Received: 4 September 2025 | Accepted: 7 January 2026

DOI: 10.1111/2041-210x.70257

Methods in Ecology and Evolution

PRACTICAL TOOLS

Automated quantification of fine root production from minirhizotron image time series

Alexander Gillert[1] | Patrick Möhl[2,3] | Bo Peters[4] | Lionel Safar[2] | Julia Bebout[1] | Stuart Schwab[1] | Erika Hiltbrunner[2] | Jürgen Kreyling[4] | Tara Javidi[5] | Elsa Cleland[1]

[1]Ecology, Behavior and Evolution Department, UC San Diego, La Jolla, California, USA

[2]Department of Environmental Sciences, University of Basel, Basel, Switzerland

[3]Lancaster Environment Centre, University of Lancaster, Lancaster, UK

[4]Institute of Botany and Landscape Ecology, Greifswald University, Greifswald, Germany

[5]Electrical and Computer Engineering, UC San Diego, La Jolla, California, USA

Correspondence
Alexander Gillert
Email: questions@alexander-gillert.com

Funding information
Schmidt Sciences LLC; Deutsche Forschungsgemeinschaft, Grant/Award Number: 531801029; Bundesministerium für Bildung und Forschung; Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Number: 31003A_182592

Handling Editor: Natalie Cooper

Abstract

1. Plant root growth accounts for a major part of the net primary production in grassland and forest ecosystems and influences the global carbon and nutrient cycles. Measuring the production of roots is inherently difficult, prone to inconsistencies and time-consuming. Notably, there are currently no methods yet to automate this task.
2. We have developed GINGER, a new method for automated estimation of the fine root production from a time series of minirhizotron images. It compares pairs of consecutive images with each other, separating new root growth from standing crop.
3. The method was evaluated on four datasets from grassland, drained fen peatland and forest ecosystems. It exhibits performance on a similar level to that of human annotators while substantially reducing the time required for the data analysis. Human annotators showed a significant degree of variability among each other, confirming that the task is subjective and error-prone.
4. For demonstration, this pipeline was applied on two real-world image datasets, spanning 2 and 3 years, to compute the total annual root production. End-to-end, including annotation and model training, GINGER reduced the required human workload from several thousand to less than 40 work hours. It could allow to scale up monitoring efforts and enable full automation in the future.

KEYWORDS
biogeochemical cycles, computer vision, machine learning, minirhizotrons, root production

1 | INTRODUCTION

Plants continuously grow new roots in search of water and nutrients, while older roots are lost to herbivory, pathogens or dieback. It has been estimated that root production accounts for 22% of the global terrestrial primary production (McCormack et al., 2015), with much higher percentages above 50% in grassland ecosystems (Mokany et al., 2006). It drives ecological processes, including carbon and nutrient cycling, and fuels soil microbial activity. It is, however, highly variable across species, biomes and strongly linked to interannual environmental conditions and their variability (McCormack et al., 2014), necessitating continuous monitoring.

Measuring root production is not straightforward, mainly due to the opaque nature of soil. Several methods have been developed over the decades, each of them comes with its own set of drawbacks and inaccuracies. *Sequential soil coring* has been the most commonly used method in the past. It is simple to perform but provides only a snapshot of standing below-ground biomass at the immediate

location and time of sampling. Observed changes are typically attributed solely to root production or mortality, not taking into account that both processes can occur simultaneously in between the sampling dates (Majdi et al., 2005). Additional sources of error include natural variation between coring locations, long sampling intervals and sample processing errors which can accumulate to threefold differences in root production estimates derived from the same soil cores (Publicover & Vogt, 1993). *Ingrowth cores* improve upon this by measuring the relative below-ground production into an initially root-free cylinder mesh from the time of its burial until its excavation. The issues with this method include a bias toward overestimation due to stimulated growth from roots that were damaged during installation, altered nutrient availability, soil structure and reduced competition (Majdi et al., 2005). Continuous measurements at the same location are not possible with this method either.

More recently, *minirhizotrons* have emerged as the method of choice for observing root dynamics (Iversen et al., 2012; Johnson et al., 2001). They allow for repeated, non-destructive, in situ monitoring of roots over time by acquiring images using a camera or scanner blade inserted into transparent tubes installed in soil. With minirhizotrons, root productivity can be measured in unit length per unit of soil area (typically $mm\,cm^{-2}$) by tracing new roots that have appeared or grown in an image observation and were not yet present or were shorter in previous observations. This method is more direct, allows observing the same location over long periods of time and in a steady state, but comes at the cost of very time-intensive data processing, which limits its adoption and scale of studies. For example, Burton et al. (2000) report that they processed only a subset of their raw image data due to time and labour constraints.

This work tackles this issue by automating the data processing step with computer vision and machine learning methods. Various automated minirhizotron data analysis pipelines have already been developed for the detection of root standing crop in below-ground images. They typically use semantic segmentation networks followed by a skeletonization postprocessing step, with some variations such as in (Gillert et al., 2021; Peters et al., 2023; Smith et al., 2020). These methods have subsequently been successfully employed in multiple ecological studies, for instance about root phenology (Möhl et al., 2022; Möhl & Hiltbrunner, 2025; Schwieger et al., 2022), with some studies reporting on their limitations (Handy et al., 2024).

From a computer vision point of view, measuring root production is a substantially harder problem statement than the detection of standing crop since a series of images need to be compared with each other to determine which roots have appeared or grown and which ones were already present in previous observations. The main difficulty hereby is that roots and the soil around them can move in between image scans, for instance due to changes in moisture. An initial proof-of-concept method has been published by Gillert et al. (2023) demonstrating the feasibility of automating this task. However, it exhibits several limitations in practice and has not yet been successfully employed in a real-world ecological study.

Here, we introduce *GINGER* (***G**uided **I**mage-based **N**ew **G**rowth **E**stimation of fine **R**oots*), which improves upon this method by changing the paradigm from the previous '*first-detect-then-track*' to the reverse '*first-track-then-detect*'. The benefit of this switch is that it helps to mitigate imperfect image registration, the process of aligning multiple images so that corresponding coordinate points match, which the hardest part of the problem. This also results in an overall simpler and faster processing pipeline. The method has been adapted from change detection methods, which are often used in satellite remote sensing. We believe it will enable continuous and automated fine root production monitoring on a larger temporal and spatial scale.

