



**HAL**  
open science

# High-throughput and automatic structural and developmental root phenotyping on Arabidopsis seedlings

Romain Fernandez, Amandine Crabos, Morgan Maillard, Philippe Nacry, Christophe Pradal

## ► To cite this version:

Romain Fernandez, Amandine Crabos, Morgan Maillard, Philippe Nacry, Christophe Pradal. High-throughput and automatic structural and developmental root phenotyping on Arabidopsis seedlings. Plant Methods, 2022, 10.1186/s13007-022-00960-5 . hal-03881548

**HAL Id: hal-03881548**

**<https://hal.inria.fr/hal-03881548>**

Submitted on 1 Dec 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

METHODOLOGY

Open Access



# High-throughput and automatic structural and developmental root phenotyping on *Arabidopsis* seedlings

Romain Fernandez<sup>1,2</sup>, Amandine Crabos<sup>3</sup>, Morgan Maillard<sup>3</sup>, Philippe Nacry<sup>3\*</sup> and Christophe Pradal<sup>1,2,4\*</sup>

## Abstract

**Background:** High-throughput phenotyping is crucial for the genetic and molecular understanding of adaptive root system development. In recent years, imaging automata have been developed to acquire the root system architecture of many genotypes grown in Petri dishes to explore the Genetic x Environment (GxE) interaction. There is now an increasing interest in understanding the dynamics of the adaptive responses, such as the organ apparition or the growth rate. However, due to the increasing complexity of root architectures in development, the accurate description of the topology, geometry, and dynamics of a growing root system remains a challenge.

**Results:** We designed a high-throughput phenotyping method, combining an imaging device and an automatic analysis pipeline based on registration and topological tracking, capable of accurately describing the topology and geometry of observed root systems in 2D + t. The method was tested on a challenging *Arabidopsis* seedling dataset, including numerous root occlusions and crossovers. Static phenes are estimated with high accuracy ( $R^2 = 0.996$  and  $0.923$  for primary and second-order roots length, respectively). These performances are similar to state-of-the-art results obtained on root systems of equal or lower complexity. In addition, our pipeline estimates dynamic phenes accurately between two successive observations ( $R^2 = 0.938$  for lateral root growth).

**Conclusions:** We designed a novel method of root tracking that accurately and automatically measures both static and dynamic parameters of the root system architecture from a novel high-throughput root phenotyping platform. It has been used to characterise developing patterns of root systems grown under various environmental conditions. It provides a solid basis to explore the GxE interaction controlling the dynamics of root system architecture adaptive responses. In future work, our approach will be adapted to a wider range of imaging configurations and species.

**Keywords:** High-throughput phenotyping, Computer vision, Dynamic analysis, Root system architecture, Tracking

## Background

Plants expand their root system to meet their needs for water and nutrients. Root system shape, structure, and plasticity impact its efficiency and condition the plant fitness in fluctuating environments [1]. The root system

architecture (RSA) is dramatically affected by spatial and temporal changes in the soil environment that, in turn, will modify its water and nutrients capture efficiency [3]. These adaptive responses, including branching angles and frequencies of lateral roots, rely on growth rate modifications of the different root classes (e.g., primary, seminal, lateral, or adventitious roots). The optimal distribution of roots in a given environment allows plants to forage efficiently for resources. There is a need to explore the link between RSA and its dynamics, plant physiology, and agronomic traits to implement plant breeding programs

\*Correspondence: philippe.nacry@inrae.fr; christophe.pradal@cirad.fr

<sup>1</sup> CIRAD, UMR AGAP Institut, 34398 Montpellier, France

<sup>3</sup> Institute for Plant Sciences of Montpellier (IPSiM), Univ Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France

Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

[19]. Such goals involve designing non-destructive high-throughput phenotyping systems. Non-destructive measurements are indispensable to periodically quantify the phenotype of the same plant all along time and then decipher the dynamics of the adaptive responses. Because roots are the below-ground organs and represent the hidden half of the plant, non-destructive root phenotyping is technically challenging for high throughput. One approach to quantify root growth and geometry under a wide range of growth conditions is to use a transparent medium, such as agarose gels, a common cultivation practice for the model plant *Arabidopsis* and other species. Several semi-automatic tools have been developed to assist root phenotyping on agar plates [10, 23, 25, 27, 35]. However, these devices are either limited in the number of plates imaged or restricted in temporal resolution. This situation hampers potentially useful phenotypical parameters of individual roots that may be linked to the temporal dynamics of growth. Here we present the HIGH Resolution ROOT Scanner (HIRROS), an automated and non-destructive visualisation automaton for 200 Petri dishes and an image analysis pipeline allowing high-throughput temporal phenotyping of *seedlings* RSA. to support the massive phenome collection required for proper genetic studies in response to fluctuating environments.

