REPLICATE SUB-DAILY MINIRHIZOTRON SAMPLING, DATA AND INTERPRETATION FOR ROOT DYNAMICS STUDIES

Richard Nair[1*] / Martin Strube[2] / Martin Hertel[2] / Olaf Kolle[2] / Victor Rolo[3] / Mirco Migliavacca[1,4]

1 Department for Biogeochemical Integration, Max-Planck-Institute for Biogeochemistry, 07745 Jena, Germany

2 Max-Planck-Institute for Biogeochemistry, 07745 Jena, Germany

3 Forest Research Group, INDEHESA, University of Extremadura, 10600, Plasencia, Spain

4 Current address: European Commission, Joint Research Centre, Ispra (VA), Italy

*Corresponding author, rnair@bgc-jena.mpg.de

Keywords: minirhizotron, root dynamics, root-shoot-synchrony, phenocam, digital repeat photography

ABSTRACT

1. Minirhizotrons (paired camera systems and buried observatories) are the best current method to make repeatable measurements of fine root dynamics in the field. Automating data collection is the only way to assess fine-scale temporal changes in fine roots in natural systems. When paired with above-ground proximal remote sensing, automatic minirhizotron data could reveal coordination between above and below ground phenology, supporting both plant allocation and carbon exchange understanding. However, there are many challenges to automating minirhizotrons for field experiments. First, instruments must be cheap enough to replicate given shallow field of view and variable natural environments. Second, automated analysis must be robust to changeable soil and root conditions because image properties extracted must have biological meaning.
2. Here, we demonstrate an automatic minirhizotron scheme, built with off-the-shelf parts and capable to sample at sub-daily resolution paired with a neural network-based image analysis method. As a proof of concept, we tested the instruments and the methodology in a mesocosm study and in two short real-world field trials alongside datasets from other instruments (e.g. proximal remote sensed leaf dynamics).
3. We show that the method produces robust daily time-series of root dynamics in the mesocosm and the field. We applied the same model across multiple instruments, demonstrating good reproducibility of measurements. Temporal variations of root cover were a strong driver of soil CO_2 efflux in mesocosm, and vegetation dynamics from above ground proximal remote sensing and minirhizotron data differed both in mesocosm and field. We found sensitivity of root cover extracted to soil moisture and sometimes time of day (potentially relating to soil moisture), which may only affect high frequency imagery and not commonly reported as encountered when training neural networks on traditional, time-distinct minirhizotron studies.
4. We discuss avenues for dealing with such issues in future field applications of minirhizotrons. If such issues are dealt with to a satisfactory manner in the field, timeseries of root biomass and traits from replicated instruments can add a new dimension to phenology studies at the ecosystem level by understanding the dynamics of root properties and traits.

INTRODUCTION

Plant phenology (seasonal patterns of recurrent events such as leaf growth and senescence) drives interannual variability of the terrestrial carbon (C) sink (Raupach et al., 2011). It determines vegetation interactions with all biogeochemical cycles above- and below- ground (e.g., C fixation, respiration, transpiration, nutrient and water uptake) (Keenan et al., 2014; Migliavacca et al., 2012). Plant phenology responds to variable environmental conditions such as climate and weather (Richardson et al., 2013), can differ above and below ground (Adair et al., 2019), and is partially determined by life history strategy (Steinaker et al., 2010). While overall plant growth is driven by whole-plant resource budgets, root and shoot phenology are not always linked (Abramoff & Finzi, 2015) as conditions differ in the two environments (e.g., Blume-Werry, Jansson, & Milbau, 2017) and because plants temporally partition resource uptake/demand, assignment and activity of individual organs (Herrmann et al., 2016).

Tools to remotely monitor ecosystem phenology above-ground are well developed, for instance 'PhenoCams' (Luo et al., 2018; Richardson et al., 2018) or satellite-derived vegetation indexes (Wu et al., 2017), but using above-ground measurements to proxy for below-ground activity is, at best, an unreliable assumption (Abramoff & Finzi, 2015). Root phenology measurements are needed in field contexts and high frequency datasets are very rare (Radville et al., 2016). This is because of the demanding, destructive, non-repeatable nature of traditional sampling. Data scarcity contributes to highly simplistic simulation of root dynamics by models, often relying on poorly calibrated root:shoot ratios, simple environmental triggers, or optimality concepts (De Kauwe et al., 2014; Walker et al., 2015), which are hard to validate without data.

Roots are also crucial for C cycling because they control decomposition, and are the main C source to soil organic matter (Dijkstra et al., 2021). Thus, misrepresentation of root dynamics also affects prediction of ecosystem capacity to sequester carbon. Unfortunately, pot and mesocosm experiments are also prone to artefacts which disproportionately affect roots (Poorter et al., 2012) so the benefits of field experiments (Schindler, 1998) are particularly acute for realism below-ground. In natural field conditions, repeatable root observations are made with minirhizotrons (buried observatories and camera systems); many other advances in root phenotyping (e.g. Le Marié et al., 2014; Liu et al., 2021; Nagel et al., 2012) are only currently possible in more controlled conditions. In the last decade, minirhizotron automation

has been possible (Allen et al., 2007). However, application of robotic minirhizotrons for monitoring below ground phenology of roots has barely advanced beyond pioneering experiments using these systems (Allen & Kitajima, 2013, 2014; Iversen et al., 2011; Vargas & Allen, 2008) and while valuable, still use low replication and infrequent imaging and manual-driven analysis (e.g. Defrenne et al., 2021). Current methods do not allow efficient replication in either time or space, partially due to financial entry cost from design specifications – for instance, targeting ectomycorrhizal fungi necessitates high magnification and limited spatial coverage of images (e.g. Allen & Kitajima, 2014). As Minirhizotrons have an extremely short 2-3 mm field of view (Taylor et al., 2014), affordable replication and attention to necessary resolution is particularly critical for scalable, real-world measurements,.

Automated observations also bring a new bottleneck: efficient workflow to process the imagery to calculate proxies of root properties of interest. A manual approach to these will inevitably not be able to annotate images as fast as they are collected and development of machine or deep learning to solve this problem is promising, but not yet standard. Recently, convolutional neural networks (CNNs) showed promising results to identify roots in a variety of settings (Bauer et al., 2021; Delory et al., 2016; Huo & Cheng, 2019; Rahmanzadeh & Shojaedini, 2016; Vincent et al., 2016). In particular, specific field soil minirhizotron imagery has already been analysed several times with good results (Gillert et al., 2021; Han et al., 2021; Peters et al., 2022; Smith et al., 2020; Wang et al., 2019), but transferability between sites is difficult to assess without widespread adoption in new settings.

In this study, we built a robotic minirhizotron system (henceforth, RMR) capable of sub-daily sampling and paired this with an established CNN method *(Rootpainter*, Smith et al., 2020) previously used in low sampling frequency studies (Alonso-Crespo et al., 2022; Han et al., 2021) to extract information relating to root density from images collected at up to sub-daily resolution. The objective was to build and test the RMR, showing that the off-the-shelf CNN was suitable for this high frequency data. The RMR was made with constraints of a limited per-instrument budget, allowing replication. There are three main challenges to using automated minirhizotrons besides collection and interpretation itself: 1) usefulness of the minirhizotron indexes to explain system functioning at fine timescales 2) comparability between different point sensors and 3) long term robustness, in terms of field operation, as automatic systems must work on schedule in adverse conditions, and image processing, as

113 high frequency and non-opportune fixed sampling times bring other issues (e.g.
114 condensation) which can be avoided in manual campaigns.

115 Here, we do not address hypothesis-driven objectives but present summary data to show the
116 RMR works in the field and produce interpretable time series. We address this through four
117 experiments. The first two experiments were in the greenhouse: experiment 1 (E1) using a
118 single RMR; Experiment 2 (E2) using 8 replicated RMRs. Objectives were 1) to test the
119 systems in different experimental settings (E1 and E2) 2) to demonstrate linkage (or not) of
120 high root density dynamics extracted from the RMR to high resolution above-ground
121 biomass dynamics (E1 and E2) and soil CO_2 flux temporal variations (i.e., potentially
122 explaining an aspect of ecosystem function through root dynamics) (E1) and 3) to examine
123 consistency between multiple uncalibrated instruments/sensors in terms of roots and their
124 dynamics observed using a single trained CNN to segment all images (E2). Furthermore,
125 we performed two additional trials under field conditions in real and sometimes adverse
126 weather conditions (E3 and E4). In E3 and E4 we were observing real soils, with
127 condensation, soil animals, root litter and other disturbances potentially affecting instrument
128 operation, human annotation and consistency of CNN segmentation across a high frequency
129 timeseries.

