video missing-actuation factor is balanced by a smaller remaining factor. The similar overall count adjustments therefore do not imply identical component errors. Because the datasets observe different sites and signal regimes as well as using different collection methods, the component contrast cannot identify a collection-method effect. Differences in site demand, cycle structure, restrictive-time share, and actuation practice may explain as much or more.

Table 7: Video pooled-censored peer estimates by session.

Session	Observed	Opportunities	First actuation	Base	α_N
Redfern/Pitt, 08–09	442	262	117	201.8	2.190
Redfern/Pitt, 12–13	353	278	86	117.6	3.002
Campbell/Riley, 08–09	317	315	107	149.1	2.126
Campbell/Riley, 12–13	366	307	117	162.3	2.255
William Henry/Harris, 08–09	288	152	96	169.4	1.700
William Henry/Harris, 12–13	237	148	84	137.1	1.729
All sessions	2,003	1,462	607	937.3	2.137

The primary video calibration statistic is the unweighted mean across the 24 leg-session factors.

Across the 24 leg-sessions, count alpha has mean 2.203, median 2.284, range 1.422–3.213, and ratio of totals 2.137. Leave-one-leg-session-out predictions total 2,067.4 and have 22.4% MAPE.

Table 8: Video estimator-matched factor distributions across 24 leg-sessions. The video factors provide an independent peer test; the 14 MCAV cases remain the transfer calibration.

Factor	Cases	Mean	Median	IQR	Range	Totals ratio
Count α_N	24	2.203	2.284	1.727–2.504	1.422–3.213	2.137
Delay $\alpha_{D,100}$	24	2.233	2.223	1.691–2.565	1.156–3.751	1.952
Delay $\alpha_{D,\text{obs}}$	24	1.396	1.479	1.132–1.589	0.870–2.044	1.286

Video count alpha correlates −0.542 with restrictive-time share, descriptively; site, period, and signal timing are confounded in this small sample.

Signal-only delay assuming 100% compliance totals 12.934 hours versus 6.626 modeled hours. Across 24 cases, $\alpha_{D,100}$ has mean 2.233, range 1.156–3.751, and pooled ratio 1.952. Recorded restrictive waiting totals 8.522 hours, giving mean $\alpha_{D,\text{obs}} = 1.396$ and pooled ratio 1.286. The recorded result is 4.412 hours (34.1%) below the signal-only compliance replay. It includes observed post-Walk startup time, while D_{100} stops at Walk onset.

SI-6. Reproducibility and Output Inventory

The primary generated MCAV files are:

- `mcav_cycle_level_validation.csv`;
- `mcav_case_model_validation.csv`;
- `mcav_model_comparison_summary.csv`;
- `mcav_model_stability_summary.csv`;
- `mcav_model_leave_one_out.csv`;
- `mcav_compliance_scenario_by_case.csv`; and
- `mcav_compliance_scenario_summary.csv`.

The primary video files are:

- `video_upstream_peer_cases.csv`;
- `video_upstream_peer_cycles.csv`;
- `video_upstream_peer_arrivals.csv`;
- `video_upstream_peer_sessions.csv`;
- `video_upstream_peer_summary.csv`; and
- `video_upstream_source_manifest.csv`.

The normalizer is `broad_results/validation/mcav_validation/scripts/normalize_mcav_validation.py`. The video reconstruction is in the video validation scripts folder. Cross-method prediction is reproduced by `broad_results/validation/scripts/compare_actuator_count_accuracy.py`.