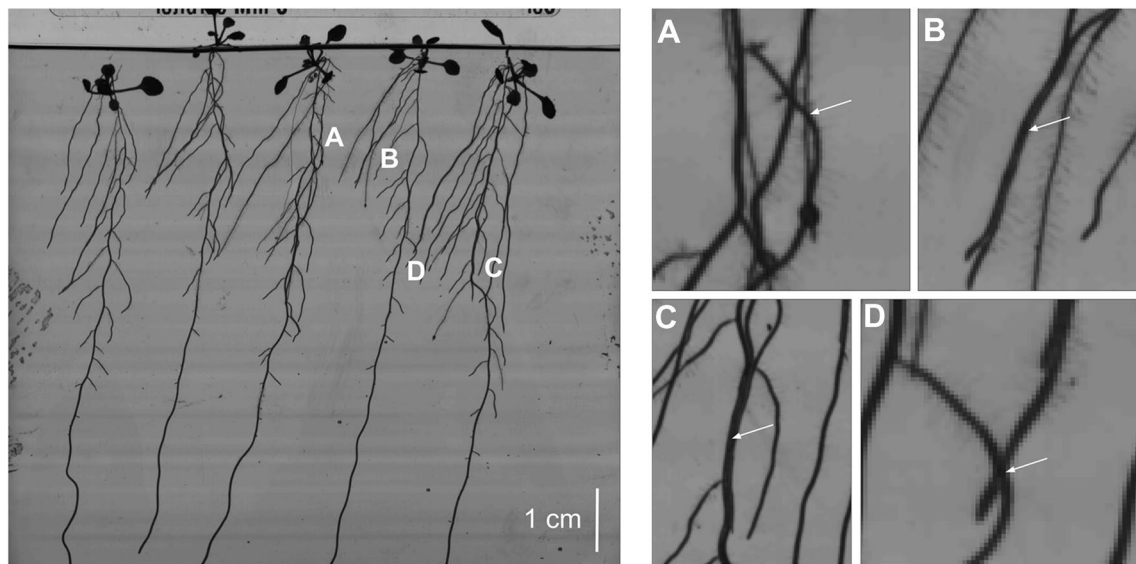
Automatic extraction of RSA is a current bottleneck and an active topic in bio-image analysis. The community attempts to develop solutions to process root systems 2D images [16, 34, 35], 3D CT observations [11, 12, 21], or multi-angle 3D reconstructions [33]; see [26] for a detailed description of standard root system phenotyping setups. The proposed processing pipelines for reconstructing RSA from 2D or 3D images typically include a step of image segmentation (root detection and background removal) and a step of RSA reconstruction by optimal path search in the segmented structures [23, 35]. The first step is a pixel classification task, previous works proposed various segmentation techniques, including pixel classifiers controlled by fixed or trainable parameters [5, 29]. Recent works exhibit significant advances using deep convolutional networks [23, 32, 35]. Unfortunately, these methods require vast training data and offer no guarantee when applied to a new dataset with different characteristics (species, observation device, bogus phenotype). Some authors partially addressed this concern using transfer learning [35], allowing the end-user to transfer the method to a new dataset after re-training the network.

The second step, aiming at RSA reconstruction, relies on algorithms extracting root paths from segmented structures, interpreting pixel sets as connected root systems to output higher-level primitives, i.e. plant organs and connections. Biologists need to assess that this

reconstruction process solves the architecture at the organ scale. This feature is required to distinguish local responses from systemic responses (elongation speed, branching density) [30, 31, 22], but offering such guarantees remains a challenge. There are several difficulties in current solutions, mainly due to the capacity to deal with root intersections and root contacts [10, 33, 35]. Each new crossing doubles the number of possible root paths, which causes a combinatorial explosion of the reconstruction problem (see Fig. 1). Some authors admitted that the root crossing issue is a significant failure cause without proposing any solution to tackle it [33]. A global optimization algorithm could provide an optimal solution to this problem. However, the most popular approach in the literature is reconstructing the different organs sequentially [35]. This method is more or less equivalent to addressing the root crossing problem with a greedy algorithm, leading to local optima.

With the development of high-throughput time-lapse observations, there is a shift to spatiotemporal data, adding another layer of complexity. The most common way to achieve spatiotemporal architecture reconstruction is to consider time points as separated data. These time points are processed using the conventional two-step pipeline before exploring the correspondences between the extracted architectures over time. This operation requires advanced matching methods [2, 10, 33] and extra processing steps that could induce a performance drop when cumulating the errors of the separate reconstructions. Some authors attempted to integrate temporal information in the reconstruction process to address this issue. Gaggion et al. [10] sidestep the geometric alignment task by fixing the specimen on the imaging device during the time course. With this setup guaranteeing a proper alignment of the successive observations, the authors enhance the temporal consistency of the obtained segmentation using a weighted average over the time series. The drawback of this method is the need to provide  $N$  observation devices for the  $N$  specimen (or for the  $N$  petri dishes, eventually), which is incompatible with high-throughput experiments. Finally, Yasrab et al. [36] proposed a novel perspective on the whole problem, training a Future-GAN network to reproduce the developmental behaviour of the root systems. This method showed high performance for geometry reconstruction but is prone to errors in topology reconstruction.

RSA complexity increases progressively in time series, making the time information a lever for solving the architecture problem: space and time can be combined to discriminate organ connections from root crossing. In this work, we propose a novel method for the automatic reconstruction of RSA integrating spatiotemporal information as a whole in an algorithm



**Fig. 1** Challenges of root system architecture reconstruction. Left, an image of Arabidopsis seedlings with a developed architecture, observed in a petri dish. **A** When observed in 2D, these root systems exhibit numerous root crossings, which induce possible root tracking errors, and can even fool experts when annotating static images without a “temporal insight.” In addition, multiple roots can follow each other for an arbitrarily long time (**B**), making them hard to disentangle. Finally, second-order roots sometimes hide below the primary root (**C**), and roots from two adjacent plants could cross (**D**). These points are severe issues for static reconstruction algorithms and greedy reconstruction processes, requiring addressing this problem with a 2D + t global algorithm

processing 2D + t data. First, a registration pipeline is built to address the alignment issue. Then, root systems are segmented and labelled with root apparition time. A novel root topological tracking algorithm is applied that combines temporal and spatial information and solves ambiguities introduced by root crossing. This pipeline is applied on 1000 time-lapse images of Arabidopsis roots grown in petri dishes ( $N=1000$  plants observed during 21 timesteps, timestep = 8 h), and the result is saved in the RSML format [17], enabling the extraction of new spatiotemporal phenes. Our automatic pipeline, evaluated over  $N=200$  plants imaged and expertized over 21 timesteps, shows state-of-the-art performances on high-throughput data. Finally, the development of architectures is reconstructed from time-lapse, providing unseen insights into root system growth.