130

131

MATERIALS AND METHODS

Minirhizotron System Design

We based our minirhizotron instrument (RMR, Fig. 1) on a movable camera design able to travel in two axes (along a minirhizotron tube, and rotationally around a minirhizotron tube). The RMR uses an (internal) 10/9.6 (external/internal) x 100 cm observatory. The system takes around 40 minutes to sample 112 separate images covering the entire tube and was built with standard electronic components. RMRs operate at any angle, we ran the RMR horizontally (E1 and E2) or at an angle (E3 and E4); generally, minirhizotrons should be installed at as shallow an angle relative to the surface as possible to minimize artefacts (Cai et al., 2016) or 45° to capture both lateral and horizontal roots (Freschet et al., 2021). The steeper the angle, the deeper the soil accessible. We also tested the RMR across a range of temperatures from -20°C to ~ 35°C (although see details later). Images (25 µm pixel^{-1}, ~1000 dpi. in our final system) are captured with a small overlap which we removed automatically. Further information about the design can be found in S1.

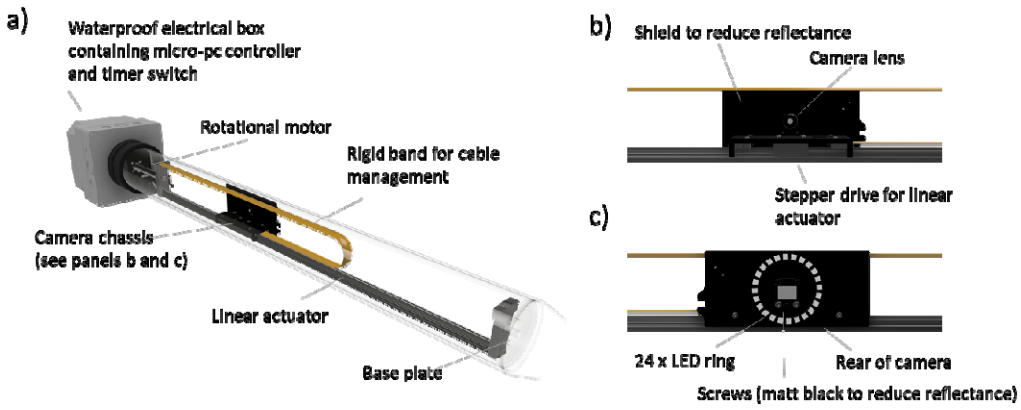

Figure 1. Conceptual diagram of the RMR. The full system (a) moves a camera chassis (b) longitudinally and rotationally to image soil in close contact with the tube. The camera chassis is designed to minimise reflectance by mounting the light source away from the focal area (c).

Greenhouse Experiments

We designed two greenhouse experiments: E1), which used a single RMR from mid-September to mid December 2018 (690 cycles, 77280 individual images) and E2), with eight RMRs from Mid-July to Mid-September 2019 (216,000 individual images). In E1 we sampled every two hours and had additional ancillary instruments, in E2 we sampled every six hours in replicate without paired ancillary instruments. Each mesocosm (a box 140 cm (L) x 40 cm (W) x 30 cm (H)) was filled with soil sieved to 0.5 cm (roots and stones removed), harvested from Meusebach, Germany (L. Eder pers. comm.). The C:N of the soil was 18:2. The soil in each mesocosm ended 2 cm above the upmost surface of the RMR. The RMR field of view spanned between 2 and 12 cm below the soil surface.

The soil was smoothed and, in both experiments, a 3-species seed mix (80/15/5 % by mass *Anthroxanthum odoratum, Plantago lanceolata, Medicago scutellata*) was scattered evenly over the surface, kept moist to germinate. Water was provided at irregular but recorded intervals and volumes via a watering can with a distributor nozzle to mimic irregular field rain events. We watered until an arbitrary date, then withheld water to move the system towards a state of (drought-induced) root senescence. This senescence period was longer in E2 (1 month) than E1 (2 weeks).

Each RMR observatory was located in a mesocosm unit (Fig. S2), extending internally 90 cm from one end. As here the RMR was horizontal, the first ~ 10 cm of the length, which is usually partially above ground, was not sampled. Aside from the camera, which we changed because of a supply issue (resulting in worse image quality in E2), and minor modifications for cable management, the RMR design was the same between E1 and E2. Both experiments were run in the greenhouse of the Max Planck Institute for Biogeochemistry, Jena, Germany.

All mesocosm units in both greenhouse experiments were included in the field of view of a standard 'PhenoCam' set up (Richardson et al., 2018; Sonnentag et al., 2012), modified for indoor use via a fish-eye lens, capturing daily images from which we obtained daily green chromatic coordinate, (GCC), that is the ratio between digital numbers in the green channel divided by the sum of the digital numbers in red, green and blue channels, commonly used to represent canopy greenness in field studies.

We also installed a combination of soil moisture and soil temperature probes (exact details in S2), as well as direct measurements of root biomass through six 4.5 cm x 13 cm ingrowth

cores in each mesocosm. Finally, in E1, we measured system gas exchange using a Li-8100A Infrared Gas Analyzer (IRGA) a Li-8100-104 opaque long-term chamber (Li-COR Biosciences, Lincoln, USA) every half-hour, processing these data via standard packages. For more information, see S2. Occasional, unplanned periods of power disruption in both experiments prevented data collection and image capture but did not affect the subsequent minirhizotron images.

Field Trials

The third (E3) and fourth trial (E4) were in field settings. Experiment 3 (E3) was conducted at Majadas de Tiétar,(Spain) in a Mediterranean wood-pasture (MDT, El-Madany et al., 2018; Nair et al., 2019). We deployed eight RMRs with specification as in E2 powered by per-instrument solar panels coupled with an external 12V battery and charge controller. We sampled twice daily from October 2019 until January 2020 (~ 22800 images per instrument). MDT is a Mediterranean ecosystem where the growing season lasts from autumn until late spring, but much undecomposed root litter remains after a dry summer (Nair et al., 2019). The soil is an Abruptic Luvisol with a sandy upper layer and a thick clay layer starting at 20-40 cm. In this first field trial, we encountered a rare technical issue with the BIOS clock when night temperatures fell below 0°C affecting accurate timekeeping (and hence, no time reference for data collected) at all subsequent sampling points. We hence show summary data from two instruments where timekeeping was not disrupted across all 4 months and a third composite instrument made from two instruments with partial timeseries (i.e. only 'true' time referenced data used, resulting in 3 instruments over 80 % of the timeseries, gapfilled via linear interpolation for periods missing data). We pair this data with site-level PhenoCam GCC from previous work at the site (Luo et al., 2018). Experiment 4 (E4) was conducted in some plots of the Jena biodiversity Experiment (Roscher et al., 2004; Weisser et al., 2017) in Jena, Germany from 1 Feb to 18 March 2022 where temperatures were low (mean air temp. 0.1 °C, min -7.5 °C). Here two instruments ran from mains power. This site is a loamy Eutric Fluvisol in the floodplain of the river Saale. The RMR here had some minor modifications to ensure correct timekeeping at low temperatures and we replaced the camera used in E2 and E3 with another model (S1). We further tested this robustness to low temperatures in -20°C cold room (not shown). In E3 and E4 observatories were installed at 40°, E3 in May 2015 and E4 in August 2020. In this manuscript we show root cover averaged over the whole depth, which reached 40 cm underground. A summary of objectives and differences between E1-E4 are in the supplementary material Table S1.

Automatic Processing of Minirhizotron Imagery

We used the percentage cover of roots (root area / total area) in whole images as our metric for analysis, similar to 'root density' from minirhizotron studies (e.g., Liedgens & Richner, 2001; Nair et al., 2019). We did this for several reasons. First, we were using GCC, a coarse image index, for above-ground timeseries and considered another coarse metric, in contrast to pixel-level consistency, suitable for comparison. Second, it was easier to produce good validation set without needing full or connective annotation of individual roots or complete accuracy in root diameters. After considering root cover between images in E1 and E2, we realised that horizontal installation meant that images on the top of the tube were different than at the bottom (see S6). Thereafter, we treated the top of the tube as 'truth' and used this 3/8 of a complete rotational series for further analyses. In E3/4, where the instruments were deployed in a conventional fashion in the field, we used the sides of the instrument (excluding 1/4 of images at on the top and bottom of the tube) for analysis and further filtered segmented objects with an area of less than 1000 pixels. Because of the camera change (S1), images in E2 and E3 were challenging to both human annotations and CNN segmentation due to low contrast and poor focus. In addition, E3 had high litter content, leading to ambiguity between roots and litter and rapid soil appearance changes as the ecosystem was released from summer drought. All pre-processing (image cropping to remove overlap) was conducted in Python 3. For root image segmentation (i.e. separation of RMR image pixels into two classes, roots and soil), we trained a CNN with a Graphical User Interface, allowing corrective annotation (*Rootpainter*, Smith et al., 2020). We started with a random model and without any pre-annotated images, training on 308 (E1), 300 (E2), 400 (E3) or 350 (E4) complete 2292 x 1944 images. Training data is a major issue for automatic minirhizotron studies which have only short intervals between images but may have sudden changes (e.g. changes in soil colour following rain). Random annotation is unlikely to capture these events and targeting these times introduces bias. Thus, high throughput of training data via corrective annotation / active learning strategies (e.g. Beluch et al., 2018; Budd et al., 2021; Ren et al., 2021) offers a major advantage in such unpredictable tasks. We trained for approximately two person-days per experiment. Details of this training are in S3. These were validated against 384/240/140/90 manually annotated images. For E3 and E4 these were not full pixel-level annotation but a 'fast-pass' annotation (max 15 min annotation time per image) deemed suitable for high throughput timeseries and compared at image level, i.e. we did not assess exact pixel matches. In field datasets, E3 and E4, we

245 applied an additional filter on segmented images, removing all contiguous objects with an
246 area of less than 1000 pixels which we treated as noise in the segmentation.

