capstract:2611.00007v1 [q-bio.QM] 4 Nov 2026

Reconciliation of endpoint counts in partially observed root systems

Mirela V. Dobreanu[1,*] Renan Costa de Vilar[2] Nadège Ilunga-Séroux[3]
Joon-ho Baek-Mori[1,4]
[1]Department of Quantitative Phenotyping, Târnava Institute of Plant Systems, Cluj-Napoca, Romania
[2]Center for Root Observation, Instituto da Serra Oriental, Curitiba, Brazil
[3]Department of Biometry, Institut des Savanes de la Lukaya, Kisantu, Democratic Republic of the Congo
[4]Division of Applied Optimization, Myeongjin Institute of Mathematics, Daejeon, Republic of Korea

November 4, 2026

Abstract

Counts of distal root endpoints from serial images fluctuate when endpoints enter and leave occluded parts of the observation window. We frame endpoint enumeration as a reconciliation problem and represent persistence and branching on a capacitated temporal graph. Each inferred movement produces two entries, so that changes in visibility can be reported without changing the inventory of roots. We evaluated the estimator on 1,440 held-out frames from 120 seedlings, with apertures specified independently of the reference counts. At 20% aperture loss, mean absolute error was 0.52 endpoints per image, compared with 1.56 for framewise counting and 0.95 for nearest-neighbor linking. Conservation gave the largest improvement when a previously recorded root entered an occlusion and later reappeared. Counts remained weakly identified when branches arose and terminated wholly within the same region. In a separate record perturbation experiment, duplicated complete histories passed arithmetic reconciliation and required corroboration against the original image evidence. We describe how to report those cases without silently propagating an opening inventory error.

Keywords: root phenotyping, network flow, occlusion, longitudinal measurement, double-entry bookkeeping

1 Introduction

A count of root tips is often an input to a developmental analysis. Tips that turn away from a plane, move behind fixtures or enter a region of poor contrast are absent from a count of the photograph, and may return in the next image without having been created in the interval. Root analysis software can trace a root path through brief gaps and describe its relationship to other roots (Lobet et al., 2011; Pound et al., 2013; Yasrab et al., 2019). It is harder to decide whether the endpoint should still be included after a section through which it might have moved becomes

∗Corresponding author: m.dobreanu@tarnava-phenotyping.example

unobservable. The distinction matters for phenotypes based on the timing of lateral emergence. Dubrovsky and Forde (2012) describe reasons to discount a change of elongation as evidence of a change in the number of developmental units.

A comparable problem arises when the object of an inventory is in transit on the date the books are closed. A physical count at the receiving depot misses it. A count at both depots may record it again after it arrives. In double-entry bookkeeping a transfer out of one account is paired with a transfer into another. In the interim, the asset may belong in an account without a physical counterparty at either depot (Yamey, 1949; Borsell and Afonso-Veira, 2020). The record gives the auditor two distinct questions to ask. Do the reported movements explain the closing balance? Is there enough evidence to support the movements used to explain it?

We apply these questions to a series of root images. An existing tip can move between visible and concealed locations. A new lateral has an entry that increases the system count, while movement through the aperture does not. A simple sum of observed endpoints does not preserve this distinction. Some lineage methods already link detections with time-dependent constraints (Jaqaman et al., 2008; Zhang et al., 2008). Our concern is the resulting inventory and the trail of evidence needed to amend it after the fact. We keep the opening balance visible, preserve every correcting entry and declare circumstances in which contemporaneous imaging cannot support a count.

2 Methods

2.1 The unit of account

The measurement target at frame t is the number N_t of intact distal root-axis endpoints located in a fixed planar observation window. A lateral enters the count when it has a distinguishable free axis endpoint separated from the main axis by at least 0.25 mm along the lateral. Root hairs and pre-emergent primordia are excluded. A stopped endpoint still belongs in the count. A cut or detached section is not an intact axis. On a branch, the existing parent tip continues to exist and the lateral adds one. This last convention differs from tracking problems in which one parent particle is replaced by two daughters (Alyun and Serebriakov, 2021). It fixes what constitutes a transaction before any estimates are made.