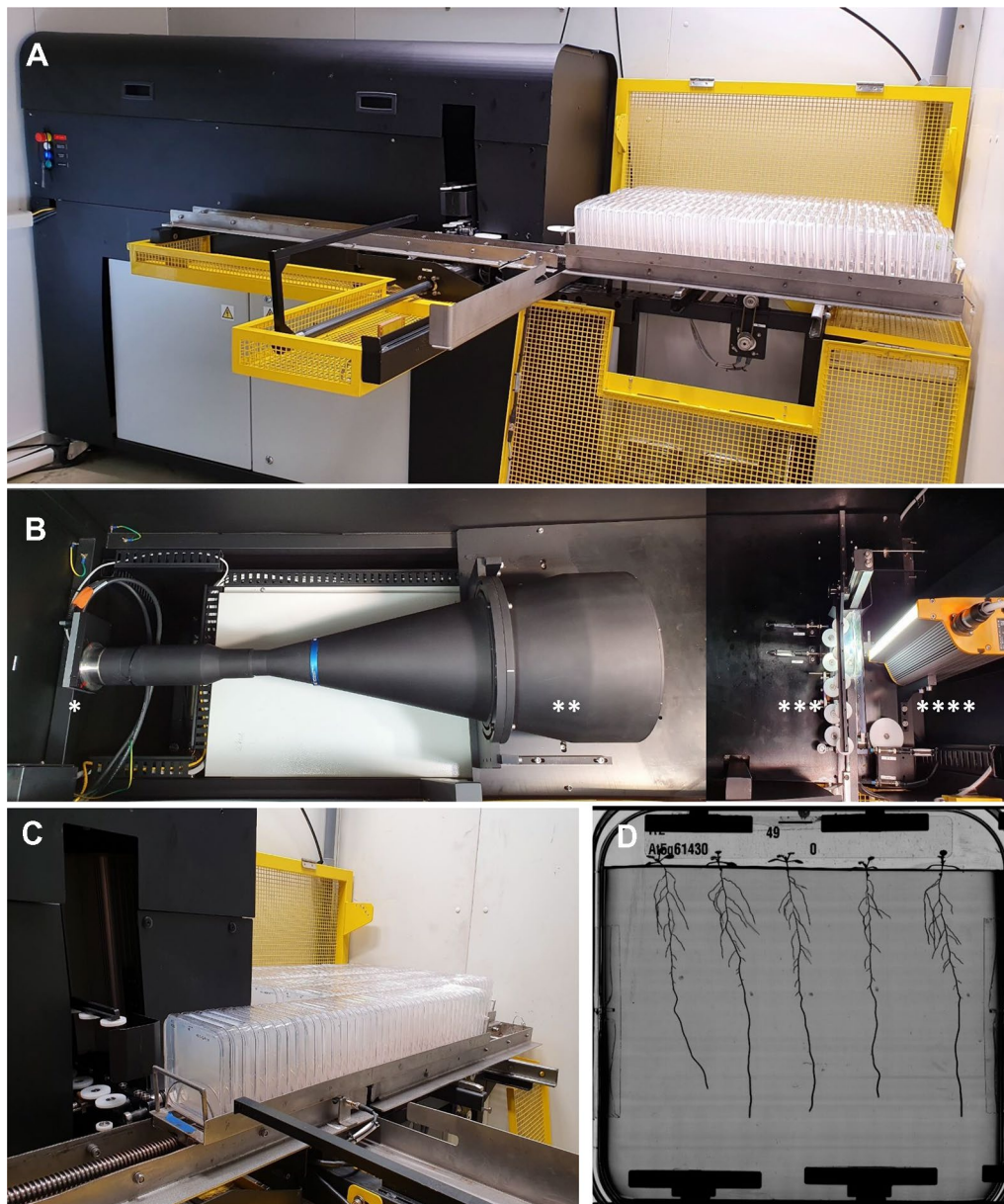
## Methods

### HIRROS: automated platform for RSA imaging

The HIGH Resolution ROOT Scanner (HIRROS) setup is a device for automated and non-destructive visualisation of root growth and architecture of *Arabidopsis thaliana* plants or small seedlings (Fig. 2). The imaging automata is located in a dedicated growth chamber allowing temperature, hygrometry, photoperiod, and light intensity to be adjusted according to the user’s requirements (Root phenotyping platform webpage: [24]).