247 ## Data analysis

248 We took a daily average across all valid images for a single 'root cover' value. For
249 timeseries, we used a further three day rolling average in common with field above-ground
250 phenology approaches (Aasen et al., 2020; Migliavacca, Galvagno, et al., 2011). To
251 calculate growth rates in E1 and E2, we took the linear slope in this average over a five-day
252 window. We did this because we observed a weak sub-daily cycle in the mesocosm
253 experiments. This was a segmentation artefact which affected root identification, particularly
254 on the sides of the observatory not caught by our validation procedure (see S7). Because
255 biomass should be almost stable on these timescales, this artefact clearly affected
256 segmentation, rather than what actually *was* roots in the soil. To analyse overall model
257 accuracy and usefulness to replace manual annotation, we used reduced major axes
258 regression analyses in the *lmodel2* package (Legendre, 2018) in R (R Core Team, 2018).
259 With this analysis we calculate the determination coefficient and the slope of observed (x,
260 annotation) vs estimated (y, CNN root cover) root fraction in the imagery, accounting for
261 similar magnitude of the error in x and y. Additionally, we quantified if systematic bias was
262 introduced by confounding variables such as time, soil moisture, and absolute roots in the
263 image by comparing these data to CNN root cover and checking for significant linear trends.
264 We worked with normalized (0 to 1) timeseries of root cover to account for variation between
265 instruments. To analyse the effect of mesocosm conditions on CO_2 efflux in E1, we fit a
266 GAM (generalized additive model (Hastie & Tibshirani, 1986; Wood, 2006), implemented via
267 *mgcv* (Pedersen et al., 2019), which allowed an independent, non-linear smooth to be fit per
268 predictor (e.g. soil respiration could occur due to more biomass (positive slope) or turnover
269 (negative slope)). For predictors, we used normalized (0 – 1) soil moisture content,
270 temperature, 5-day slope of normalized root cover and normalized GCC and the mean
271 normalized 'biomass index' (mean of normalized GCC and normalized root cover, due to
272 high concurvity), fitting a univariate smooth for each without interactions. Otherwise,
273 concurvity in all cases was less than 0.7. We used the restricted maximum likelihood
274 (REML) method to estimate smooths to reduce overfitting. We compared the variable
275 importance in these models using the *varIMP* function in the *caret* package (Kuhn, 2008).
276 For overall timeseries in E3 and E4, we gap-filled an aggregate timeseries level accounting

277 for the occasional missing cycle via linear interpolation between successive points. The
278 modelling analysis above used only the non-gap filled datapoints.

279

RESULTS

Experiment 1: Biological Interpretation

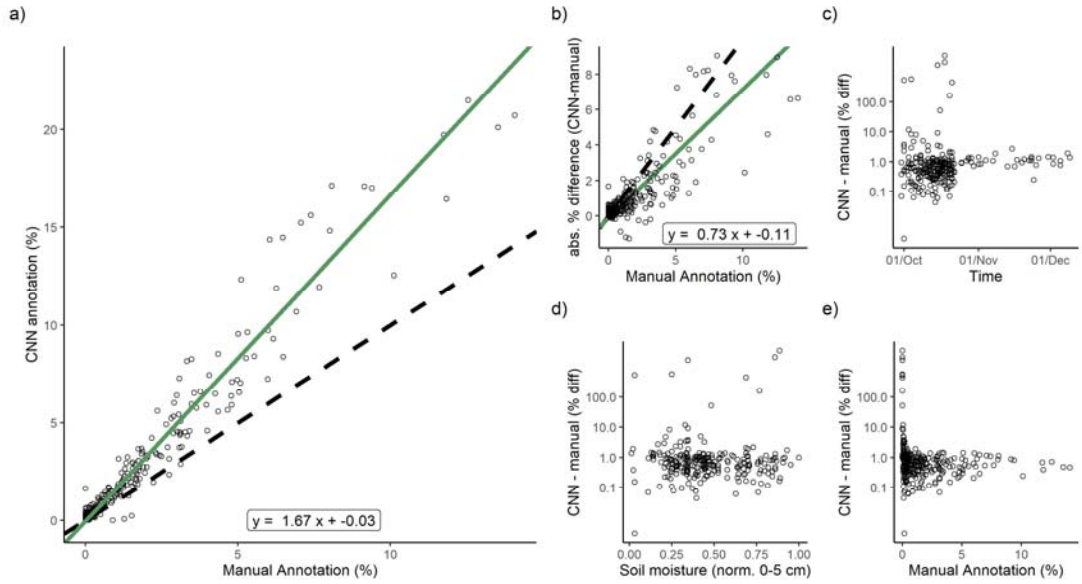

Figure 2. a) Validation of CNN-annotation against independent manual root annotation for root pixels. The green line shows a reduced major axis regression of the two variables, while the dashed line is a 1:1 relationship. Statistics show equation of significant RMA fit. b) absolute percentage difference between manual annotation and CNN-classified pixels, and relative percent difference standardised to manual annotation over c) time, d) soil moisture content in the 0-5 cm soil at time of sampling and e) manual cover. The increased absolute error at higher manual cover was expected. The clustering of data towards the early part of panel c) is due to an uneven validation dataset.

In E1, CNN segmentation corresponded well to manual annotation (R^2 = 95%, Fig. 2a). However, the CNN identified significantly more pixels as roots (slope = 1.67, $P < 0.01$) than humans. While there was an expected relationship between manual cover and absolute difference between CNN and manual mark-up (bigger values could be more wrong, Fig. 2b), there was no effect of time, soil moisture content or manual cover on the error relative to the manual cover (Fig. 2c, 2d, 2e). Hence, we could trust the segmentation on average over the whole time series in terms of dynamics, and if adjustmed by a linear transformation, in terms of magnitude.

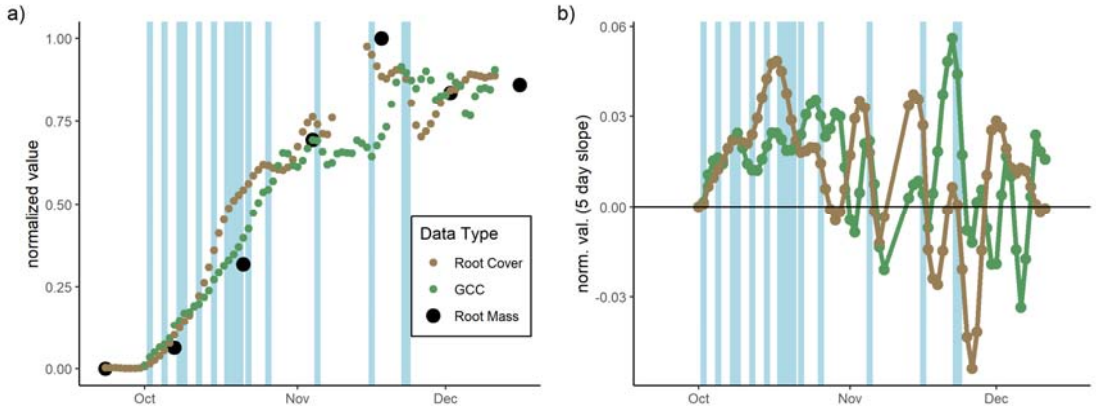

Figure 3. A time series of root data from the mesocosm experiment. Panel a) shows the relative, 3-day smoothed root cover, GCC, and root mass showing parallel development. Panel b) shows the 5-day slope of these values (i.e., rate of change) where at some periods GCC is increasing faster than root cover and at others root cover is increasing faster than GCC. Blue lines show watering events which are scaled in width relative to the volume of water applied.

We also paired the CNN-segmented 'root cover' time series with the GCC and the absolute root mass measurements (Fig. 3a). We found a good match between normalized root mass and root cover (Pearson r = 0.96). When compared to watering, some of the instability in root cover appeared to be related to the watering events late in the mesocosm (Fig 3a), where CNN-root cover dropped following after watering events, a potential effect not caught by comparison with SWC nor visible in human mark-up.

We also compared rates of change ('growth'). In general, this differed between the two indices (Fig. 3b) except at the start of the experiment. As previously mentioned, from the start of November there was a decrease in root cover growth following watering. Root cover growth continued to be positive at the end of the experiment even when the GCC was stable or decreasing as the soil slowly dried out.