2 | METHODS

2.1 | Data

We utilize two main image datasets in this study. The first one (A) was obtained from a multi-year climate change experiment in an alpine grassland at 2480 m elevation in the Swiss Alps (Möhl & Hiltbrunner, 2025). No permits were required. Minirhizotron tubes (50 cm long, 56 mm diameter) were installed at a 45° angle in July 2019. Root images were taken with a narrow-gauge root imager (1200 DPI, CID-602, CID BioScience, US). Two images were taken per tube to cover the entire length, which corresponded to a soil depth of ca. 20 cm and stitched together to a size of approximately 15,000 × 8748 pixels. For the measurements of annual root production we used a total of 215 images from 5 tubes sampled at a weekly interval spanning from July to October 2021 and from June to September in 2022 and 2023. Example images in Figures 1a, 2 and 5b,c.

The second dataset (B) was acquired in a coastal fen peatland in north-eastern Germany that was drained almost a century ago, resulting from the construction of a dyke (nature reserve 'Karrendorfer Wiesen') (Schwieger et al., 2022). The site is dominated by grasses and was fenced off to exclude large herbivores. Permission was granted by the Lower Nature Conservation Authority Vorpommern-Greifswald (permit number 60.5/21/250/2017/01) and the land owner Succow Foundation. Minirhizotron tubes (50 cm long, 70 mm diameter) were installed at a 45° angle in August 2017 and image data were recorded with an image scanner (CI-600 In-Situ Root Imager; CID Bio-science Inc.) at 300 dpi at 3 depths down to ca 50 cm, resulting in individual images of size 2550 × 2273 pixels. For the measurements of annual root production, we used a total of 678 images from 15 tubes sampled at a monthly interval spanning 2 years from 2018 to 2019, from May to November each. Example images in Figures 1b and 5d.

Two additional datasets are used for evaluation: (C) was acquired in multiple beech forests in north-eastern Germany and north-western Poland with the same acquisition setup as (B). Permissions were granted by the regional forest management authorities Forst Brandenburg including the Landeskompetenzzentrum Forst Eberswalde, Landesforst Mecklenburg-Vorpommern, forest management of the University of Greifswald, National Forest Holding of Poland's State Forests in Szczecin, Gdańsk and Toruń. This dataset

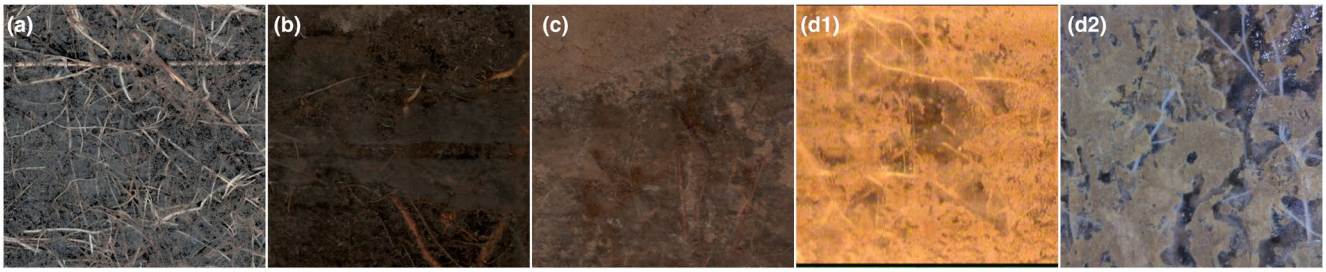

FIGURE 1 Example images of the datasets used in this paper. (a) alpine grassland, (b) drained fen peatland, (c) beech forest, (d1) agricultural setting (camera by Bartz), (d2) agricultural setting (camera by Vienna Scientific).

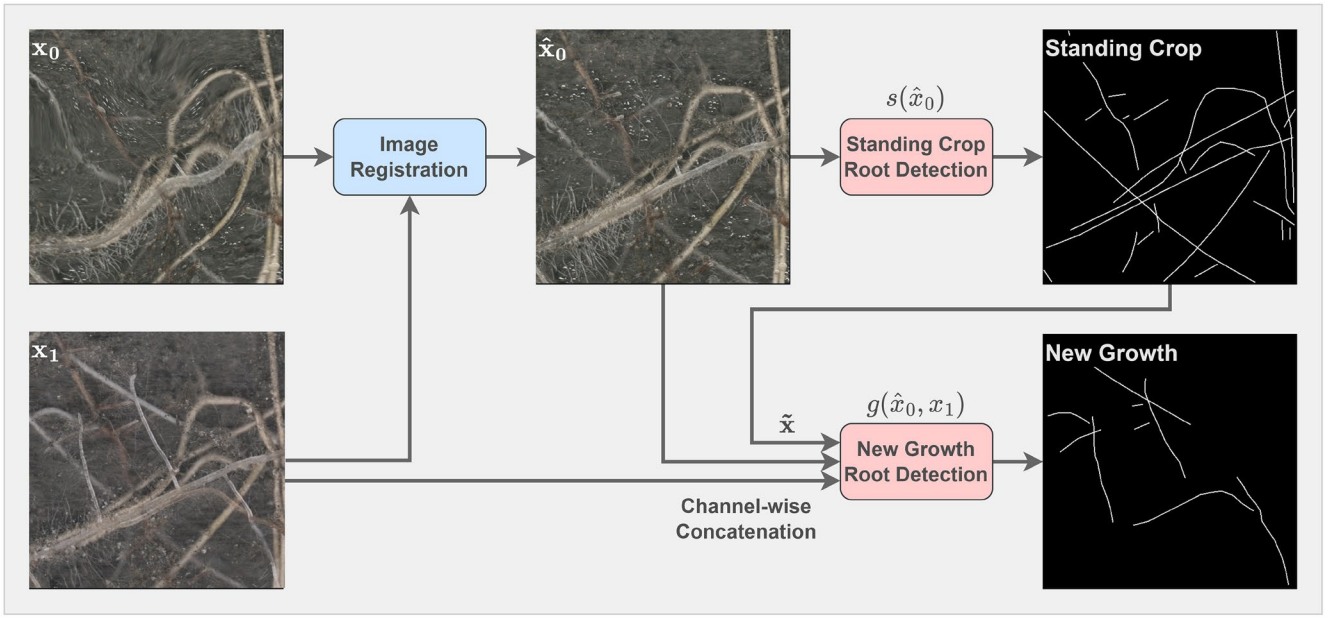

FIGURE 2 Overview of the GINGER processing pipeline. (Image x_0 is artificially warped to better illustrate the problem. In reality the movement is more subtle, but large enough to cause issues.)

contains many roots that are difficult to discern and to distinguish from fungal mycelia. Example images in Figures 1c and 5a.