The opening inventory was taken from an unmasked image obtained one hour before the first scored frame. In the primary validation two analysts supplied its endpoint coordinates after reconciliation; the coordinates and count went to all methods allowed to use observations before frame one. The experiment therefore tests closure conditional on an externally established opening balance. Appendix B examines sensitivity to this condition. We record visible count V_t and hidden count H_t separately, and $N_t = V_t + H_t$. A tip intersecting the window edge belongs to the count if its distal point meets the coordinate convention in Section 2.3, regardless of how much of its axis remains visible.

2.2 Posting and closure

Write B_t for lateral births, I_t for entries across the field boundary, and R_t for exits or corroborated loss of intact endpoints during the interval ending at frame t. Each quantity is nonnegative. The close is the recursion

$$N_t = N_{t-1} + B_t + I_t - R_t. \tag{1}$$

Occlusion cannot change N_t. Let D_t and U_t count transfers from visible to hidden and from hidden to visible. Split B_t, I_t and R_t by the corresponding location at posting, using subscripts v and h. The subsidiary accounts obey

$$\begin{aligned} V_t &= V_{t-1} + B_{v,t} + I_{v,t} - R_{v,t} - D_t + U_t, \\ H_t &= H_{t-1} + B_{h,t} + I_{h,t} - R_{h,t} + D_t - U_t. \end{aligned} \tag{2}$$

The signs are determined by destination and origin. Adding the lines in Eq. (2) gives Eq. (1); assigning disappearance of the image to R_t at first loss of visibility would not. Postings associated with openings, births, entries, exits and final settlements have an external account, so they too have two sides. It is possible for the hidden account to have a positive closing balance after the last recorded frame. The experiment has ended, and nothing further is implied about the root.

We construct one split node per candidate tip detection and per masked cell at each time, with an occupancy arc joining the two halves of each node. Node capacities are one for detections and large enough for the candidate set for mask cells. The spatial cells are 1 mm squares on the registered field. Arcs join positions at successive frames when mutually admissible under a limit of 2.8 mm axial travel per imaging interval. Boundary arcs terminate a trajectory only near the fixed field edge or at an annotated break motif. Birth arcs start at candidate points on prior-frame axes, including axis segments projected into the mask from admissible boundary intersections. Within a fully unseen patch we include a more expensive birth arc for each masked cell that a projected segment could reach. We allow motion in either direction relative to a mask edge. Direction of travel along a root axis, when estimable, enters as a cost. The correspondence can be wrong if a buried axis turns through a high curvature, and small costs do not prove anatomical continuity (Kertész-Papp and Rehagen, 2018).

The initial inventory has fixed occupied nodes. The end of the last frame closes every active trajectory to a settlement node, separately from biological exits. An observed candidate can be omitted from the solution. Occupying a candidate with estimated detection probability p adds the cost $\log((1-p)/p)$ relative to omitting it. For hidden cells the occupancy cost comes from a fitted, nonnegative penalty per tip per frame in the absence of an observed endpoint. We used linear costs and made the construction independent of past assignments to the same cell. Transition costs use the empirical distributions of displacement and angle fitted in development images. Opening coordinates carry no detection cost. Start and exit fees depend on the kind of observed evidence attached to the arc. These terms are a scoring rule for consistent explanations of the images, with no claim that the fitted costs yield calibrated posterior probabilities.

For a directed graph with incidence matrix M, arc capacities u, opening and closing demands b, and cost vector c, we solve

$$\hat{f} \in \arg\min_{f} c^{\top} f, \qquad M f = b, \qquad 0 \le f \le u. \tag{3}$$

Opening nodes each supply one unit to the graph. A common reservoir has the corresponding net demand. Additional trajectories start with an entry or birth arc from the reservoir and return on an exit or settlement arc. The occupancy arcs report the inventory, with every accepted transition written once at each end. Integer capacities and net supplies admit an integer optimizer by total unimodularity (Ahuja et al., 1993). We retain the objective and both postings. For repeatability, ties are resolved in a deterministic arc order by frame and position. Tied solutions can have different counts.