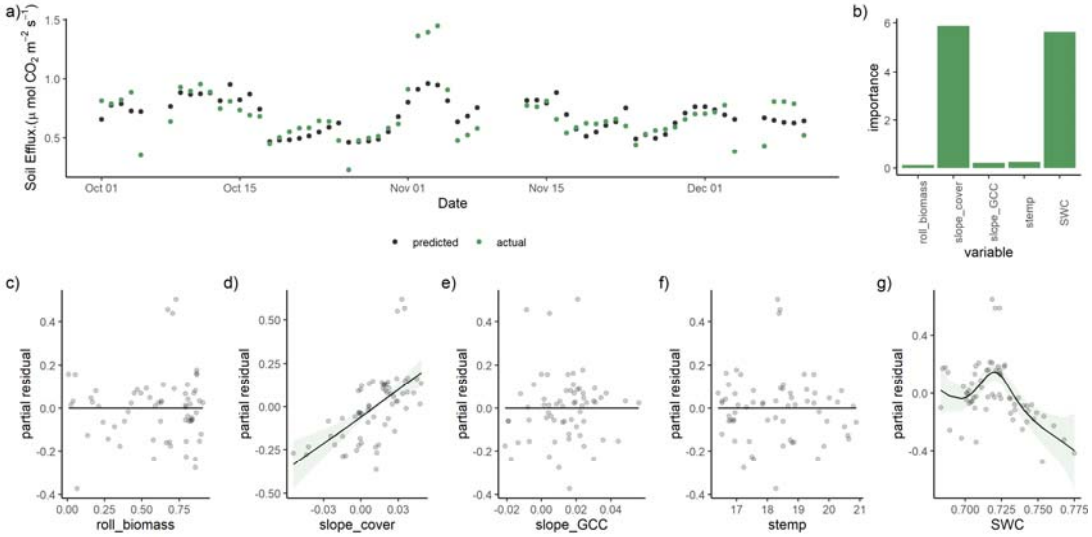

Figure 4. Summary of the GAM model fit for E1. Root cover (roll_cover) correlated with GCC so we used a mean of the two in the main model (roll biomass). We were able to predict daily CO_2 flux fairly well with an R^2 of 46% (panel a). The model contained univariate smooths for rolling biomass (roll_biomass, shown vs partial residuals in panel c), the 5-day slopes (i.e. growth rate) of root cover (slope_cover, panel d) and GCC (slope_GCC, panel e), soil temperature (stemp, panel f) and soil moisture content (SWC, panel g). Shaded area shows 2 x SE. When variable importance was considered (panel b), the slope of root cover (i.e. change in amount of roots, birth or death) was a much better predictor than that of biomass.

Using the GAM on Experiment 1, we were able to explain 46% of the total variation in CO_2 soil efflux with the best fitted model (Fig. 4a). While we did not use a time structure the model performed worse close to data gaps. In this model, soil moisture content ($P < 0.001$) and root growth (slope_cover) ($P < 0.001$) were significant. Root growth (slope_cover) was most important in determining soil CO_2 efflux (Fig. 4b), increasing root growth rate had a positive effect on soil CO_2 efflux (Fig. 4d). In contrast, absolute biomass, and GCC slopes, our proxy for leaf growth rate, did not have an effect (Fig 4c, 4e) on soil CO_2 efflux.

Experiment 2: Instrument Consistency

E2 did not achieve as good a fit against manual annotation in E2 (R^2 = 69 %, with a steeper slope = 2.4, validation in Fig S4). The higher bias compared to E1 might be related to the worse quality of the imagery due to the hardware limitations.

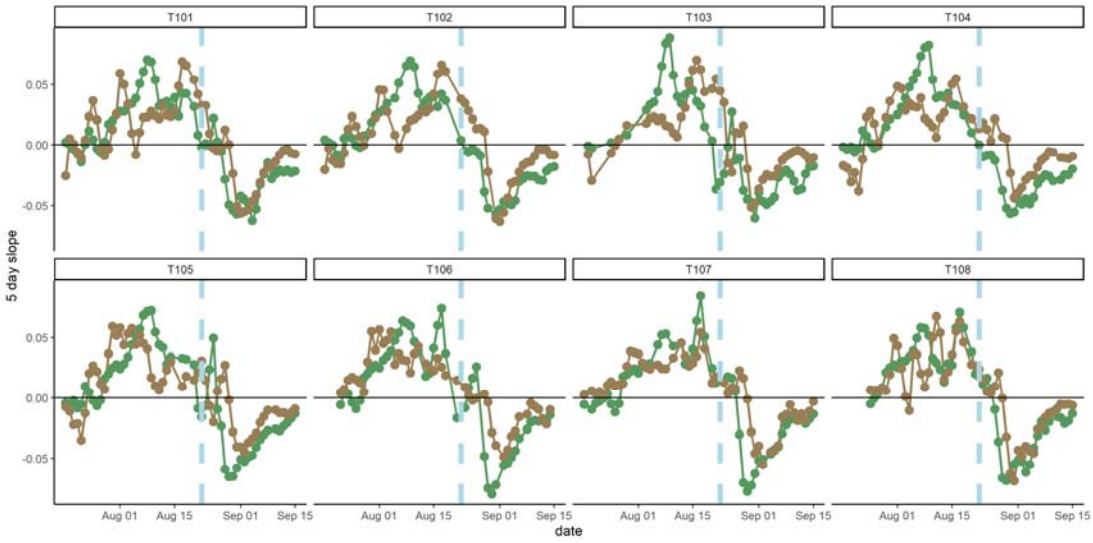

Figure 5. 5-day slope (growth rate) of above- and below-ground indexes across 8 mesocosms. Vertical blue line indicates the last watering date, horizontal black line is 0 and the transition point between positive growth rates (above-) and a decrease in the index (below-) indicating yellowing or disappearance of roots identified by the CNN trained on 'living' roots.

Overall, a similar time series was extractable from each mesocosm (Fig. S5). Correlation between GCC and root cover was between 0.7 and 0.96. Notably, GCC increased faster than root cover in the first four mesocosms. The mesocosms were arranged in numerical order, suggesting that a gradient (e.g., light) within the greenhouse may have driven a difference. Once watering ceased, root cover declined less steeply (Fig. 5) than GCC in all mesocosms and root growth rate remained positive for longer than GCC growth rate (Fig. 5), indicating a continued production of roots even as the above-ground began to yellow.

The roots extracted from ingrowth cores at the start of the experiment matched the overall time series (when normalized between mesocosms, Fig. S5b). However, the last data point, collected at harvest did not, and the measured root biomass was still high when the root cover index had dropped.

Experiment 3: Root Litter Field Trial

The first field trial was conducted in a Mediterranean ecosystem and so the period studied was early in the growing year. Overall, the CNN performed satisfactorily, with an image-level R^2 of 78% against the non-pixel level annotation data (S4) and as before a slight net overprediction. Examining the mean total change in root cover over the whole 40 cm depth

sampled by the minirhizotron tube, we observed growth period starting in November-December, exact dates variable between instruments and after the initial green-up of the above-ground system as detected by GCC (Figure 6). Certain short-term instability in this index appeared to be related to periods of rainfall (as in E1) but in general roots were being produced over the period of data collection. We additionally filtered out a short period (19-21 December) where all images across all instruments were very poorly illuminated. In general, temporal variability and noise in root cover was of similar relative magnitude to site-level GCC. Root growth generally was lagged following positive GCC change, unlike E1, potentially due to the conventional installation angle here and subsequent better coverage of the whole soil. This trial was halted by the failure of accurate timekeeping across all instruments, which we subsequently addressed in E4.