Dataset (D) has been acquired by Lärm et al. (2023) in an agricultural setting. We split it into two subsets due to the use of different root imagers: (D1) was acquired with a camera-based root imager by Bartz Technology Corporation, (D2) was acquired with a camera by Vienna Scientific Instruments GmbH. We refer to Lärm et al. (2023) for more information. Example images in Figure 1d1,d2, respectively.

2.2 | Pipeline architecture

Our image analysis pipeline consists of 3 main processing steps. A schematic overview is shown in Figure 2. Given as inputs are two images x_0 and x_1 from the same minirhizotron tube, acquired at different dates, where x_0 is the chronologically earlier one. The main contribution of our pipeline is the detection of new roots that have appeared or grown in x_1 but were not yet present or were shorter in x_0.

The first step is to align the two images onto each other in order to remove the movement of roots and soil. This is crucial for the following step to help determine which roots are new. We have selected the RoMa model (Edstedt et al., 2024) for this task. This is a foundation model for dense image matching, meaning that it can map the location of each pixel from one image onto another. It leverages image features returned by a frozen DINOv2 (Oquab et al. (2023)), a transformer model pretrained under self-supervision via contrastive learning. These features have been found to be substantially more robust than local features trained from scratch and to easily transfer to unseen domains. An additional convolutional encoder improves the matching of fine details. We refer to Edstedt et al. (2024) for a more detailed technical description of this model. As a foundation model, RoMa has been trained on a large volume of diverse images and generalizes well on unseen images. We have found it to perform sufficiently well on minirhizotron images even without retraining.

The second model s is responsible for the detection of the standing crop roots in the aligned image $\hat{x}_0$. Following previous work on

automated root detection, we employ a U-Net semantic segmentation network (Ronneberger et al., 2015), postprocessed by a thinning (or skeletonization) algorithm (Zhang & Suen, 1984) to reduce its output to single pixel wide lines.

The third and final model g is responsible for detecting only the newly grown roots in the image x_1, ignoring those that were present in x_0. To this end, g receives as input a 7-channel tensor $\tilde{x}$, which consists of the aligned image $\hat{x}_0$, the original image x_1 and a grayscale image containing the detected standing crop roots, that is $s(\hat{x}_0)$. Those 3 images are concatenated along the channel dimension, or more intuitively speaking, they are overlayed onto each other. $\tilde{x}$ is then forwarded to g, which computes a new segmentation map representing only newly grown roots in x_1. During the training phase, it automatically learns to compare and to search for differences in the images. Model g has the same architecture as s, except for the number of input channels.

To avoid hardware limitations, this pipeline automatically processes high-resolution images on overlapping image patches of size 512×512 pixels and the result is stitched to the original size afterwards.

The outputs of the networks s and g are binary images. We use the following equations to extract relevant statistics related to root growth:

$$\begin{aligned} \text{annual root production} &= \sum_{i=0}^{n-1} \left| g(\hat{x}_i, x_{i+1}) \right| \\ \text{maximum root standing crop} &= \max_j \left| s(x_j) \right| \\ \text{relative root production} &= \frac{\text{annual root production}}{\text{maximum root standing crop}} \end{aligned} \tag{1}$$

where $|\ |$ denotes the sum of output pixels after postprocessing with the thinning algorithm and n the number of images in a year.

2.3 | Training

The neural network training is performed in two stages. First, the standing crop detection model s is trained via standard supervised training as in (Smith et al. (2020), Gillert et al. (2021) and Peters et al. (2023)). The growth detection model g requires inputs from s and thus needs to be trained afterwards, also supervised. During training, both images are augmented independently from each other to simulate environmental changes over time. Augmentations consist of nonlinear image deformation such as in Smith et al. (2020) to simulate movement of roots or soil, colour and brightness jitter to simulate different acquisition parameters and 90° rotations. The image registration model RoMa is not trained.

For dataset A, the standing crop model s was trained on a total of 78 sub-images of size 512×512 pixels and the growth detection model g on 338 pairs of sub-images of the same size. For dataset B, the corresponding training set sizes were 136 for s and 154 for g. Both models were trained with the AdamW optimizer (Loshchilov & Hutter, 2017), starting learning rate of 0.001 and cosine learning rate annealing. The models s were trained for 1000 epochs, and g for 5000 epochs since the task is harder.

Annotation was performed with the GIMP open-source graphics editor. For standing crop annotations, all visible roots in an image were traced with the 'Paths' tool, disregarding visual appearance and root diameter. Annotation of newly grown roots is significantly more challenging, since x_0 and x_1 are usually nearly identical, making it difficult to spot subtle changes, a perceptual phenomenon known in psychology as *change blindness* (Simons & Levin, 1997). Therefore, the two consecutive images were loaded as separate GIMP layers. By switching between both layers, changes become easier to detect. Annotators were tasked to trace only newly grown roots or root segments in x_1 that were not present or shorter in x_0.

2.4 | Evaluation

We compute the evaluation metrics from root lengths rather than overlapping pixels or individual roots. Commonly used pixelwise semantic segmentation metrics such as intersection over union (IoU) cannot be applied in this case, since we are dealing with single pixel wide lines. A single pixel shift could result in no intersecting pixels. Instead, we follow the evaluation protocol for skeletonized curvilinear structures by Youssef et al. (2015). It redefines true positives, false positives and false negatives as follows:

$$TP = S \cap \delta_\rho(S^*),\ \ FP = S \cap \overline{\delta_\rho(S^*)},\ \ FN = S^* \cap \delta_\rho(S) \tag{2}$$

where S is the binary skeletonized output of a network (either standing crop or new growth), S^* the annotation by an individual human annotator in rasterized form, $\overline{S^*}$ the inverse of the binary image S^* and δ_ρ the morphological dilation operation with a structuring element of size ρ. We empirically set $\rho = 7$. For more details on this evaluation protocol, see Youssef et al. (2015). From those values, we compute the commonly used precision (P) and recall (R) metrics in addition to over- or underestimated of total root length: $L = \frac{\text{length detected}}{\text{length annotated}} - 1$.