HIRROS is designed to use either 200 small ( $120 \times 120$  mm) or 72 large ( $245 \times 245$  mm) square Petri dishes, which are filled with agar medium and placed upright in holders to allow root growth along a vertical plane at the surface of the culture medium. Once the medium is solidified, a one-centimetre wide strip of medium is removed to avoid contact of cotyledons and leaves with the culture medium and maintain shoots vertically (Fig. 2). Plate lids are wiped with tween20 to reduce condensation and water droplet formation. Five pre-germinated seedlings are evenly distributed, and the plates are sealed with four pieces of tape to allow gas exchanges. The Petri dishes are then placed into holders to maintain them vertically and make them transportable within the automated HIRROS setup. Holder motions are driven by worms allowing lateral and forward movements. The imaging system that is protected from parasitic light is located beside the culture zone (Fig. 2). Plates are transferred to the imaging cart. For illumination, the plates are backlit with a white collimated LED (Fig. 2). This allows a high contrast (even for thin and almost translucent Arabidopsis roots) and reduces reflection on the plate’s lid. Plates are imaged using a 16MP linear camera coupled to a telecentric lens located at 50 cm of the plate (Fig. 1). This optical setup increases the depth of field. It suppresses image distortions allowing imaging of small and large plates without adjustment. It has a





**Fig. 2** The High Resolution ROot Scanner (HIRROS) setup for automated and non-destructive visualisation of root architecture of seedlings grown in agar plates. Global view of the device (**A**) with the culture zone on the right side and the imaging zone on the left side. Overview of the imaging zone (**B**), which is composed of a 16Mpixel linear camera (\*), a telecentric objective of a diameter of 25 cm (\*\*), a mobile imaging cart (\*\*\*) holding plates during imaging, and a collimated LED backlight (\*\*\*\*). The square Petri dishes are maintained vertically in metal racks. They are stored in the culture zone (right side of the device). The rack moves to the imaging area (**C**) for image acquisition, and the plate is pushed in the imaging with a moving arm. Once the image is captured, the plate moves back to its initial location, and the rack moves to image the next plate. Once all plates are imaged, the rack moves back to the culture zone to proceed to the next one. Representative original grayscale image of Arabidopsis seedlings taken by the camera (**D**). 200 Petri dishes of 12 × 12 cm or 72 Petri dishes of 24 × 24 cm containing up to 1000 Arabidopsis plants fit into HIRROS

resolution of 19  $\mu\text{m}/\text{px}$  (image size 8 MB, TIFF for small plates and 32 MB for large plates). Metadata describing the experiment associated with each image and stored locally in a folder labelled with the experiment information and the imaging date. Illumination is switched on

and off automatically and synchronised with the camera during image acquisition. Once imaged, the plate moves back to the holder that moves one position further to place the next Petri dish in the row in front of the imaging cart until all plates are recorded. The holder moves

back to the culture zone, and the next holder enters the imaging process. The imaging frequency can be adjusted between 2 and 24 h, and up to 1000 plants can be imaged in less than 80 min.

### Image processing pipeline

Our processing pipeline includes four steps (see Fig. 3): (1) time series registration, (2) 2D+t segmentation, (3) root topological tracking, and (4) architecture reconstruction. It is fully automatic. The batch computation starts as soon as the user provides the indication of the source image directory.

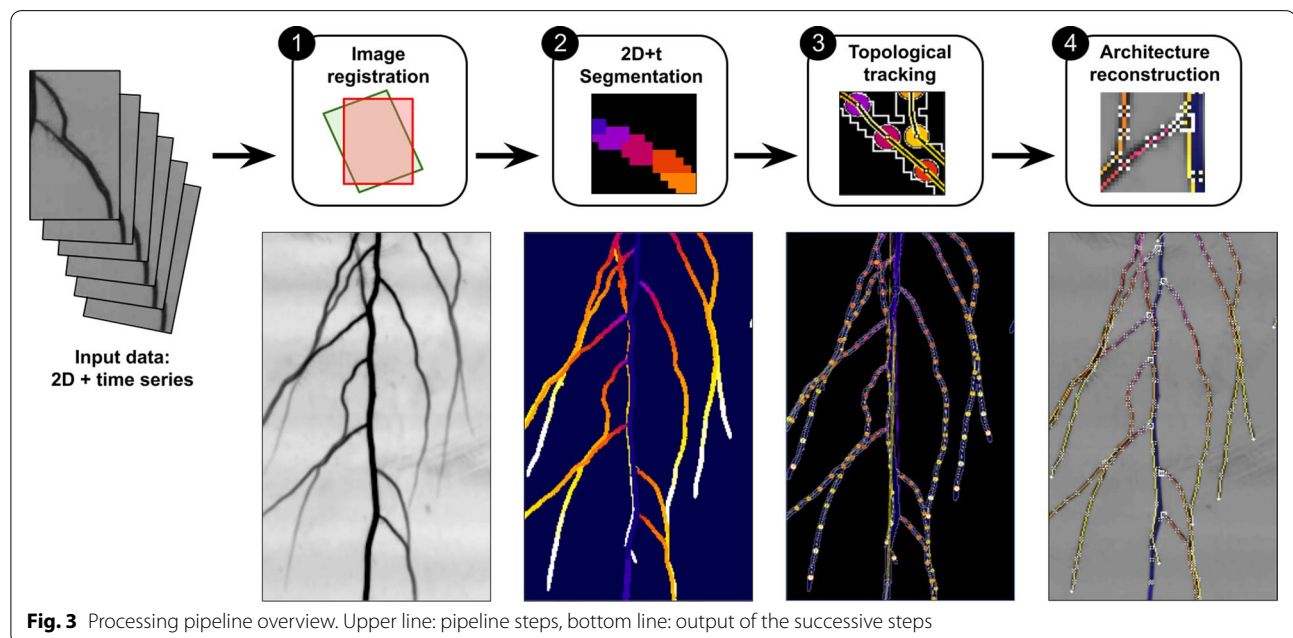
First, all the images in the time series are registered together in order to align root organs in all images to ease the tracking process. Then, in the second step, root pixels are separated from the background (segmentation), and an apparition time is associated to the root pixel that corresponds to the time associated with the image in the time series. A single 2D labelled image is output where each pixel is labelled with a root apparition time.