When evaluating a suspect transaction we forbid that arc and resolve Eq. (3). We also force it in a second solve if it was initially absent. These solves exchange one bound at a time and do not require a different graph. The cost gap expresses support relative to the competing explanation. A low gap may identify a motion ambiguity that has no effect on N_t. Therefore the review record includes both the cost gap and the maximum count difference between the two flow solutions. Only the first item is in objective units. We distinguish a count discrepancy from a choice of path since a plant can have a fully reconciled count and two plausible anatomical pedigrees (Öhlund Varga and Neves-Arouca, 2022).

2.3 Specimens, images and masks

The corpus comprised 240 planar growth series with one seedling per observation window. Each of the three contributing sites supplied 40 garden pea and 40 garden cress seedlings. Within each site-by-species cell, 10 plants went to cost fitting, 10 to interval calibration, and 20 to the final test. Assignments were made by seedling before annotations were released. Thus fitting, calibration and testing had 60, 60 and 120 independent plant units. Each series had a reference opening image and 12 hourly analysis frames. The field was the same 100 mm by 140 mm rectangle after registration to dotted perimeter fiducials. The left and top borders were included and the right and bottom excluded. Changing which half-open rectangle was chosen moved a single reference endpoint in the held-out series.

Two annotators traced one clean image sequence independently, without seeing each other's work. The production reference used joint adjudication of discrepant endpoint identities after inspection of the high-resolution image sequence. A third annotator gave the deciding count on 19 held-out frames after the first adjudication pass. We used a registration tolerance of 0.12 mm to match the traces. Tip identity relied on continuity to the same proximal axis in the clean series. The reference includes both location and timing but can still miss primordia below the length cutoff, especially at a tight bifurcation. We export these traces as Root System Markup Language records (Lobet et al., 2015). The boundaries of the observation window and the properties of the ground truth are part of the benchmark definition.

The estimation images were resampled to 0.10 mm per pixel and then overlaid with one of four aperture treatments: 0, 10, 20 or 35% lost field area. For each positive loss setting we randomized mask shape in balanced groups of 40 plants: a strip, two strips and a cluster of blocks. Shapes were balanced within site-by-species strata to within one plant. The placement generator did not use any plant image. It was frozen before inspecting the held-out series. A mask changed position at a prescribed frame boundary in frames 5 and 9, allowing both entry and reappearance observations. When a moving mask covered a stationary endpoint, the posting was still a transfer to hidden status. The count definition uses the fixed field, and only the observation mask moves. The reference labels were generated from images taken before downsampling and masking.

We extracted all models' endpoint candidates from the same image arrays using a trained pixel segmentation model with a held-out threshold for visible tip proposals (Ronneberger et al., 2015). Unmasked aperture geometry was supplied to every estimator. No method observed annotations from the analyzed frames. In a separate physical occlusion subset, 24 additional pea series were imaged with their support fixtures in place and checked at close against high-resolution photographs taken after moving a support plate. That subset was planned for qualitative error characterization. It did not contribute to fit, calibration, or the primary holdout comparisons. A photograph taken after moving a plate can displace a tip. We compare to the adjudicated total only when the before-and-after registration error stays under the annotation tolerance.

2.4 Comparators, estimands and intervals

Framewise counting uses the detections judged to be tips at a threshold chosen in development series. Nearest-neighbor linking with a two-frame gap permits an existing track to be maintained without a current detection. We solve an assignment at each step using distance and a fixed penalty for longer gaps (Kuhn, 1955). An exponentially smoothed count is also included. Its weight minimizes the same development loss. The latter uses no anatomical prior, and rounds to the nearest count only when summarizing errors. None of these comparators uses adjudicated post-opening counts. All comparisons are with our common detector, so a literature system's result under a different detector cannot be read directly from Table 1. The root image extraction studies (Pound et al., 2013; Yasrab et al., 2019) motivate the shared candidate representation.