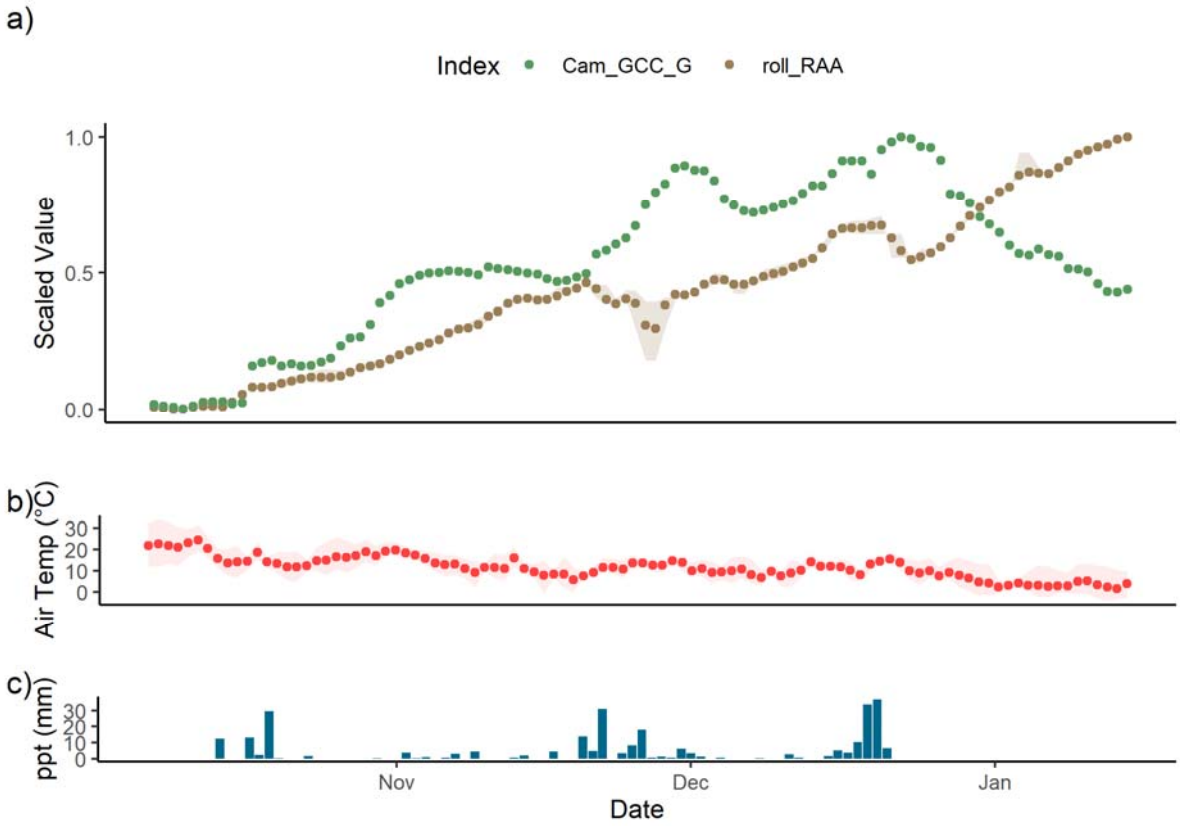

Figure 6. a) 3-Day normalized GCC (green chromatic coordinate, green) and mean root cover (brown) from 3 instruments with at the MDT site. Error bands on root cover correspond to maximum and minimum mean segmented root cover in the aggregation period. The high errors in late November are due to instrument drop-out, reduction to 2 instruments in this period and subsequent instrument replacement, A short period (2 days) is removed in mid-December as batteries failed to charge in very cloudy weather. Root growth began later in the Mediterranean growing season than leaf growth and continued even when GCC was declining in midwinter. There was some instability in the root index which may have followed c) precipitation, although this was not larger than relative instability in GCC.

343

344

345 ## Experiment 4: Low Temperature Field Trial

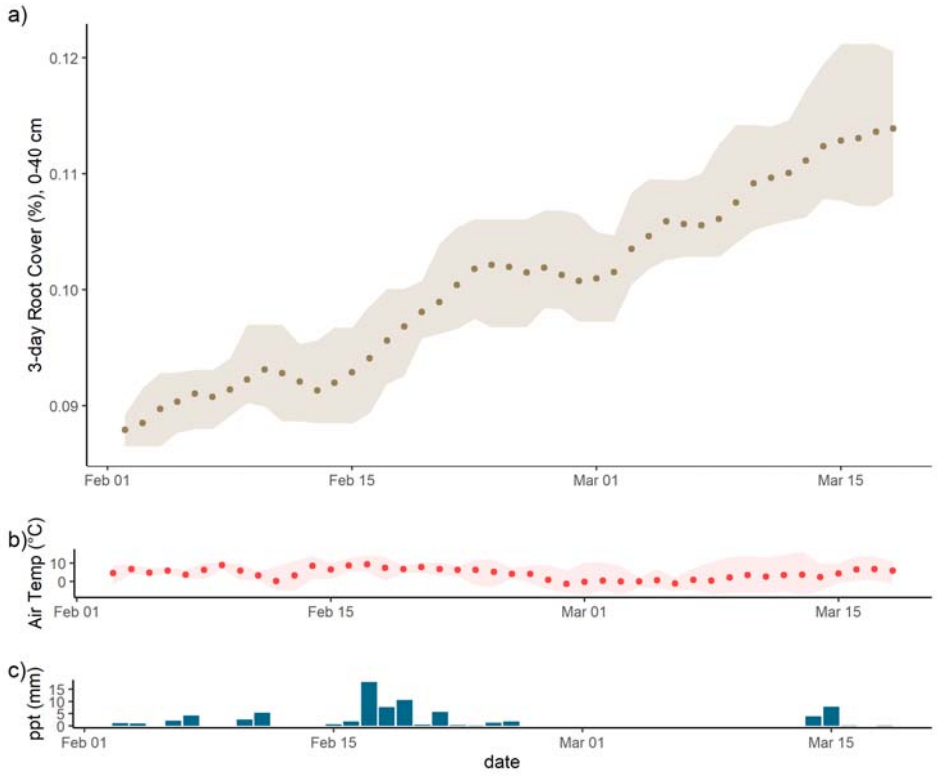

Figure 7. a) 3-Day mean normalized root cover (brown) from 2 instruments at JE showing growth in February and March. Error band shows the maximum and minimum segmented RAA during the aggregation period. This winter trial shows robustness of the RMR to cold temperatures, the mean air temperature in b) is shown with an error band corresponding to maximum and minimum daily temperatures which fell as low as -9°C. Unlike E3, it did not appear that in this experiment there was a sensitivity in root cover to precipitation.

346

347 E4 ran under winter conditions more severe than those which had previously caused
348 timekeeping failures, without issue. Here we achieved agreement between CNN-
349 segmentation and image-level annotation with a R^2 of 85 % (Figure S4) against human
350 observations. We observed root growth through February and March (Figure 7) with periodic
351 increases in root cover and periods of no net growth which did not qualitatively appear to be
352 linked to site conditions. We had occasional unexplained periods where the measurement
353 cycle did not start; this affected ~ 5 % of observational periods with no relationship to
354 temperature or humidity. The images were confounded by condensation and soil animals,
355 which we were largely able to successfully train around and the timeseries consistency (S7)
356 and did not have the same potential vulnerability to soil moisture and precipitation as E1 and

357 E3. There was no relationship between the root cover we segmented over the whole 0-40
358 cm depth sampled and air temperature, precipitation, nor soil moisture, and hence the
359 observations in Figure 7 can be considered real root growth in this time.

DISCUSSION

Use of automatic minirhizotrons

The minirhizotron (MR) technique is the best method to measure phenological properties of fine roots (Freschet et al., 2021). Automation of the whole minirhizotron workflow, from image collection to analysis, is essential to make high time resolution data available. So far there has been no work pairing indexes of root data on the same temporal and spatial scales as above ground vegetation indexes, despite their usefulness for whole system phenology interpretation.

Here, we presented a newly built automatic minirhizotron system. The system is designed for relatively affordable budgets and is robust in adverse field climatic conditions (after modification) and could therefore be replicated. We produced similar data sets from these replicated instruments despite independent sensors (Fig. 4) and produced interpretable and reliable time series in the field (Figure 6, Figure 7). Methods such as PhenoCams (Richardson et al., 2018; Sonnentag et al., 2012) or network for measurements of spectral vegetation indices (Gamon et al., 2015) are standardized for a single instrument, but automatic minirhizotrons must combine robotics and imaging across multiple point sensors, both advancing fast. Imaging targets belowground in other studies may also require different magnification, sampling time, demands on power supplies. Hence, we recommend against standardisation of every aspect of design if time and expertise is available to build minirhizotrons to exact specification and budgets. In particular, cameras could be tooled to specific application, which must in any case be validated against annotated properties of interest for CNN training while retaining fundamental principles of design.

Our system was designed for replication, a key shortcoming of many field studies (Filazzola & Cahill Jr, 2021; Yang et al., 2022), especially acute for minirhizotrons given the short field of view, numerous potential artefacts (Joslin & Wolfe, 1999) and huge range of soil and root appearances. The necessary modifications and COVID-19 led to field trials E3 and E4 not using all instruments from E2 but we observed consistency and reliability in patterns within instruments used. We also fully analysed the suitable data. While we discarded 5/8 (E1/E2) or 1/4 (E3/E4) of the interpreted series due to the angle of the observatory, we covered every sampling timepoint and these represent to our knowledge the highest frequency minirhizotron dataset so far analysed. Data collected at such high frequency have multiple

391 issues in this simple analysis with implications for application of such techniques to more
392 complex situations. The majority of this discussion concerns these issues.