For the validation set, we have used 24 pairs of sub-images, approximately 2048×4400 pixels in size, for dataset A and 18 full-sized image pairs (2550×2273 pixels) for dataset B. Images were selected to be spread out throughout the growing seasons and from tubes that were not present in the training set to minimize bias. After this pre-selection, the selection of the validation images was random.

As previously noticed by Peters et al. (2023) and Handy et al. (2024), there is a substantial degree of disagreement between human annotators when measuring root standing crop in minirhizotron images. For instance, Handy et al. (2024) report a Concordance Correlation Coefficient (CCC) (Lawrence & Lin, 1989) value of only 0.56 even among two very experienced experts, indicating only moderate correlation. Since we are dealing with a new problem statement (detection of new growth rather than standing crop) and to put our metrics into perspective, our datasets A and B were annotated independently by three different annotators, two postdocs and one graduate student. All of them were at least on an

intermediate level of working with plant roots, which we define similarly to Handy et al. (2024) as having processed at least 30 images and 300 work hours with intact roots prior to the study. Annotators were only given technical instructions and interpreted the image data based on individual judgement. No time limits were given. We compare their annotations among each other in the same way as GINGER.

3 | RESULTS

3.1 | Fine root production

The root standing crop and growth lengths for datasets A and B as measured by our proposed method are presented in Figure 3. Table 1 shows the absolute and relative root production and maximum standing crop. We point out that many data points reveal some detected root growth even though the overall standing crop declines or stagnates, indicating simultaneous growth and disappearance of roots. At some time points, such as in week 42 in 2018 (dataset B), the increase in standing crop is larger than the detected growth. In those cases, *g* failed to detect some new root growth due to inconsistent acquisition parameters.

3.2 | Evaluation

A comparison of the outputs produced by GINGER against the manually measured root growth shows a high degree of agreement in general (Figure 4), but with a small bias toward overestimation. We attribute the lower performance on dataset C to many ambiguous cases in which roots and fungal mycelia are difficult to distinguish. Looking closer, precision and recall values in Table 2 reveal that our method makes several

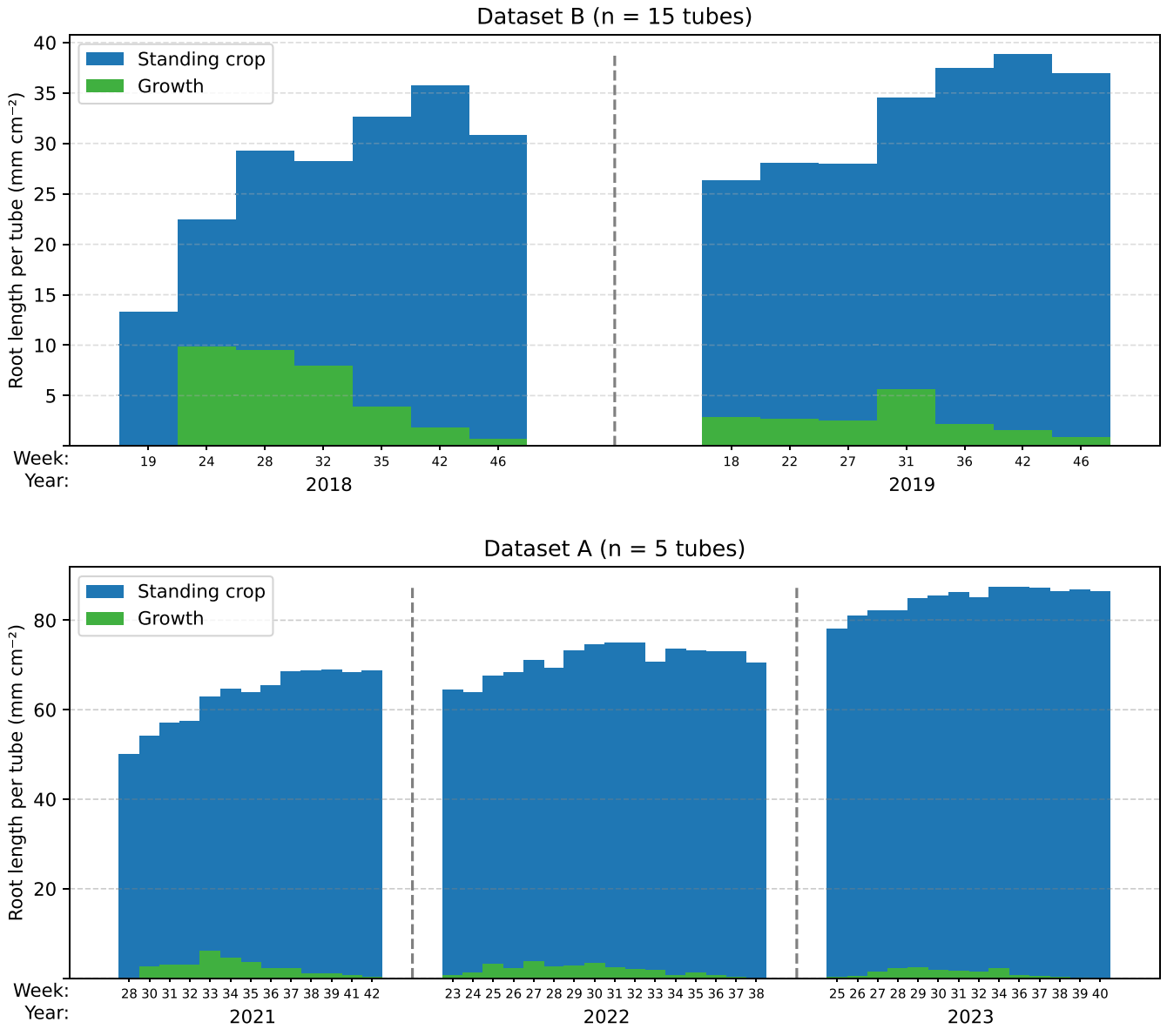

FIGURE 3 Time series of detected standing crop and absolute and relative root production for datasets A and B.

TABLE 1 Detected average lengths per area (mm cm^{-2}) of absolute and relative root production and maximum standing crop.