For a fixed aperture fraction, absolute error L_i is averaged over the 12 analyzed frames of each plant and then equally over test plants. We report errors in tips per image, and bootstrap plants within site and species (5,000 draws). Differences use the same plant draws across estimators. We include signed error to detect systematic undercounting, since a constant positive correction can improve absolute error on an imbalanced corpus and still give biased differences between subgroups (Bland and Altman, 1986). We estimated signed error on the very same frames as absolute error.

The 20% analysis also reports a simultaneous prediction interval for the 12 counts of a held-out plant. The score Q_i is its maximum framewise absolute count error, and the radius q is a calibration rank within each site-by-species stratum (Lei et al., 2018). In one stratum there are 10 calibration plants, so for nominal 90% we take the highest of the 10 values, the $lceil(10+1)0.90ceil$th rank. These bands have at least 90% marginal coverage per new series exchangeable with the calibration series within stratum, conditional on the fitted estimator. The band intersects the nonnegative integers. Its coverage is simultaneous across the imaged window and marginal over plants within the stratum. It does not identify which inferred tips are real or offer a guarantee at a new site.

3 Results

3.1 Estimation under controlled occlusion

Across the 1,440 images at 20% loss, the baseline detector undercounted by an average of 1.29 tips. The mean absolute error was 1.56. At that same aperture the gap linker brought mean absolute error down to 0.95. Reconciliation yielded 0.52, which is a paired reduction of 0.43 tips (95% CI 0.34–0.52) relative to linking. The per-site reductions were in the same direction: among the 40 plants per site, total absolute errors averaged 0.48, 0.55 and 0.53 for reconciliation and 0.90, 1.00 and 0.95 for linking. The estimates weight sites by equal plant counts. Table 1 shows the signed error and fraction of frames exactly recovered. Coverage did not enter model selection.

The framewise method had the lowest mean absolute error in completely visible images at 0.23, compared to 0.25 for reconciliation (Figure 1). Most frames offered little scope for a correction when there were no blind areas. There were some mistakes in the inferred trajectory, usually where a spur below the reference length threshold appeared twice in adjacent frames. The flow treated joint persistence as evidence in favor of a tip; in these cases the stricter framewise threshold gave the right count. Figure 2 compares mask forms and finds that the difference across masks grew as the lost area increased. For the most extensive strips, a lateral could enter blind space at one edge and leave through the opposite edge well after the ordinary two-frame linking interval. In the presence of a compact block at the same area, many tracks departed along the

edge they used to enter, with a shorter invisible duration. Mean placement loss alone is insufficient to characterize either pattern.

Table 1: Count recovery on 120 test plants under 20% aperture loss. Intervals for mean absolute error are stratified plant-bootstrap intervals.

Estimator	Mean absolute error (tips)	Mean signed error (tips)	Exact frames (%)
Framewise threshold	1.56 (1.39–1.73)	−1.29	24.4
Exponentially smoothed count	1.31 (1.17–1.45)	−1.13	31.7
Nearest-neighbor gap linking	0.95 (0.84–1.06)	−0.76	43.2
Reconciled inventory	0.52 (0.45–0.59)	−0.23	61.5