393 Interpreting High Resolution Root Image Data

394 Neural-network methods, like many other root phenotyping applications (Atkinson et al.,
395 2019) are an obvious approach to analysing minirhizotron images and necessary to exploit
396 the high volume of imagery from automated sampling. The main objective here was to pair
397 existing methods with our automated sampling, and while we did not develop a new CNN
398 method, we applied an established segmentation algorithm (Smith et al., 2020) for root
399 images, chosen because of the simple GUI for corrective annotation. In general, this method
400 performed well (Fig. 2, Figure S4) despite the complexity of our images and the lack of time
401 series consideration. We chose the root pixel cover metric for comparison because of the
402 aims of comparing with a similarly coarse above-ground index. More interpretive root
403 phenotyping (e.g. Seethepalli et al., 2021) can also be paired to this CNN, and may be
404 possible from automated measurements with more thorough training, although this was not
405 the aim of this study.

406 Automatic root identification necessary from automated sampling are demanding on training
407 data but also reliant on validation. We used annotated images produced for other CNN trials
408 as validation data (E1/E2) and a fast annotation not true to pixel level (E3/E4) so especially
409 for these latter trials some difference between segmented and annotated images should be
410 expected. Minirhizotron image annotations are human-best practice identification and not a
411 direct property of study systems and may be outperformed in both consistency and accuracy
412 in identifying true properties of interest by an appropriately trained CNN. For field studies,
413 where above-ground indexes are similarly coarse in space and time (e.g. 3-day averaging
414 phenocam data (Aasen et al., 2020)), a similar accuracy as experiment 1 (96%) or
415 experiment 4 (85 %) even experiment 2 (72%) or 3 (78%) with the less powerful camera is
416 sufficient for generating plausible root time series (e.g. Fig 3, Fig S5, Fig 6, .Fig 7.).

417 In the mesocosm E1, we found a good agreement in physical validation between root mass
418 and minirhizotron root cover (Fig. 4a, r = 0.96), and also in the growth period of Experiment
419 2 (but not the last, replicated measurement, discussed later). We also found a reasonable
420 correlation (E1 95 % and E2: 70 % respectively) between CNN and manual annotation.
421 While in all four cases the CNN segmentation overestimated root cover compared to
422 validation (Fig 2, Fig S4), we were not validating nor training for true pixel-level identification

423 because the training data from the GUI method was not annotated in the same way as the
424 full image validation data. In any case, human root image interpretation is potentially biased
425 by annotator and consistent CNN interpretation has major advantages for throughput for root
426 dynamics.

427 While mesocosms are *not* good replacements for field root measurements (Poorter et al.,
428 2012) and we did not validate E3 or E4 against direct root measurements, some further
429 consideration is due to using minirhizotrons for dynamic biomass pools. There are
430 differences in properties measurable from minirhizotrons and other methods (e.g. Addo-
431 Danso et al., 2016; Milchunas, 2009; Nair et al., 2019), but minirhizotrons are a reasonable
432 index of biomass with appropriate conversions used (e.g. Brown et al., 2009; Johnson et al.,
433 2001; Lee et al., 2017; Sullivan & Welker, 2005). Indeed, root identification relates directly to
434 biological structures in a close to 2D surface around a minirhizotron observatory. This can
435 potentially be paired more easily with density and C contents than leaf greenness of a whole
436 canopy, if the short range of the instruments can be offset by replication. A key aspect of
437 future studies will be to ascertain whether well-known artefacts of roots around
438 minirhizotrons (Rytter & Rytter, 2012; Vamerali et al., 2009) can be understood sufficiently to
439 correct for these issues, and indeed if these artefacts vary with phenology or environmental
440 conditions.

441 A further issue with our segmentation of high frequency images related to unrealistic
442 variation after rain/watering (Fig. 3, Fig 6) and in E1 and E2, between different hours of the
443 day (Fig. S2). Like similar issues using classifiers for NIR-enabled minirhizotron images
444 (Svane et al., 2019), sudden changes in soil colour/reflectance led to potential erroneous
445 pixel identification and subsequent unrealistic changes in roots identified. This may affect
446 short term biomass change, but we note that in Fig 6 this instability is not larger than similar
447 short-term patterns in greenness above-ground, potentially due to illumination conditions.
448 For the sub-daily artefact, this was not explainable by soil moisture from the surface sensors
449 (Fig. 2c), perhaps because very local condensation at minirhizotron surfaces was not
450 represented. Indeed, sub-daily error only became apparent after roots had colonized the
451 sides of the observatories, suggesting that the CNN was misidentifying pixels in close
452 proximity to roots. Inspection of the images confirms this explanation (Fig. S7). Diel variation
453 in root diameter (Huck et al., 1970) or hydraulic redistribution (passive movement of water
454 via roots from areas of wet to dry soil (Ryel, 2004)) at night and transpiration during the day
455 drying these areas may provide an explanation which could bias a CNN more than human

annotation. If this was a segmentation artefact which affected the immediate 'rhizosphere' only, this would not have been detected by our soil moisture sensors.

There are several potential solutions. The first is to aggregate, smooth or throw out data to remove such effects, which disregards information from high resolution sampling but is common practice at ecosystem level e.g. using smoothing splines (Migliavacca, Galvagno, et al., 2011). However, if one wants to analyse sub-daily data this is not a viable solution unless one uses time series decomposition (e.g. Biriukova et al., 2021). Secondly, one could post-process with human intervention, applying an adjustment or a separately trained model to periods with problematic changes. Such an approach needs to interfere with the image index so is not favourable. Thirdly, one could train on more data, particularly around periods of difficulty. This exacerbates training data limits, especially if randomly selecting data or if problematic events are rare but important. Fourthly, consistency in segmentation between sequential images could be used to filter for 'true' observations, if true changes during a period of instability are unidirectional (i.e. basing interpretation on objective priors about root growth). Finally, a wholly different model structure incorporating either other factors such as soil moisture or rainfall (i.e. describing when these issues could occur) or time series information (e.g. via recurrent network architectures such as Long-Short Term Memory approaches (Hochreiter & Schmidhuber, 1997), multivariate time series classification (Ruiz et al., 2021) or other time series classification approaches (Fawaz et al., 2019)) could be built. This has the advantage of processing data without bias (if training data is selected fairly) at the cost of a more complex model and/or more variables to measure alongside root imagery.

Field Robustness of Our Techniques

Both automation of measurement and image analysis are more advanced in simple artificial systems than field measurements, which are complicated by soil appearance, litter, soil animals, and hydrology. Instruments which run in controlled environments such as the greenhouse in E1 and E2 and analysis methods also need to be robust in the field to study phenology. In the two short field studies we show here we demonstrate viability of the technique; while E3 was compromised by camera quality and the structural complex system, we were able to produce a time series both plausible from previous work at the site –early season root growth has been previously shown to continue through during winter, unlike above-ground vegetation indexes in most years (Nair et al., 2019) - and reasonable against human mark-up (S4). Here the instruments operated consistently on solar power for 4

489 months with an overall 95 % uptime. While we do not show a whole phenological year and
490 do not include root death in the summer drought, this period contained a large amount of
491 undecomposed root litter due to the arid summer, confusing for both human annotators and
492 the CNN. Difficult minirhizotron images are troublesome for even reasonably experienced
493 annotators (Peters et al., 2022) - a key advantage of an automated approach is consistency
494 (given learned features) across a whole dataset. The target for our approach was a robust
495 index comparable to above-ground digital repeat photography (e.g. Migliavacca, Galvagno,
496 et al., 2011; Sonnentag et al., 2012) rather than exact matches of segmentation to
497 annotation. A key difference due to the scale disparity is necessity of segmenting roots from
498 soil rather than a simple image index such as 'greenness' on a defined region of interest.
499 Because we were segmenting complex features, minor variations in images setting (e.g. due
500 to soil water, soil animals) may introduce short term variation in the segmentation; indeed it
501 seemed, if no root death is assumed from E4 (which did not appear to have a rainfall bias,
502 perhaps due to the shorter dataset with more wet weather, thus easier to train in rain events)
503 a small systematic error) may have arisen over the timeseries (e.g. the drop around Feb13).
504 Future efforts could benchmark an acceptable level of consistency for representativeness of
505 phenology rather than requiring the same level of accuracy on a high-resolution dataset as
506 expected in a finer scale analysis also achievable from minirhizotrons, for example root
507 architectural traits.

508 The E3 field trial revealed an unexpected and random vulnerability of BIOS timekeeping to
509 low temperatures which we corrected by adding a GPS clock. In E4 the system was able to
510 run successfully at night-time temperatures of as low as -7.5°C without issues. In this
511 relatively 'easy' (stone-free and loamy) soil we were able to achieve a much closer match
512 with manual mark-up – even using our 'root cover' comparison, which could be expected to
513 be more variable due to the contribution of errors in root volume which would not be seen if
514 we compared root lengths. We could train around soil animals and condensation
515 (quantitatively assessed) in this wet and cold part of the year and indeed found that roots
516 were growing even in this temperate winter with low air temperatures for an increase of ~ 12
517 % in total root area observed over the six weeks of the trial. Indeed, there were periods of
518 net increase in root cover of several days interspersed with periods without net growth. This
519 did not appear to be related to meteorological conditions, suggesting that i) these conditions
520 were unlikely to be seriously biasing our CNN annotation and ii) this early-season growth
521 may be driven by intrinsic cues rather than photosynthesis. Longer timeseries enabled by
522 such devices can enable a deeper understanding of the coupled above-belowground action

of the carbon cycle in plants. While we note our field indexes (especially E3) could be unstable, this was due to the low replication and differences between instruments rather than individual timeseries inconsistency. Wider application of such devices is hence reliant on reasonable per-instrument costs, the rationale for development of our instrument to start with.