	Dataset A ($n = 5$ tubes)			Dataset B ($n = 15$ tubes)	
	2021	2022	2023	2018	2019
Annual root production (mm cm^{-2})	30.45 (±5.13)	29.31 (±5.13)	15.85 (±4.22)	33.72 (±17.54)	18.36 (±8.84)
Maximum root standing crop (mm cm^{-2})	70.32 (±13.78)	75.84 (±9.20)	88.53 (±4.33)	38.98 (±16.96)	41.14 (±20.28)
Relative root production (mm cm^{-2} year^{-1})	0.443 (±0.092)	0.393 (±0.093)	0.179 (±0.045)	0.864 (±0.189)	0.472 (±0.119)

Note: Standard deviations in brackets (the values are averaged per tube and thus do not perfectly match up).

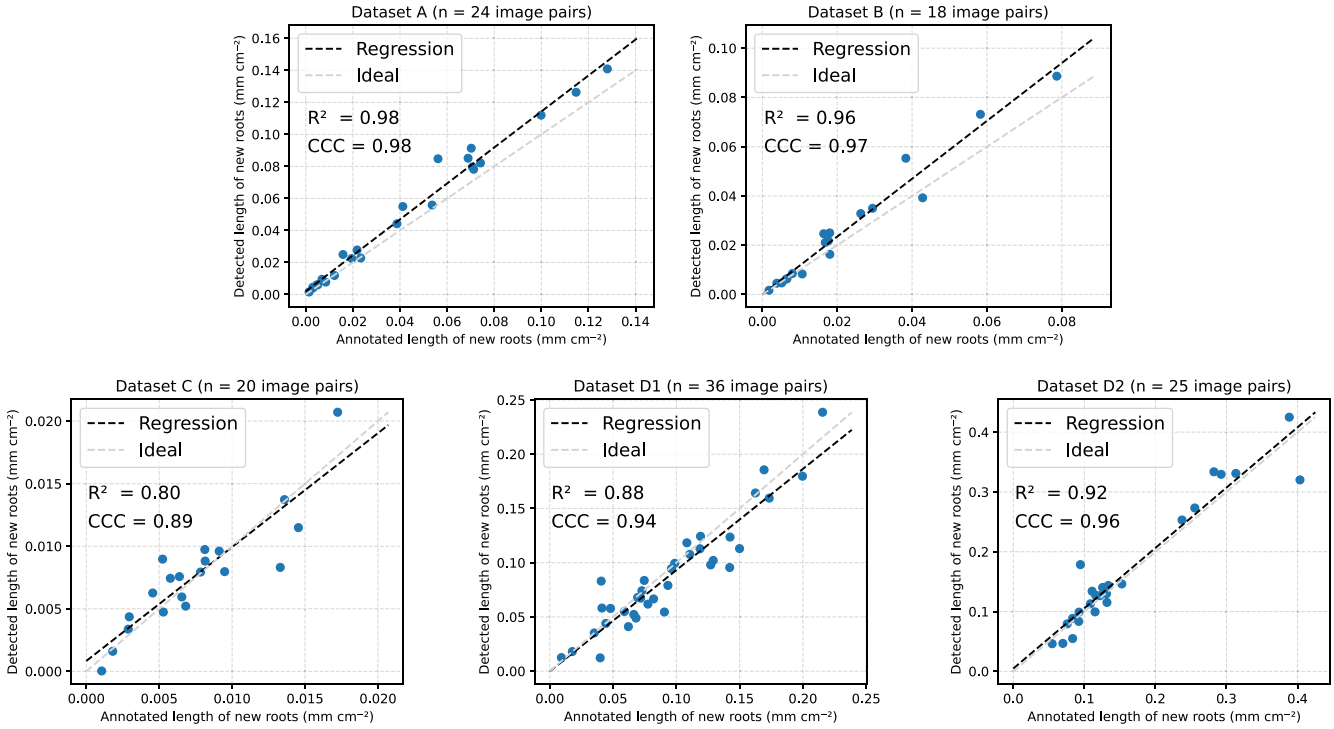

FIGURE 4 Relationship between detected and annotated total length of newly grown roots. Each point represents a validation image pair. Values for datasets A and B are averaged across three annotators, values for datasets C, D1 and D2 by annotator 1. CCC refers to the concordance correlation coefficient as defined by Lawrence and Lin (1989).

mistakes, mainly misclassifying old roots as new (false positives) as well as overlooking some (false negatives); examples are shown in Figure 5b. Those mistakes mostly cancel each other out when looking only at the total length values. The precision and recall values among the annotators ranged from ca 60% to 83%, while GINGER showed an overall lower performance but on a similar level. It agreed the most with annotator 1, who provided the training annotations.

As illustrated in Figure 5, the main reason for false positives produced by GINGER was imperfect image registration due to sometimes large movements of roots. For human annotators, the most common type of mistakes was false negatives, that is overlooking small, hardly visible roots due to change blindness or fatigue.

We used a Nvidia GeForce 4090 GPU for training and inference. For dataset A, the processing of a full-sized image pair took 40 s on average, resulting in ca 2.3 h processing time for the full dataset spanning 3 years and 5 tubes. Training of both models took ca. 15 h in total, in addition to ca. 18 work hours for the annotation of the training data. Human annotators required on average 25 min to annotate only the newly grown roots in a single cropped validation image pair. Extrapolating this number to the full dataset, this would have required ca. 1225 work hours of processing time, in addition to standing crop measurements which we estimate would have exceeded 3000 work hours.

4 | DISCUSSION

Our work presents the first automated pipeline for the quantification of fine root production with minirhizotrons. While the proposed method GINGER does not achieve perfect accuracy, this reflects the substantial variability observed among human annotators. The inter-annotator comparison suggests that the task itself is inherently challenging and prone to subjectivity. Nevertheless, our results demonstrate that the method provides usable outputs at a fraction of the time required for manual processing and mistakes could be further corrected with minimal manual postprocessing effort as explained in the supplement.

TABLE 2 Evaluation metrics and comparison among human annotators for all datasets.