The two final steps work on graph and tree data structures. In step 3, the label image is converted into a graph representing the successive growing parts of roots. The adjacency relations of the root segments are investigated by an optimization algorithm to estimate the topology of the root system architecture. Finally, in step 4 we refine the geometrical description of this architecture by computing the central line of the identified organs. The detailed description of these steps is presented below.

### Time series registration

The pipeline's first operation corrects the root system's misalignment in the 2D+t image sequence  $I = \{I_i\}_{i=1\dots N_t}$ , induced by the displacement of the plate in the imaging automata. We identify two sources of misalignment: (i) movements of the box relative to the imaging device and (ii) deformations of the root system in the culture medium induced by plate displacement. We correct these distortions using a registration pipeline developed with the FijiYama plugin libraries [8].

We fix the geometry of the last image  $I_{N_t}$ , considered the reference image because it contains all the structures present in the previous images. We determine  $(N_t - 1)$  geometric transformations  $\{T_t\}$  to align the other images with the reference image  $I_{N_t}$ . Each estimated transformation  $T_t$  comprises a rigid transformation (translation + rotation), and a non-linear transformation (dense vector field) estimated successively. In this setup, we assume that the rigid transformation is required to correct the movement of the plate, while the non-linear transformation is required to correct the slight deformations of the root system in the culture medium. These transformations are computed from local correspondences using the blockmatching algorithm [28], integrated into a daisy-chain strategy: each image is registered with the next image in the sequence. Then the successive transformations are composed and applied at once for each image.



**Time series segmentation**

The image time series shows the roots growing over time. The difference between two successive images highlights the newly appeared segments of the roots. By working on these differences over the image sequence, one can distinguish the roots from the image’s background, thus segmenting roots and associating each root segment and apparition time. Thus, we define the two purposes of this segmentation step: (i) identify which image pixels are traversed by the root systems during the sequence (binary segmentation of RSA), and (ii) estimate the corresponding traversal times (see Fig. 4).

In our dataset, background pixels have medium intensity while root pixels have low intensity. However, segmenting roots by thresholding each image separately is not an option, as other structures (box border, condensation, dirtiness) may have similar characteristics to roots, and pixels intensity varies between the root central line (low intensities) and root border (medium intensities).

Nevertheless, the registered image sequence allows studying the time sequence of intensities of each pixel and detecting in this sequence the event associated with the detection of a root (i.e., when a root reaches this pixel). The time sequence of intensity of a pixel traversed by a root begins with high values corresponding to the background, then undergoes a sudden decrease of intensity at the moment when the root reaches this pixel, and finally finishes with low values corresponding to root tissues. We process each pixel individually and identify this pattern by computing the maximum mean shift over its time series of intensity. We identify as “root pixels” those whose mean shift value is greater than a fixed threshold  $s_1 = 25$  (given that the average intensity difference measured between the centerlines of the roots and the background is approximately 100). Among these pixels, we compute the difference between successive values in the

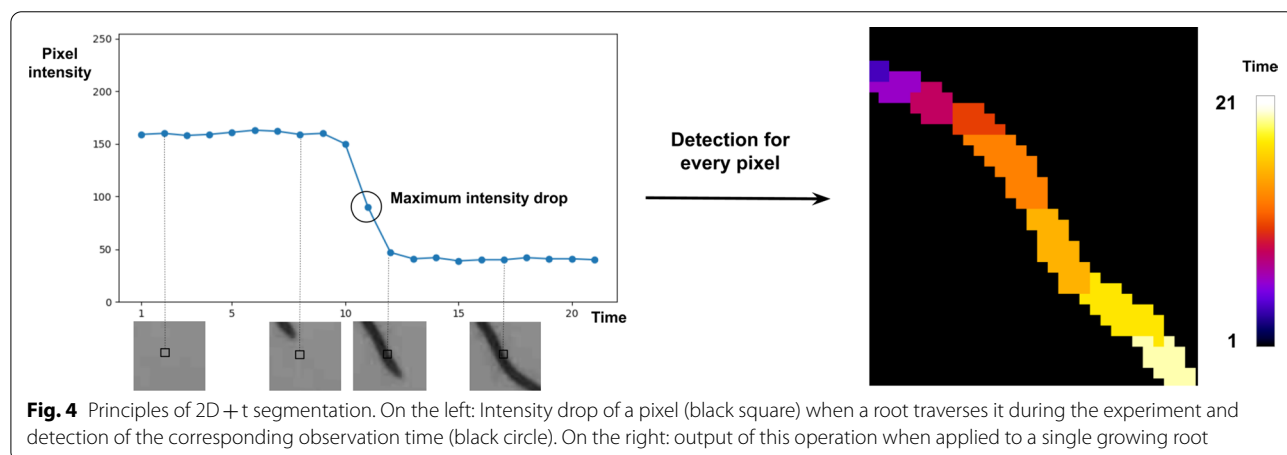
time sequence and identify the observation time following the maximal intensity drop. If this maximum intensity drop is lower than a fixed threshold  $s_2 = 10$ , the pixel is rejected and identified as background. Else, the index of the maximal intensity drop is identified as the “apparition time” of the root at this given pixel location. The pixels selected by these two thresholding operations constitute a binary segmentation of the root systems. Each selected pixel is given a label defining the root apparition time (see Fig. 4).