Relevance for Ecology and Carbon Cycling Knowledge

While the experiments described herein were to test the RMR, and caution against over-interpretation of our field trials due to the lower than desired replication caused by events outside our control, we show the potential of these kind of high-resolution root measurements. In E1 we observed short-term de-synchronicity from our two image indexes above- and below- ground, indicating that development of roots and shoots was not synchronous, although the peak was reached at similar times (Fig. 4b, Fig. 7). This matches some conventional minirhizotron studies which find asynchrony in non-woody vegetation (Sloan et al., 2016; Steinaker et al., 2010; Steinaker & Wilson, 2008). In E3 we noticed a similar phenomenon in relative directions of the two indexes which may relate to either patterns of assignment or a systematic artefact and should be investigated further. We note that E1 was not a field experiment and we were not measuring true phenology, rather development from seed, and E3 was only one part of the overall phenological year and so care must be taken in applying any conclusions to a wide scale. Nonetheless, the decoupling between GCC and roots was very consistent between instruments during the dry down in E1 and the overall pattern in E3 was similar to previous manual minirhizotron observations at the site (Nair et al., 2019) when compared to GCC. In E1, Asynchrony was also happening on an extremely short scale of several days. Such fine distinctions potentially have impact on interpretation of C uptake and respiration in field experiments and should be tested in these contexts. From modelling E1, in our system change in roots also were a better indicator of soil CO_2 evolution than leaf greenness which may indicate that link between photosynthesis and soil/ecosystem respiration (Bahn et al., 2008, 2009; Migliavacca, Reichstein, et al., 2011) was driven by root growth. In other words, growth respiration was more important than maintenance respiration for variation in CO_2 efflux measured. In terms of general trends (but not absolute mass), our root cover index was also a reasonable proxy for root biomass so with appropriate calibration automatic minirhizotron imagery could be interpreted as changes in C pools and provide data to inform allocation in vegetation models. Interestingly the ‘harvest’ data point in E2 did not match the minirhizotron

index, even when scaled. This is potentially explainable by the CNN, trained on live roots, outperforming humans sorting physical samples. Because the living/dead distinction is of functional interest, further development e.g. training models specifically on dead roots, using multi-classifier segmentation or generating labels using NIR or other spectral images (Arnold et al., 2017; Bodner et al., 2017) and segmenting using RGB images may allow a segmentation of field datasets necessary for functional quantification of root biomass and its complex and dynamic contribution to ecosystem C cycling.

ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 748893 (Marie Sklodowska-Curie Individual Fellowship to R. Nair). This research also received funding through the Alexander von Humboldt Foundation and the Max Planck Research Prize 2013 to M. Reichstein. VR was supported by the regional government of Extremadura (Spain) through a "Talento" fellowship (TA18022). We also thank Y. Luo for processing the phenocam-GCC data at MDT, A. Smith for development and user-friendly implementation of the *Rootpainter* CNN, M.Schrumpf for her very useful comments on this manuscript and M. Reichstein for his support for the project in general.

AUTHORS CONTRIBUTIONS

RN and MM conceived the study. RN led all data acquisition and analysis. Field work was supported by VR. MS, MH and OK developed and tested the instrument. Manuscript preparation was led by RN with major input from MM and VR. All authors gave final approval for publication.

DATA AVAILABILITY STATEMENT

A dataset with validation data used in this study, along with a small dataset of unannotated images will be uploaded to Zenodo.org on final publication. The full image dataset is not sharable using current services due to its size, but the authors are happy to share this to developers of root segmentation tools on request.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

BIBLIOGRAPHY

Allen, M. F., & Kitajima, K. (2014). Net Primary Production of Ectomycorrhizas in a California Forest. *Fungal Ecology*, *10*(1), 81–90. https://doi.org/10.1016/j.funeco.2014.01.007

Allen, M. F., Vargas, R., a. Graham, E., Swenson, W., Hamilton, M., Taggart, M., Harmon, T. C., Rat'Ko, A., Rundel, P., Fulkerson, B., & Estrin, D. (2007). Soil Sensor Technology: Life within a Pixel. *BioScience*, *57*(10), 859. https://doi.org/10.1641/B571008

Arnold, T., Leitner, R., & Bodner, G. (2017). Near infrared hyperspectral imaging system for root phenotyping. *Sensing for Agriculture and Food Quality and Safety IX*, *10217*, 94–99. https://doi.org/10.1117/12.2262441

Atkinson, J. A., Pound, M. P., Bennett, M. J., & Wells, D. M. (2019). Uncovering the hidden half of plants using new advances in root phenotyping. *Current Opinion in Biotechnology*, *55*, 1–8. https://doi.org/10.1016/j.copbio.2018.06.002

Blume-Werry, G., Jansson, R., & Milbau, A. (2017). Root phenology unresponsive to earlier snowmelt despite advanced above-ground phenology in two subarctic plant communities. *Functional Ecology*, *31*, 1493–1502. https://doi.org/10.1111/1365-2435.12853

Bodner, G., Alsalem, M., Nakhforoosh, A., Arnold, T., & Leitner, D. (2017). RGB and Spectral Root Imaging for Plant Phenotyping and Physiological Research: Experimental Setup and Imaging Protocols. *Journal of Visualized Experiments : JoVE*, *126*, 56251. https://doi.org/10.3791/56251

608 Brown, A. L. P., Day, F. P., & Stover, D. B. (2009). Fine Root Biomass Estimates from
609 Minirhizotron Imagery in a Shrub Ecosystem Exposed to Elevated CO2. *Plant and*
610 *Soil*, *317*(1–2), 145–153. https://doi.org/10.1007/s11104-008-9795-x
611 Cai, G., Vanderborght, J., Klotzsche, A., van der Kruk, J., Neumann, J., Hermes, N., &
612 Vereecken, H. (2016). Construction of Minirhizotron Facilities for Investigating Root
613 Zone Processes. *Vadose Zone Journal*, *15*(9), 0.
614 https://doi.org/10.2136/vzj2016.05.0043
615 Delory, B. M., Baudson, C., Brostaux, Y., Lobet, G., du Jardin, P., Pagès, L., & Delaplace, P.
616 (2016). ArchiDART: An R Package for the Automated Computation of Plant Root
617 Architectural Traits. *Plant and Soil*, *398*(1–2), 351–365.
618 https://doi.org/10.1007/s11104-015-2673-4
619 Dijkstra, F. A., Zhu, B., & Cheng, W. (2021). Root effects on soil organic carbon: A double-
620 edged sword. *New Phytologist*, *230*(1), 60–65. https://doi.org/10.1111/nph.17082
621 Filazzola, A., & Cahill Jr, J. F. (2021). Replication in field ecology: Identifying challenges and
622 proposing solutions. *Methods in Ecology and Evolution*, *12*(10), 1780–1792.
623 https://doi.org/10.1111/2041-210X.13657
624 Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, *1*(3),
625 297–310. https://doi.org/10.1214/ss/1177013604
626 Herrmann, S., Grams, T. E. E., Tarkka, M. T., Angay, O., Bacht, M., Bönn, M., Feldhahn, L.,
627 Graf, M., Kurth, F., Maboreke, H., Mailander, S., Recht, S., Fleischmann, F., Ruess,
628 L., Schädler, M., Scheu, S., Schrey, S., & Buscot, F. (2016). Endogenous rhythmic
629 growth, a trait suitable for the study of interplays between multitrophic interactions
630 and tree development. *Perspectives in Plant Ecology, Evolution and Systematics*, *19*,
631 40–48. https://doi.org/10.1016/j.ppees.2016.02.003