	Dataset A ($n = 24$ image pairs)								
	Annotator 1			Annotator 2			Annotator 3		
	L	P	R	L	P	R	L	P	R
GINGER	+9.2%	73.5%	85.1%	+13.7%	66.9%	76.7%	+29.7%	58.2%	75.4%
Annotator 1				+3.3%	77.1%	75.2%	+17.6%	67.5%	73.1%
Annotator 2	−3.1%	75.4%	77.3%				+14.3%	60.6%	68.0%
Annotator 3	−14.9%	73.3%	67.7%	−12.6%	68.0%	60.6%			
	Dataset B ($n = 18$ image pairs)								
	Annotator 1			Annotator 2			Annotator 3		
	L	P	R	L	P	R	L	P	R
GINGER	+6.9%	70.2%	73.4%	+16.5%	65.3%	68.6%	+31.1%	57.9%	73.4%
Annotator 1				+8.2%	79.7%	79.2%	+22.1%	69.6%	83.5%
Annotator 2	−7.6%	79.3%	79.7%				+13.6%	65.8%	79.9%
Annotator 3	−18.4%	83.5%	69.6%	−12.1%	79.9%	65.8%			
	Dataset C ($n = 20$)			Dataset D1 ($n = 36$)			Dataset D2 ($n = 25$)		
	L	P	R	L	P	R	L	P	R
GINGER	+2.1%	68.5%	66.7%	−6.9%	70.6%	66.3%	+3.7%	68.7%	69.8%

Note: Annotator 1 supplied the annotations for neural network training and for datasets C, D1, D2. L: over- or underestimation of the total length of new roots, P: average precision, R: average recall Values are not perfectly symmetric because of the dilation operation in the evaluation method (see Equation 2).

Our results also confirm measurements via destructive methods are likely to overestimate root productivity (Majdi et al., 2005). In the two ecosystems studied here, root standing crop was still increasing even after four (dataset A) and two (dataset B) years after installation of the rhizotrons, suggesting that a steady state was not yet reached. Simultaneously, net root growth declined over time, pointing to a potential overestimation of root production in the initial years. A similar bias is to be expected for ingrowth cores, which are usually excavated after comparably short periods of 1–2 years. Noninvasive monitoring and long-term experiments are thus necessary to reliably quantify root production in a steady state.

Given the substantial variability between human annotators we suggest that future studies employ measures to increase harmonization, such as detailed annotation guidelines and training sessions before actual annotation (Ahmadzadeh et al., 2025). Priorities for further improvements of the pipeline include a better image registration method. RoMa is able to counteract soil movement relatively well. It struggles in cases where roots move independently of the surrounding soil, which we have found to be the main reason for false positives. Additional pitfalls of our method include double-counting of roots that disappear (or just do not get detected) in one image and re-appear in a following image. This can happen due to occlusions but we have found this to be a rare occurrence. This happens more often because of inconsistent image quality, which highlights the need for standardized image acquisition protocols.

For best results, we recommend minirhizotron root imagers based on scanner blades that can cover a large area of soil in a single image at 600 to 1600 dpi. If camera-based imagers are used instead, we recommend undistorting and seamlessly stitching the images as a preprocessing step to maximize the areas of visual overlap between time steps. Retraining on a new site is usually required; the instructions can be found in the supplement.

A limitation of GINGER is that it only measures root lengths, whereas other traits such as the root diameter are also relevant to make inferences about carbon or nutrient cycling (Lorène et al., 2025). Minirhizotrons excel in the detection of fine roots (Johnson et al., 2001), which are active in water and nutrient uptake and highly important in ecosystem carbon cycling, while coarse roots provide stability in the ground and grow much less in length. Moreover, many plant species form other below-ground organs such as rhizomes, tubers or taproots that largely differ from fine roots in functionality and longevity (Ottaviani et al., 2024). The spatial and temporal development of these below-ground organs cannot be covered by GINGER, nor by the minirhizotron technique in general, as their occurrence in the soil is extremely heterogeneous and hardly observable in the small sections visible through minirhizotrons.

The datasets in this work were still relatively short term and small in scale. An image processing pipeline measuring root production, such as the one described here, can enable full automation when paired with robotic minirhizotrons (Defrenne et al., 2021; Nair et al., 2023), which are able to acquire image data without supervision. Images could then be transmitted to a remote server or processed by a local computing unit on site. This will allow scaling up monitoring efforts and to more precisely understand variability of root dynamics in different biomes and environmental conditions.