We run a topological and geometrical characterisation of the binary segmentation of the root systems to eliminate various outliers: nightly condensation and dirtiness falling on the dish during the experiment. We compute the 4-connected components and reject the pixels of components smaller than 2000 pixels. In our setup, this threshold is barely equivalent to rejecting any candidate root system in which cumulative organ length at the last observation time point is lower than 38 mm, which is suited to our experiment.

This segmentation step results in a 2D labelled image  $S$  (see Fig. 4), whose pixels  $p$  are given a label  $S(p)$  between 0 and  $N_t$ ; the label “0” identifies non-root pixels, and the other possible labels identify the index of the first image where the root is present.

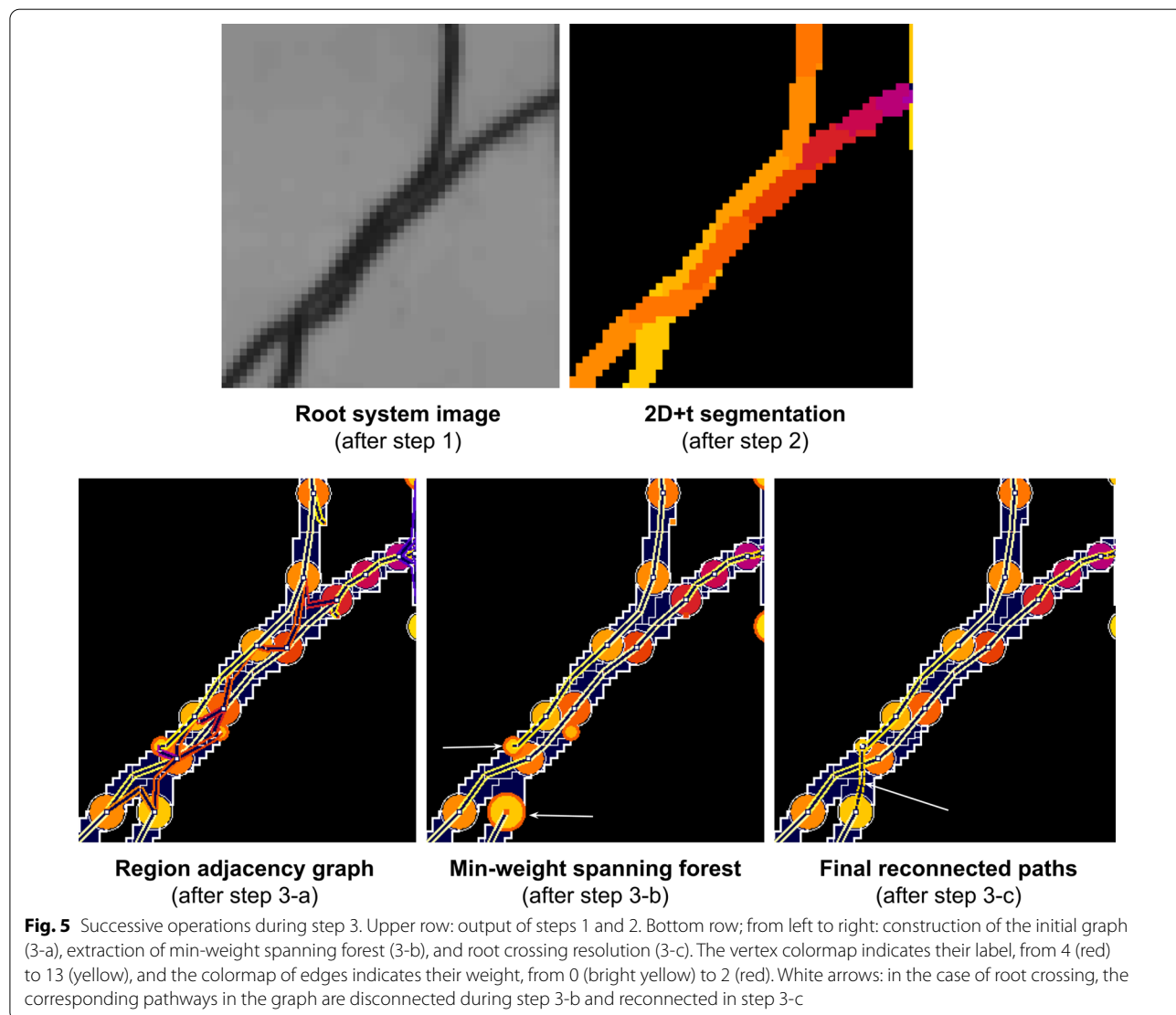
**Root topological tracking**

In this section, the pipeline converts the 2D labelled image into a graph representing the root system architecture. This graph is extracted by investigating the adjacency relations between the detected root growing sections, and connecting these sections in the most plausible way, using a connection metric used to compute a minimal cost tree with an optimisation algorithm. Multi-particle tracking can be challenging, especially if the number of particles is not constant, which requires



managing the creation and destruction of tracks in time [18]. An additional difficulty comes when the temporal sampling is sparse with respect to the particle velocity, inducing possible confusion between crossing tracks or neighbouring tracks. For the tracking of roots in the time series, we propose a solution based on a global tracking algorithm working on the full-time series. In the previous step, the pipeline outputs the growing architecture as a fixed architecture with time-stamped elements. During root topological tracking, the method tracks structures by using these timestamps. In steps 3-a, the algorithm computes a directed Region Adjacency Graph equivalent to the segmentation. In 3-b, it extracts a covering forest of minimal weight from it, and provides the skeleton of the primary and secondary roots. Then in 3-c, it processes the root crossings by a global optimization algorithm (see Fig. 5).