632 Huo, C., & Cheng, W. (2019). Improved root turnover assessment using field scanning
633 rhizotrons with branch order analysis. *Ecosphere*, *10*(8), e02793.
634 https://doi.org/10.1002/ecs2.2793
635 Iversen, C. M., Murphy, M. T., Allen, M. F., Childs, J., Eissenstat, D. M., a. Lilleskov, E.,
636 Sarjala, T. M., Sloan, V. L., & Sullivan, P. F. (2011). Advancing the Use of
637 Minirhizotrons in Wetlands. *Plant and Soil*, *352*, 23–39.
638 https://doi.org/10.1007/s11104-011-0953-1
639 Johnson, M. G., Tingey, D. T., Phillips, D. L., & Storm, M. J. (2001). Advancing fine root
640 research with minirhizotrons. *Environmental and Experimental Botany*, *45*(3), 263–
641 289. https://doi.org/10.1016/S0098-8472(01)00077-6
642 Joslin, J. D., & Wolfe, M. H. (1999). Disturbances During Minirhizotron Installation Can Affect
643 Root Observation Data. *Soil Science Society of America Journal*, *63*(1), 218–221.
644 https://doi.org/10.2136/sssaj1999.03615995006300010031x
645 Lee, C. G., Suzuki, S., Noguchi, K., & Inubushi, K. (2017). Estimation of Fine Root Biomass
646 Using a Minirhizotron Technique among Three Vegetation Types in a Cool-
647 Temperate Brackish Marsh. *Soil Science and Plant Nutrition*, *62*(5–6), 465–470.
648 https://doi.org/10.1080/00380768.2016.1205957
649 Legendre, P. (2018). *lmodel2: Model II Regression* (1.7-3) [Computer software].
650 https://CRAN.R-project.org/package=lmodel2
651 Luo, Y., El-Madany, T., Ma, X., Nair, R., Jung, M., Weber, U., Filippa, G., Bucher, S. F.,
652 Moreno, G., Cremonese, E., Carrara, A., Gonzalez-Cascon, R., Escudero, Y. C.,
653 Galvagno, M., Pacheco-Labrador, J., Martín, M. P., Perez-Priego, O., Reichstein, M.,
654 Richardson, A. D., … Migliavacca, M. (2020). Nutrients and water availability
655 constrain the seasonality of vegetation activity in a Mediterranean ecosystem. *Global*
656 *Change Biology*, *26*(8), 4379–4400. https://doi.org/10.1111/gcb.15138

Luo, Y., El-Madany, T. S., Filippa, G., Ma, X., Ahrens, B., Carrara, A., Gonzalez-cascon, R., Cremonese, E., Galvagno, M., & Tiana, W. (2018). Using Near-Infrared-Enabled Digital Repeat Photography to Track Structural and Physiological Phenology in Mediterranean Tree – Grass Ecosystems. *Remote Sensing*, *10*, 1293. https://doi.org/10.3390/rs10081293

Nair, R. K. F., Morris, K. A., Hertel, M., Luo, Y., Moreno, G., Reichstein, M., Schrumpf, M., & Migliavacca, M. (2019). N:P stoichiometry and habitat effects on Mediterranean savanna seasonal root dynamics. *Biogeosciences*, *16*(9), 1883–1901. https://doi.org/10.5194/bg-16-1883-2019

Nijland, W., de Jong, R., de Jong, S. M., Wulder, M. A., Bater, C. W., & Coops, N. C. (2014). Monitoring plant condition and phenology using infrared sensitive consumer grade digital cameras. *Agricultural and Forest Meteorology*, *184*, 98–106. https://doi.org/10.1016/j.agrformet.2013.09.007

Pedersen, E. J., Miller, D. L., Simpson, G. L., & Ross, N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ*, *7*, e6876. https://doi.org/10.7717/peerj.6876

Poorter, H., Bühler, J., van Dusschoten, D., Jos Climent, & Postma, Johannes A. (2012). Pot size matters: A meta-analysis of the effects of rooting volume on plant growth. *Functional Plant Biology*, *39*, 839–850. https://doi.org/10.1071/FP12049.

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*.

Radville, L., McCormack, M. L., Post, E., & Eissenstat, D. M. (2016). Root Phenology in a Changing Climate. *Journal of Experimental Botany*, *67*(12), erw062-. https://doi.org/10.1093/jxb/erw062

Rahman, G., Sohag, H., Chowdhury, R., Wahid, K. A., Dinh, A., Arcand, M., & Vail, S. (2020). SoilCam: A Fully Automated Minirhizotron using Multispectral Imaging for Root Activity Monitoring. *Sensors*, *20*(3), 787. https://doi.org/10.3390/s20030787

683 Rahmanzadeh, H., & Shojaedini, S. V. (2016). *Novel Automated Method for Minirhizotron*
684 *Image Analysis: Root Detection Using Curvelet Transform*. *29*(03), 337–346.
685 Raupach, M. R., Canadell, J. G., Ciais, P., Friedlingstein, P., Rayner, P. J., & Trudinger, C.
686 M. (2011). The Relationship between Peak Warming and Cumulative CO2
687 Emissions, and Its Use to Quantify Vulnerabilities in the Carbon-Climate-Human
688 System. *Tellus, Series B: Chemical and Physical Meteorology*, *63*(2), 145–164.
689 https://doi.org/10.1111/j.1600-0889.2010.00521.x
690 Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M.,
691 Johnston, M. R., Keenan, T. F., Klosterman, S. T., Kosmala, M., Melaas, E. K.,
692 Friedl, M. A., & Frolking, S. (2018). Tracking Vegetation Phenology across Diverse
693 North American Biomes Using PhenoCam Imagery. *Scientific Data*, *5*, 1–24.
694 https://doi.org/10.1038/sdata.2018.28
695 Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M.
696 (2013). Climate change, phenology, and phenological control of vegetation
697 feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173.
698 https://doi.org/10.1016/j.agrformet.2012.09.012
699 Rytter, R.-M., & Rytter, L. (2012). Quantitative estimates of root densities at minirhizotrons
700 differ from those in the bulk soil. *Plant and Soil*, *350*(1), 205–220.
701 https://doi.org/10.1007/s11104-011-0896-6
702 Sloan, V. L., Fletcher, B. J., & Phoenix, G. K. (2016). Contrasting synchrony in root and leaf
703 phenology across multiple sub-Arctic plant communities. *Journal of Ecology*, *104*(1),
704 239–248. https://doi.org/10.1111/1365-2745.12506
705 Smith, A. G., Han, E., Petersen, J., Olsen, N. A. F., Giese, C., Athmann, M., Dresbøll, D. B.,
706 & Thorup-Kristensen, K. (2020). RootPainter: Deep Learning Segmentation of
707 Biological Images with Corrective Annotation. *BioRxiv*, 2020.04.16.044461.
708 https://doi.org/10.1101/2020.04.16.044461

709 Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H.,
710 Milliman, T., O'Keefe, J., & Richardson, A. D. (2012). Digital Repeat Photography for
711 Phenological Research in Forest Ecosystems. *Agricultural and Forest Meteorology*,
712 *152*(1), 159–177. https://doi.org/10.1016/j.agrformet.2011.09.009
713 Steinaker, D. F., Wilson, S. D., & Peltzer, D. A. (2010). Asynchronicity in Root and Shoot
714 Phenology in Grasses and Woody Plants. *Global Change Biology*, *16*(8), 2241–2251.
715 https://doi.org/10.1111/j.1365-2486.2009.02065.x
716 Sullivan, P. F., & Welker, J. M. (2005). Warming chambers stimulate early season growth of
717 an arctic sedge: Results of a minirhizotron field study. *Oecologia*, *142*(4), 616–626.
718 https://doi.org/10.1007/s00442-004-1764-3
719 Svane, S. F., Dam, E. B., Carstensen, J. M., & Thorup-Kristensen, K. (2019). A multispectral
720 camera system for automated minirhizotron image analysis. *Plant and Soil*, *441*,
721 657–672. https://doi.org/10.1007/s11104-019-04132-8
722 Taylor, B. N., Beidler, K. V., Strand, A. E., & Pritchard, S. G. (2014). Improved Scaling of
723 Minirhizotron Data Using an Empirically-Derived Depth of Field and Correcting for the
724 Underestimation of Root Diameters. *Plant and Soil*, *374*(1–2), 941–948.
725 https://doi.org/10.1007/s11104-013-1930-7
726 The GIMP Development Team. (2019). *GIMP*. https://www.gimp.org
727 Upchurch, D. R., & Ritchie, J. T. (1984). *Battery-Operated Color Video Camera for Root*
728 *Observations in Mini-Rhizotrons*. 1015–1017.
729 Vamerali, T., Ganis, A., & Mosca, G. (2009). *Methods For Thresholding Minirhizotron Root*
730 *Images*. *September*, 1–2.
731 Vincent, C., Rowland, D., Na, C., & Schaffer, B. (2016). A High-Throughput Method to
732 Quantify Root Hair Area in Digital Images Taken in Situ. *Plant and Soil*.
733 https://doi.org/10.1007/s11104-016-3016-9

734 Wang, T., Rostamza, M., Song, Z., Wang, L., McNickle, G., Iyer-Pascuzzi, A. S., Qiu, Z., &
735 Jin, J. (2019). SegRoot: A high throughput segmentation method for root image
736 analysis. *Computers and Electronics in Agriculture*, *162*, 845–854.
737 https://doi.org/10.1016/j.compag.2019.05.017
738 Yang, Y., Hillebrand, H., Lagisz, M., Cleasby, I., & Nakagawa, S. (n.d.). Low statistical power
739 and overestimated anthropogenic impacts, exacerbated by publication bias,
740 dominate field studies in global change biology. *Global Change Biology*, *n/a*(n/a).
741 https://doi.org/10.1111/gcb.15972
742

744