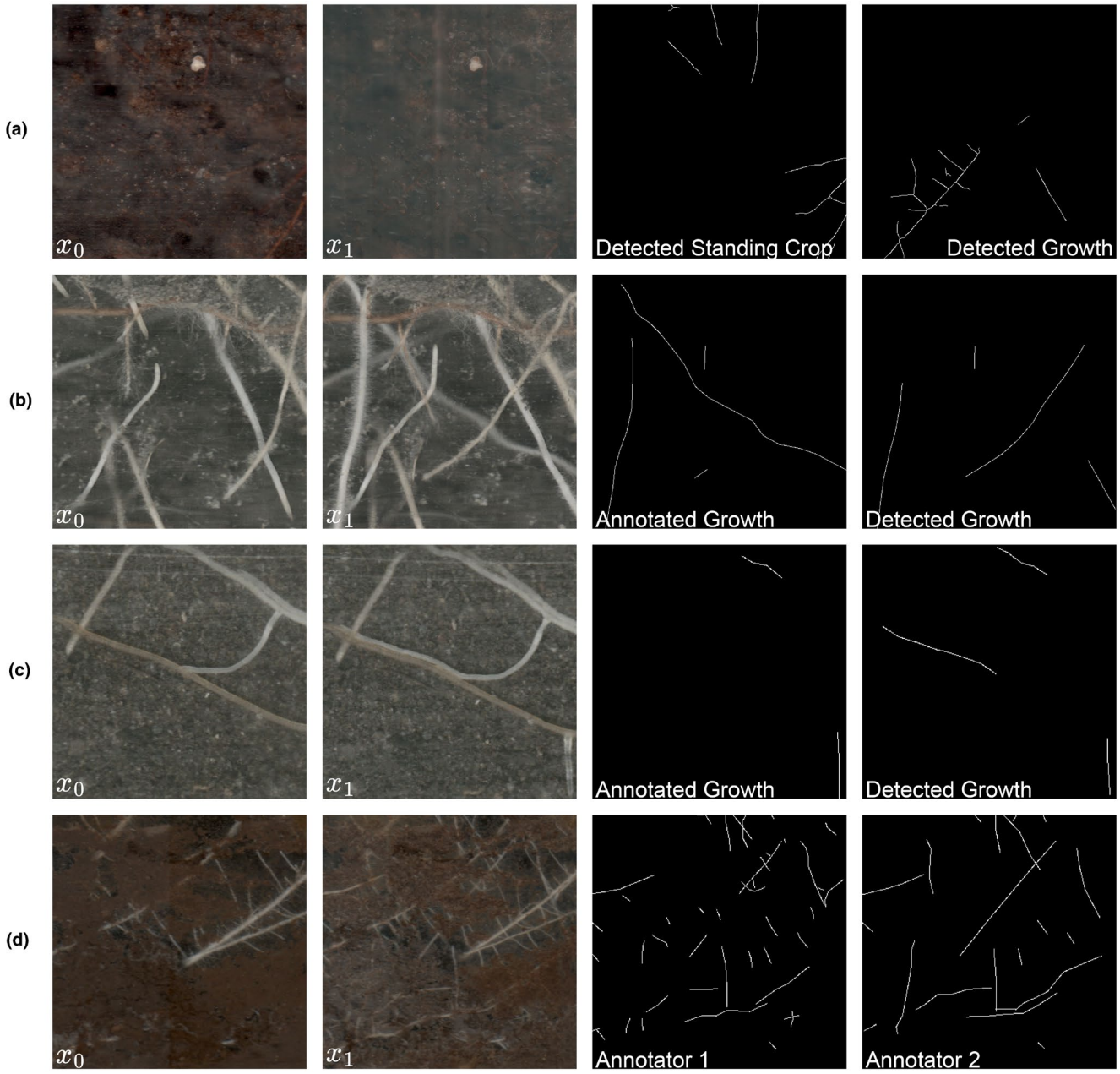

FIGURE 5 Example images (cropped) and typical mistakes. (a) GINGER is able to detect roots with low contrast to soil. (b) GINGER incorrectly classified a moving root as new (false positive) and failed to detect an out-of-focus blurred root (false negative). (c) A typical mistake by human annotators is overlooking roots that are in close proximity to others. All three annotators missed the root in the centre. (d) Differences in annotators' performance.

Considering the important role of plant root production for the global C cycle, it is striking that we still lack robust approaches to characterize this process at broad scales. With the pipeline presented here, we provide a high-throughput way to assess root production based on direct, non-destructive observations from the widely employed minirhizotron technique. This approach greatly shortens processing time for analysing root images and thus enables a much wider application. Long-term studies in which root production reaches a steady state are still rare but necessary. When combined with machine-learning tools such as GINGER, they will allow for a more accurate and representative representation of root productivity across global ecosystems.

AUTHOR CONTRIBUTIONS

Alexander Gillert conceived the idea for the algorithm and wrote the source code, advised by Tara Javidi; Patrick Möhl, Bo Peters, Lionel Safar, Erika Hiltbrunner and Jürgen Kreyling collected the image data; Alexander Gillert, Julia Bebout and Stuart Schwab annotated the training and validation images; Alexander Gillert and Lionel Safar evaluated the performance; Patrick Möhl, Erika Hiltbrunner, Jürgen Kreyling and Elsa Cleland advised and assisted on ecological background and implications; Alexander Gillert led the writing of the manuscript with edits from all authors.

ACKNOWLEDGEMENTS

This research received support through Schmidt Sciences LLC, Roots@ArtIGROW funded by BMFTR through the WIR!-Bündnis ArtIFARM (BMBF) and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 531801029—TRR 410 (WETSCAPES2.0). Data acquisition of dataset A was funded by the Swiss National Science Foundation (SNSF, 31003A_182592). We thank the community Obergoms (Valais, Switzerland) for offering the alpine grassland for this field experiment.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210x.70257.

DATA AVAILABILITY STATEMENT

Data available via https://doi.org/10.5281/zenodo.17014187 (Gillert et al., 2026b). Source code archived on https://doi.org/10.5281/zenodo.18349858 (Gillert et al., 2026a), for the most recent version please refer to https://github.com/alexander-g/GINGER. The corresponding author can provide additional technical support.

STATEMENT ON INCLUSION

Our study brings together authors from a number of different countries, including scientists based in the country where the primary data were collected.

ORCID

Alexander Gillert https://orcid.org/0009-0000-8777-2588
Bo Peters https://orcid.org/0000-0001-7597-4211
Julia Bebout https://orcid.org/0009-0006-1893-0670
Jürgen Kreyling https://orcid.org/0000-0001-8489-7289
Elsa Cleland https://orcid.org/0000-0003-3920-0029

REFERENCES

Ahmadzadeh, A., Adhyapak, R., Iraji, A., Chaurasiya, K., Aparna, V., & Martens, P. C. (2025). A guide for manual annotation of scientific imagery: How to prepare for large projects. ArXiv, abs/2508.14801.

Burton, A., Pregitzer, K., & Hendrick, R. (2000). Relationships between fine root dynamics and nitrogen availability in michigan northern hardwood forests. *Oecologia*, *125*, 389–399.

Defrenne, C. E., Childs, J., Fernandez, C. W., Taggart, M., Nettles, W. R., Allen, M. F., Hanson, P. J., & Iversen, C. M. (2021). High-resolution minirhizotrons advance our understanding of root-fungal dynamics in an experimentally warmed peatland. *Plants, People, Planet*, *3*(5), 640–652.

Edstedt, J., Sun, Q., Bökman, G., Wadenbäck, M., & Felsberg, M. (2024). RoMa: Robust dense feature matching. IEEE Conference on Computer Vision and Pattern Recognition.

Gillert, A., Möhl, P., Peters, B., Safar, L., Bebout, J., Schwab, S., Hiltbrunner, E., Kreyling, J., Javidi, T., & Cleland, E. (2026a). Automated quantification of fine root production from minirhizotron image time series (code).

Gillert, A., Möhl, P., Peters, B., Safar, L., Bebout, J., Schwab, S., Hiltbrunner, E., Kreyling, J., Javidi, T., & Cleland, E. (2026b). Automated quantification of fine root production from minirhizotron image time series (data). *Methods in Ecology and Evolution*. https://doi.org/10.1037/ppm0000185.

Gillert, A., Peters, B., von Lukas, U. F., & Kreyling, J. (2021). Identification and measurement of individual roots in minirhizotron images of dense root systems. In *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)* (pp. 1323–1331). IEEE.