*Representation of roots growth space using a weighted directed region adjacency graph* After registration, we expect that the root system appears in the sequence of images as a sequence of Russian dolls: assuming that a root does not disappear (by dying or being removed) if we note  $P_t$  the set of pixels covered by the roots at the time  $t$ ,  $P_t \subseteq P_{t+k}$  ( $k > 0$ ), and the difference  $P_{t+k} \setminus P_t$  is the additional surface of root system generated by root tips growth between  $t$  and  $t + k$ . Thus, we assume that a connected component of pixels in  $S$  sharing the same label  $t$  is an additional surface covered between the observation time number  $t$  and  $t + 1$  by a part of the root systems. With this assumption, root trajectories can be tracked by identifying the optimal pixel paths in  $S$ , beginning with pixels labelled 1 (sequence start) and finishing with the label  $N_t$  (last observation time point). We formalise a global solution to the tracking problem by searching a minimum cost





spanning forest in the weighted directed region adjacency graph  $G = (V_G, E_G)$  of the labelled image  $S$  (see Fig. 5). We build  $G$  as follows:

- Each 8-connected component of pixels  $p \in S$  sharing the same label  $t(p)$  in  $S$  is represented by a vertex  $v \in V_G$  with label  $t(v) = t(p)$ .
- Two vertices  $v_1, v_2$  are connected with a (directed) edge  $e = (v_1, v_2)$  if and only if  $t(v_1) < t(v_2)$  (edges are directed according to forward time direction) and the two corresponding 8-connected components in  $S$  are 4-neighbours (they share at least a pixel edge).

The weight  $p(e)$  of an edge  $e = (v_1, v_2)$  is defined by:

$$p(e) = \frac{|\#v_1 - \#v_2|}{\#v_1 + \#v_2} + (t(v_2) - t(v_1) - 1) + \frac{v_1 v_2 \cdot \underline{y}}{|v_1 v_2|}$$

with

- $\#v$  the number of pixels contained in the corresponding connected component in  $S$ .
- $t(v)$  the label of  $v$ .
- $v_1 v_2$  is the vector from  $v_1$  to  $v_2$  and  $\underline{y}$  is the Y-axis normalised vector.

The weight  $p(e)$  of an edge is small (the edge is more “likely” to belong to an actual root trajectory) if (i) the edge connects two vertices which have a comparable surface, (ii) the two vertices labels (structures apparition time) are close, and (iii) the vector connecting the target from the source is directed downwards, which is the preferred orientation of root growth.

**Extraction of root system skeleton** At this step, we extract the root systems skeleton in the graph by identifying the paths corresponding to primary roots and lateral roots without considering lateral root crossings. We describe this reconstruction algorithm for architecture composed of primary roots and lateral roots (second-order organs inserted on the primary), which is suited to arabidopsis seedlings.

**Primary roots:** The primary root of seedlings is the first organ of the root system and tends to be larger than the lateral roots, have a higher growth rate and point downwards. Consequently, the corresponding path in the graph is composed of edges of minimum weight  $p(e)$ . At the beginning of the experiment, we assume that the primary axes of the  $N$  plants ( $N_p=5$  plants per box in our setup) have no branching. We then identify the corresponding vertices in  $G$ , which are the  $N_p$  vertices associated with the largest connected component of timestamp 1 in the segmentation  $S$ . From each of such vertices, we

extract the path of the primary axis, which is the longest greedy min-cost path. Vertices and edges traversed in this step are labelled primary, and all other edges and vertices in  $G$  are labelled lateral. In agreement with the root hierarchy, all lateral edges incident to a primary vertex are invalid and eliminated from the graph. Finally, the algorithm filters out outliers by computing an extraction of connected components in  $G$  from the primary vertices and rejecting all vertices not included in the final components.

**Lateral roots:** The algorithm identifies the path corresponding to lateral roots by selecting in the graph the vertices and edges that are likely to correspond to actual lateral root paths (see Fig. 5). This action of classifying between “relevant” and “irrelevant” edges is an action of graph pruning that we run in two steps. First, we aim to extract a subpart of the graph whose topology corresponds to that of a set of root systems, which is a forest structure. To that end, we extract a forest from the graph by computing the minimal directed spanning forest  $F = (V_F, E_F)$  using the Edmonds algorithm [7]. In  $F$ , each vertex has a single incoming arc. Second, we try to ensure that each path corresponding to a lateral root has no branching, according to our dataset composed of Arabidopsis seedlings. In order to avoid the branching of lateral roots, we check each lateral vertex and keep only their min-cost outgoing arc.