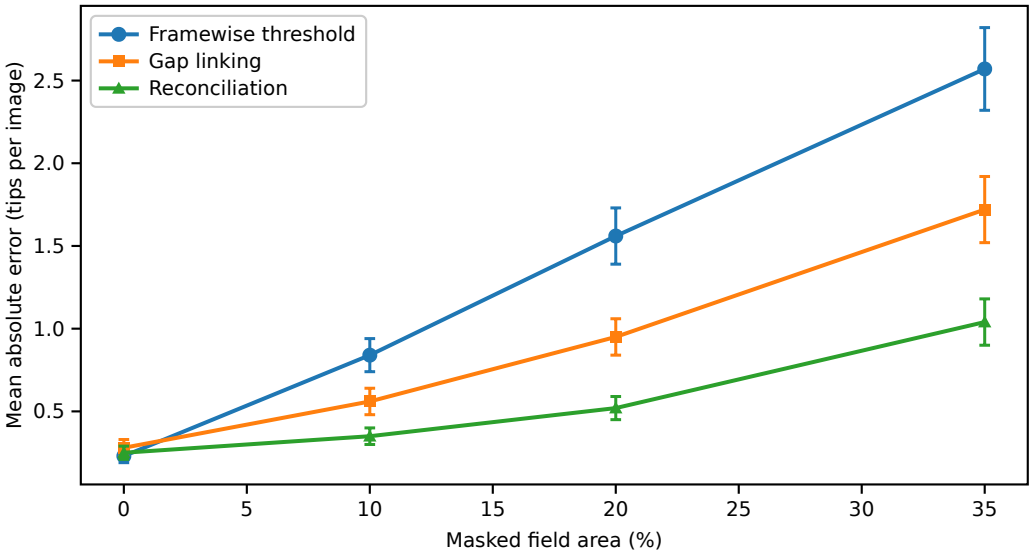


Figure 1: Absolute count error in all 120 test plants at each aperture loss. Points are means over plant-level frame averages, with pointwise 95% bootstrap intervals.

For 111 of the 120 test plants, all 12 reference counts were inside the calibrated interval. There were 37, 39 and 35 covered plants by site respectively. The median interval width across frames was four tips, expressed as the numerical upper endpoint minus lower endpoint. Eight of the nine failures involved undercounts at the edge of the fixed window or under a wide strip. Three were found in the same species-by-site stratum. Since there were only 20 held-out plants in any stratum we do not estimate subgroup coverage to more precision than those counts warrant. The nine failed plants contained 14 missed frame counts altogether.

We used the separate plate series to examine what the fitted arc costs treated as an evidentiary defect. A reliable match between pre- and post-movement images was recorded for 19 of the 24 series. In 13 of the 19, the final estimated count equalled the manually adjudicated count. The remaining estimates were four undercounts and two overcounts, each by one tip. In one of the latter cases, successive captures on opposite sides of a beveled fixture made the same tip appear eligible for two origins. Both proposed routes entered the same mask bank at approximately the same angle. The branch account could have backed either, though there was no image showing

both roots simultaneously. Section 3.2 gives an isolated test of what in the ledger could detect a repeated use of the same evidence.

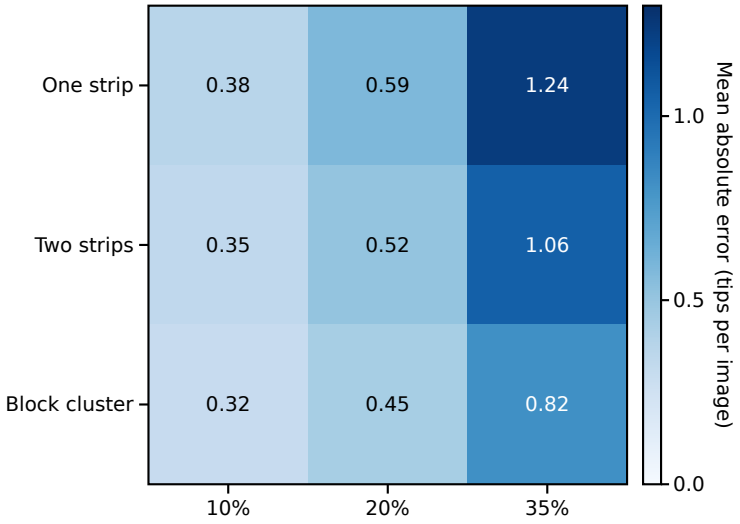


Figure 2: Mean absolute errors for the reconciled inventory by occlusion shape. The three shapes were assigned to 40 test plants each at each loss level. Color limits are shared across the nine cells.