Gillert, A., Peters, B., von Lukas, U. F., Kreyling, J., & Blume-Werry, G. (2023). Tracking growth and decay of plant roots in minirhizotron images. In *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (pp. 3688–3697). IEEE.

Handy, G., Carter, I., Mackenzie, A. R., Esquivel-Muelbert, A., Smith, A. G., Yaffar, D., Childs, J., & Arnaud, M. (2024). Variation in forest root image annotation by experts, novices, and ai. *Plant Methods*, *20*(1), 154.

Iversen, C. M., Murphy, M. T., Allen, M. F., Childs, J., Eissenstat, D. M., Lilleskov, E. A., Sarjala, T. M., Sloan, V. L., & Sullivan, P. F. (2012). Advancing the use of minirhizotrons in wetlands. *Plant and Soil*, *352*, 23–39.

Johnson, M. G., Tingey, D. T., Phillips, D. L., & Storm, M. J. (2001). Advancing fine root research with minirhizotrons. *Environmental and Experimental Botany*, *45*(3), 263–289.

Lärm, L., Bauer, F. M., Hermes, N., van der Kruk, J., Vereecken, H., Vanderborght, J., Nguyen, T. H., Lopez, G., Seidel, S. J., Ewert, F. A., Schnepf, A., & Klotzsche, A. (2023). Multi-year belowground data of minirhizotron facilities in selhausen. *Scientific Data*, *10*, 672.

Lawrence, I., & Lin, K. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, *45*, 255–268.

Lorène, S., Garynechèle, S., Isabelle, B., Mickaël, H., & Christophe, J. (2025). High-throughput monitoring of root diameter reveals the temporal dynamics of root decomposition. *Annals of Botany*, *136*, 1163–1173.

Loshchilov, I., & Hutter, F. (2017). Fixing weight decay regularization in adam. arXiv Preprint arXiv:1711.05101, 5(5): 5.

Majdi, H., Pregitzer, K., Moren, A.-S., Nylund, J.-E., & Ågren, G. I. (2005). Measuring fine root turnover in forest ecosystems. *Plant and Soil*, *276*, 1–8.

McCormack, M. L., Adams, T. S., Smithwick, E. A., & Eissenstat, D. M. (2014). Variability in root production, phenology, and turnover rate among 12 temperate tree species. *Ecology*, *95*(8), 2224–2235.

McCormack, M. L., Dickie, I. A., Eissenstat, D. M., Fahey, T. J., Fernandez, C. W., Guo, D., Helmisaari, H.-S., Hobbie, E. A., Iversen, C. M., Jackson, R. B., Leppälammi-Kujansuu, J., Norby, R. J., Phillips, R. P., Pregitzer, K. S., Pritchard, S. G., Rewald, B., & Zadworny, M. (2015). Redefining fine roots improves understanding of below-ground contributions to terrestrial biosphere processes. *The New Phytologist*, *207*(3), 505–518.

Möhl, P., & Hiltbrunner, E. (2025). Recurrent summer drought temporarily stimulates fine root growth but enhances winter root losses in alpine grassland. *Frontiers in Plant Science*, *16*, 1625076.

Möhl, P., Von Büren, R. S., & Hiltbrunner, E. (2022). Growth of alpine grassland will start and stop earlier under climate warming. *Nature Communications*, *13*(1), 7398.

Mokany, K., Raison, R. J., & Prokushkin, A. S. (2006). Critical analysis of root: shoot ratios in terrestrial biomes. *Global Change Biology*, *12*, 84–96.

Nair, R., Strube, M., Hertel, M., Kolle, O., Rolo, V., & Migliavacca, M. (2023). High frequency root dynamics: Sampling and interpretation using replicated robotic minirhizotrons. *Journal of Experimental Botany*, *74*(3), 769–786.

Oquab, M., Darcet, T., Moutakanni, T., Vo, H. V., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Howes, R., Huang, P.-Y., Xu, H., Sharma, V., Li, S.-W., Galuba, W., Rabbat, M.,

Assran, M., Ballas, N., … Bojanowski, P. (2023). Dinov2: Learning robust visual features without supervision.

Ottaviani, G., Klimešová, J., Charles-Dominique, T., Millan, M., Harris, T., & Silveira, F. A. (2024). The underestimated global importance of plant belowground coarse organs in open biomes for ecosystem functioning and conservation. *Perspectives in Ecology and Conservation*, *22*(2), 118–121.

Peters, B., Blume-Werry, G., Gillert, A., Schwieger, S., von Lukas, U. F., & Kreyling, J. (2023). As good as human experts in detecting plant roots in minirhizotron images but efficient and reproducible: The convolutional neural network "rootdetector". *Scientific Reports*, *13*(1), 1399.

Publicover, D. A., & Vogt, K. A. (1993). A comparison of methods for estimating forest fine root production with respect to sources of error. *Canadian Journal of Forest Research*, *23*(6), 1179–1186.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234–241). Springer.

Schwieger, S., Kreyling, J., Peters, B., Gillert, A., von Freiherr Lukas, U., Jurasinski, G., Köhn, D., & Blume-Werry, G. (2022). Rewetting prolongs root growing season in minerotrophic peatlands and mitigates negative drought effects. *Journal of Applied Ecology*, *59*(8), 2106–2116.

Simons, D. J., & Levin, D. T. (1997). Change blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267.

Smith, A. G., Petersen, J., Selvan, R., & Rasmussen, C. R. (2020). Segmentation of roots in soil with u-net. *Plant Methods*, *16*(1), 13.

Youssef, R., Ricordeau, A., Sevestre-Ghalila, S., & Benazza-Benyahya, A. (2015). Evaluation protocol of skeletonization applied to grayscale curvilinear structures. In *2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA)* (pp. 1–6). IEEE.

Zhang, T. Y., & Suen, C. Y. (1984). A fast parallel algorithm for thinning digital patterns. *Communications of the ACM*, *27*(3), 236–239.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Text S1. Usage instructions for the described software method GINGER.

How to cite this article: Gillert, A., Möhl, P., Peters, B., Safar, L., Bebout, J., Schwab, S., Hiltbrunner, E., Kreyling, J., Javidi, T., & Cleland, E. (2026). Automated quantification of fine root production from minirhizotron image time series. *Methods in Ecology and Evolution*, *00*, 1–10. https://doi.org/10.1111/2041-210x.70257