SI-7. Relation to Existing Push-Button Count Work

Table 9: Comparison with empirical pedestrian push-button volume studies

Study	Data and signal measure	Reported conversion or performance
Li and Wu [2021]	Two Arizona midblock crossings; Poisson arrivals plus button and signal timing events	Mean absolute error 2.27 and 1.78 pedestrians/hour; no single adjustment factor
Singleton and Runa [2021]	90 Utah intersections; nonlinear models of controller activity	Prediction–observation correlation 0.84; mean absolute error 3.0 pedestrians/hour
Kothuri et al. [2024], Vahedi Saheli et al. [2026]	65 Oregon collection sites, 64 in the article analysis; 15-second-filtered presses	Quadratic empirical model; correlation 0.86; mean absolute error 2.4 crosswalk users/hour
Present MCAV	14 cases; theory-based censored first-actuation-latency estimator	$\alpha_N = 2.165$; leave-one-case-out mean absolute percentage error 23.9%
Present video	24 leg-sessions; same theory-based estimator	$\alpha_N = 2.137$; leave-one-case-out mean absolute percentage error 22.4%

The closest published estimator in structure is Li and Wu [2021]: pedestrian arrivals are Poisson and signal-controller button and timing events supply the observations. It was developed for one- and two-stage button-activated midblock crossings and explicitly reconstructs missing signal cycles. The present estimator instead uses flashing-don't-walk-to-first-actuation latency, right-censors no-actuation cycles, extends the rate over total exposure, and calibrates both count and delay.

The multi-site Utah and Oregon studies fit empirical nonlinear relationships rather than a single expansion factor [Singleton and Runa, 2021, Kothuri et al., 2024, Vahedi Saheli et al., 2026]. The present method differs by deriving its base estimate from Poisson waiting-time theory, then empirically calibrating the remaining difference between arrival events and people. Their reported error metrics also differ from the present leave-one-case-out mean absolute percentage error. The common finding is that site context matters and a universal people-per-call factor is not transferable without qualification.

For delay, Karimpour et al. [2022] used high-resolution smart-sensor events and a finite-mixture model at four Arizona intersections. The reported mean absolute and root-mean-square errors were 10 and 13 seconds, respectively. That study validates delay directly and does not report a multiplicative adjustment comparable with $\alpha_{D,100}$ or $\alpha_{D,\text{obs}}$. The present study therefore adds a first-actuation-based delay adjustment rather than replicating an existing published alpha.

References

Keng Io Choy and David M. Levinson. Pedestrian non-compliance at signalised intersections in sydney. Manuscript and public analysis archive, 2026. URL https://github.com/chyy512/PedestrianNoncomplianceAnalysis.

Abolfazl Karimpour, Jason C. Anderson, Sirisha Kothuri, and Yao-Jan Wu. Estimating pedestrian delay at signalized intersections using high-resolution event-based data: A finite mixture modeling method. *Journal of Intelligent Transportation Systems*, 26(5):511–528, 2022. doi: 10.1080/15472450.2021.1926246.

Sirisha Kothuri, Patrick Singleton, Mahyar Vahedi Saheli, Elizabeth Yates, and Joseph Broach. Active transportation counts from existing on-street signal and detection infrastructure. Technical Report FHWA-OR-RD-24-03, Oregon Department of Transportation, February 2024. URL https://www.oregon.gov/odot/Programs/ResearchDocuments/SPR857_ACTIVE-TRANSPORTATION-COUNTS.pdf.

Xiaofeng Li and Yao-Jan Wu. Real-time estimation of pedestrian volume at button-activated midblock crosswalks using traffic controller event-based data. *Transportation Research Part C: Emerging Technologies*, 122:102876, 2021. doi: 10.1016/j.trc.2020.102876.

Patrick A. Singleton and Ferdousy Runa. Pedestrian traffic signal data accurately estimates pedestrian crossing volumes. *Transportation Research Record*, 2675(6):429–440, 2021. doi: 10.1177/0361198121994126.

Mahyar Vahedi Saheli, Elizabeth O'Neill Yates, Patrick Singleton, Sirisha Kothuri, and Joseph Broach. Pedestrian volumes from push-button traffic signal data in oregon: Estimating models and assessing model transferability. *Journal of Transportation Engineering, Part A: Systems*, 152(9), 2026. doi: 10.1061/JTEPBS.TEENG-9266.