3.2 Exceptions in the record

We tested the record against two candidate defects in Table 2. First, take a regular withdrawal from visibility into occlusion and omit just its destination posting. Trial balance catches the omission because the sum of the signed entries for the transaction does not vanish. The check does not prove that either coordinate on an otherwise double-posted transfer is correct. To demonstrate the difference, we constructed a second error. A complete history observed to arise after opening was copied with a new root identifier, a new reference to the same initial branch image, and both sides of every posting. This can produce a balanced false result throughout the period. It is the same problem as counting an invoice twice after it has been given a second reference number (Kaim-Soler and Adebisi, 2017).

For this check alone, we prepared 90 export files from distinct held-out plants with at least one image-supported emergence and one confirmed visible-to-hidden transfer. We assigned 30 to each test condition after selection. Each file used the 20% mask and the fitted graph. We audited the file as exported. We did not feed the altered files to the estimator or change the plant-level validation metrics. One-sided omissions were all caught by the paired transaction test. Duplicated histories never failed that test. Image-evidence reuse caught 28 of the copies; in two cases the injection linked to a second, overlapping proposal for the same tip. The retained segmentation file still supported both proposals, so its evidence was insufficient to settle the outstanding identity.

Table 2: Checks on quarantined exports. Each column reports the number of exports flagged out of 30. A file can be flagged by both checks.

Check	Unchanged	One-sided omission	Duplicated history
Transaction balances	0	30	0
No repeated exclusive evidence	1	0	28
Either check	1	30	28

For the repeated-evidence audit we started with primary-image proposal indices and compacted indices linked by their optical match group. The group definitions were chosen before the audit. Each visible endpoint proposes at most one live trajectory. A detection certificate assigned exclusively to more than one trajectory in one frame triggers the check. The one flagged unchanged record contained two candidates at a close emergence, members of the same group, which the high-resolution reference traced to two separate axes. We therefore report a flagged frame as an exception, pending inspection of its evidence, even if it balances. We included both defective forms because a validation set with dropped entries would have been unusually favorable to our cover sheet.

3.3 Exposure to a change in accounting policy

Leave out the projected-axis support for births under cover, but preserve all other costs, capacities and postings. The error then depends on how much of the hidden time series begins with a root no person has seen. At 20% mask loss, limiting births to directly visible evidence increases holdout mean absolute error from 0.52 to 0.68. The increase is 0.16 tips (95% CI 0.09–0.23). Letting disappearance carry an unrestricted retirement arc instead yields 0.71. This second change drives bias more negative, from −0.23 to −0.43 tips per frame. Figure 3 shows signed error for the same solutions split by site. A permission to write off a tip, used as an aid to computation, has a quantitative effect on what the record says was present.

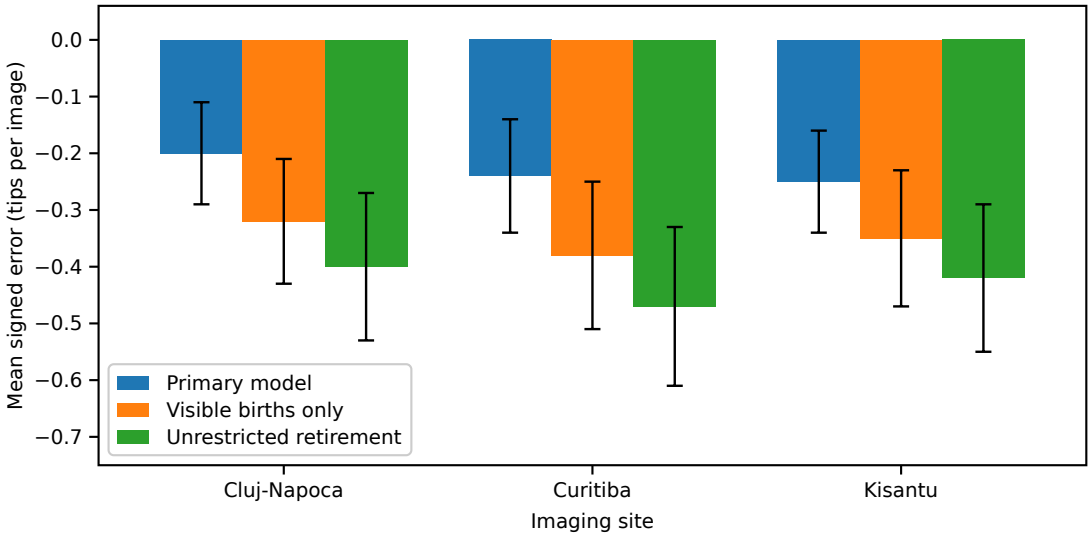


Figure 3: Mean signed error when selected recording constraints are changed. Bars show mean counts minus reference counts, with pointwise 95% bootstrap intervals over plants at each site.

4 Discussion

An optical observation and the posting that records it need different identifiers. When an analyst corrects the optical identity, the counterpart postings may all have to be reversed, including on frames where nobody had seen a tip.

Our primary experiments condition on a closed opening count and designed aperture patterns. A natural occlusion can depend on the plant. Roots tend to collect near support edges, and two paths passing under the same opaque component can occlude each other after leaving it. When growth rate drives mask exposure, the probability of reappearance may be smaller for the phenotype of interest. We have not estimated error for unresolved trajectories via any missing-at-random assumption (Rubin, 1976). The count band can absorb some estimation error in a population undergoing the same observation process as calibration. A different pattern calls for a different calibration series. It may also call for changing the target from individual identities to bounds on a total.

The table reviewer could detect an unmatched debit without inspecting morphology. A balanced claim to a second root required independent evidence. The exported record therefore retains image links for accepted postings as well as exceptions. Appendix A gives a case where the reviewer cannot establish an accurate count even with a correct opening and a fully visible close.

Data and code availability

The benchmark bundles contain the scoring mask bitmaps, proposed detection coordinates, reference counts and adjudicated traces. Derived benchmark materials and analysis code are available from the corresponding author. The export includes an adjustment log with evidence identifiers. Construction, solution and export at 20% mask loss took 0.62 s per 12-frame series (median over the 120 holdout plants on one CPU core, excluding the shared detector). Median times for optional per-transaction review were 0.04 s per forced or forbidden solve. The build includes the deterministic capacity and cost tables.

A Closure and identifiability

Subtracting $V_{t-1} + H_{t-1}$ from $V_t + H_t$ in Eq. (2) removes both D_t and U_t. Summing across intervals gives

$$N_t - N_0 = \sum_{j=1}^{t} (B_j + I_j - R_j). \tag{4}$$

Thus the complete set of external postings identifies the count trajectory without identifying any particular internal transfer. The converse fails. Suppose one potential lateral emerges in the covered part of the field and is later lost after a corroborated disturbance that occurred while the field was still covered. It has no effect on N_0 or the terminal count. It adds a unit to B_s and R_u, at different times $s < u$. Inserting those postings and the matching hidden occupancies leaves Eq. (2) balanced and increases N_t by one for $s \le t < u$. Balancing neither chooses the right history nor identifies whether a reported birth has support. Even a fully visible close at the final frame cannot resolve that combination of internal counts. This counterexample uses root loss as

allowed by the measurement definition. In a series that excludes loss, the same ambiguity is possible if the lateral crosses the observation-window boundary under cover and is never seen.

B An error in the opening inventory

In 40 held-out series we omitted a recognized lateral endpoint from the reported openings. The root was chosen as the opening lateral with the smallest reference index. This simulation concerns reporting of the opening and leaves the observed reference images unchanged. With all later external postings held fixed, Eq. (4) makes the discrepancy exactly one tip on every subsequent line of the total. Eight test series had independent direct evidence of the missing lateral in frame one. When allowed to re-estimate the subsequent flow at fixed opening count, the estimator posted six of those tips as new laterals. Reinstating the original identifier requires changing the opening count and reversing the estimated new birth where it was posted. Counting the recovered root only in the current image corrects the final balance while assigning the wrong emergence interval. We retain both the initial and amended records so that users analyzing birth times can apply a different choice from users analyzing the closing count. The corrected opening record is part of any amended analysis output.

Acknowledgments

We thank Valéria Grinhauz, Merveille Kabinda-Loy and the imaging technicians who repeated the perimeter checks after moving the support plates. The Monte Claro Scientific Instruments Trust provided access to two imaging stations.

References

Ravindra K. Ahuja, Thomas L. Magnanti, and James B. Orlin. *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993.

Dilara Alyun and Leonid Serebriakov. Distinguishing lineage branching from tip replacement in planar growth series. *Studies in Quantitative Organography*, 8:166–184, 2021.

J. Martin Bland and Douglas G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327:307–310, 1986.

Annetta Borsell and Célio Afonso-Veira. Inventory in transit and the timing of independent physical counts. *Journal of Inventory Assurance*, 17:21–39, 2020.

Joseph G. Dubrovsky and Brian G. Forde. Quantitative analysis of lateral root development: pitfalls and how to avoid them. *The Plant Cell*, 24:4–14, 2012.

Khuloud Jaqaman, Dinah Loerke, Marcel Mettlen, Hirotaka Kuwata, Sergio Grinstein, Sandra L. Schmid, and Gaudenz Danuser. Robust single-particle tracking in live-cell time-lapse sequences. *Nature Methods*, 5:695–702, 2008.

Bruna Kaim-Soler and Tomilayo Adebisi. Independence of documentary evidence in serial reconciliations. *Review of Applied Assurance*, 6:89–105, 2017.

Zsuzsa Kertész-Papp and Nils Rehagen. Continuity of axis identity through partial optical obstruction. *Journal of Phenotypic Measurement*, 12:342–359, 2018.

Harold W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2:83–97, 1955.

Jing Lei, Max G'Sell, Alessandro Rinaldo, Ryan J. Tibshirani, and Larry Wasserman. Distribution-free predictive inference for regression. *Journal of the American Statistical Association*, 113:1094–1111, 2018.

Guillaume Lobet, Loïc Pagès, and Xavier Draye. A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant Physiology*, 157:29–39, 2011.

Guillaume Lobet, Michael P. Pound, Julien Diener, Christophe Pradal, Xavier Draye, Christophe Godin, Mathieu Javaux, Daniel Leitner, Félicien Meunier, Philippe Nacry, Tony P. Pridmore, and Andrea Schnepf. Root System Markup Language: toward a unified root architecture description language. *Plant Physiology*, 167:617–627, 2015.

Emil Öhlund Varga and Lina Neves-Arouca. Aggregate identifiability in repeated counts without individual linkage. *Quarterly Journal of Measurement Theory*, 19:137–154, 2022.

Michael P. Pound, Andrew P. French, Jonathan A. Atkinson, Darren M. Wells, Malcolm J. Bennett, and Tony Pridmore. RootNav: navigating images of complex root architectures. *Plant Physiology*, 162:1802–1814, 2013.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015.

Donald B. Rubin. Inference and missing data. *Biometrika*, 63:581–592, 1976.

Basil S. Yamey. Scientific bookkeeping and the rise of capitalism. *The Economic History Review*, 1:99–113, 1949.

Robail Yasrab, Jonathan A. Atkinson, Darren M. Wells, Andrew P. French, Tony P. Pridmore, and Michael P. Pound. RootNav 2.0: deep learning for automatic navigation of complex plant root architectures. *GigaScience*, 8:giz123, 2019.

Li Zhang, Yuan Li, and Ramakant Nevatia. Global data association for multi-object tracking using network flows. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.