**Resolution of root crossing ambiguities** By pruning the graph, parts of actual root paths corresponding to root crossings could be disconnected. This point is investigated and processed in this step using a global algorithm. If a root  $r_2$  crossed the path of another root  $r_1$  during the time sequence, assuming that  $r_2$  traversed the area at least one timestep after  $r_1$ , the region adjacency graph in 3-a) shows a typical “X” pattern, where the path of  $r_2$  cross at least one edge in the wrong direction. This pattern leads to an upstream disconnection of the  $r_2$  path and a downstream disconnection of the  $r_2$  path during graph pruning in 3-b), which lets an additional root start and root end in the graph (see Fig. 5). To process these cases using a global algorithm, we first identify the two following vertices sets in : the “root starts”  $V_{start}$ , including every lateral vertex with no incoming arc, and the “root stops”  $V_{stop}$ , including every vertex with no outgoing arc. As these elements are possibly parts of disconnected pathways, we target to identify the best pairwise matches (root stops with root starts) in order to reconnect them. We define a connection cost function to compute the optimal set of pairwise matches between elements of these two sets. The cost function is defined to indicate the likelihood of the connection, computed as the weighted average of multiple geometrical and topological continuity features, replaced

by majoring values when the elements to connect are not long enough to provide the needed information. The features used for the cost function are the edge weight function  $p(a)$  (see 3-a), the angle between root start and root end, the relative speed difference between the root start and the root stop, and a topological score. The topological score is computed by evaluating the minimal path between the root end and root start in the original region adjacency graph, evaluating the number of edges crossed in the right or wrong direction, and the corresponding temporal distance. We evaluate this cost function for all the possible ordered pairs  $(v_{stop}, v_{start})$  in order to build a cost matrix of all possible one-to-one connections of root starts and root stops. This cost matrix is used to compute the optimal pairwise matching using the Hungarian algorithm [14]. The algorithm works as follows:

---

**Function: Iterative reconnection of isolated root start and dead-ends**

```

→ Compute the cost matrix C, which values are the connection cost of all possible
pairs (root stop, root start)

→ While there is at least one root stop and at least one root start left, do :
    → Compute the optimal set S of connections using the Hungarian algorithm
    → In S, select the connection of minimum cost and build the
        corresponding edge  $e = (v_{stop}, v_{start})$ .
    → If  $(t(v_{stop}) = t(v_{start}))$  and (insertion of e create a cycle in F):
        → reject e and stop the algorithm.
    → Else Add the new edge e to the graph
    → In C, update the values that could have been modified by adding a.

end

```

While the Hungarian algorithm provides a global and optimal set of connections, we do not apply this set in one go. We iterate over these steps: computation of the Hungarian algorithm, connection of the pair of minimal cost, and update of the cost matrix. That way, each new connection adds knowledge to the system. This point is critical for tackling some challenging cases, like when a single root is reconstructed from a series of numerous small isolated elements comprising both a root start and a root end. This kind of situation happens for example,

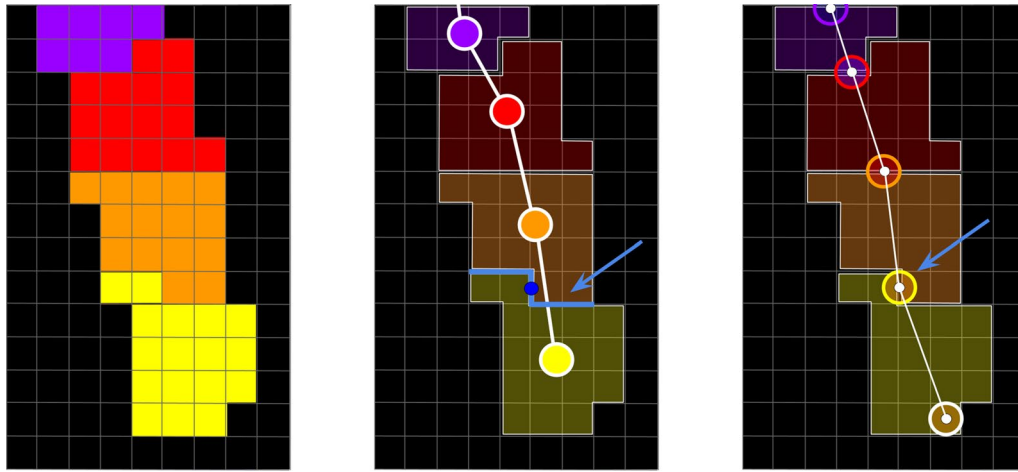
when a lateral root grows hidden below a primary root (see results, Fig. 7D).

**Artefacts rejection** We compute statistics over the identified lateral root paths and proceed to detection of outliers in order to reject spurious lateral roots. To that end, we collect data characterising the geometry and dynamics of these paths, measured at each timestep: length, speed, and surface. These scalar values are binned with respect to the temporal distance from the emergence of the path. Then we compute robust statistics within these subgroups. We compute the double-sided MADe [15] and detect values exceeding twenty-five times the median absolute deviation around the median. These abnormal values are identified as markers of spurious roots. The corresponding roots are then rejected as artefacts.

---

**Architecture reconstruction**

The reconstruction of the spatiotemporal model of the RSA involves a precise description of the topology and geometry of the root systems. At step 4, the pipeline refines the geometric description of the root system architecture by identifying the successive positions of root tips and interpolating the path of the root tips to extract the central lines of the organs. Vertices and edges of the forest  $F$  resulting from step 3 describe the root system topology. However, they do not provide the precise



**Fig. 6** Estimation of root central line  $2D + t$  coordinates. Left: labelled segmentation, comprising successive growth intervals of a single root, from purple (earliest interval) to yellow (latest interval). Middle: corresponding elements (edges and vertices) in the min-weighted spanning forest  $F$  extracted from the region adjacency graph. Right: Corresponding elements in  $I$ , the dual graph of  $F$ , whose spatial coordinates of vertices are estimated from the geometry of the separation between successive growth intervals (blue arrow and blue line)

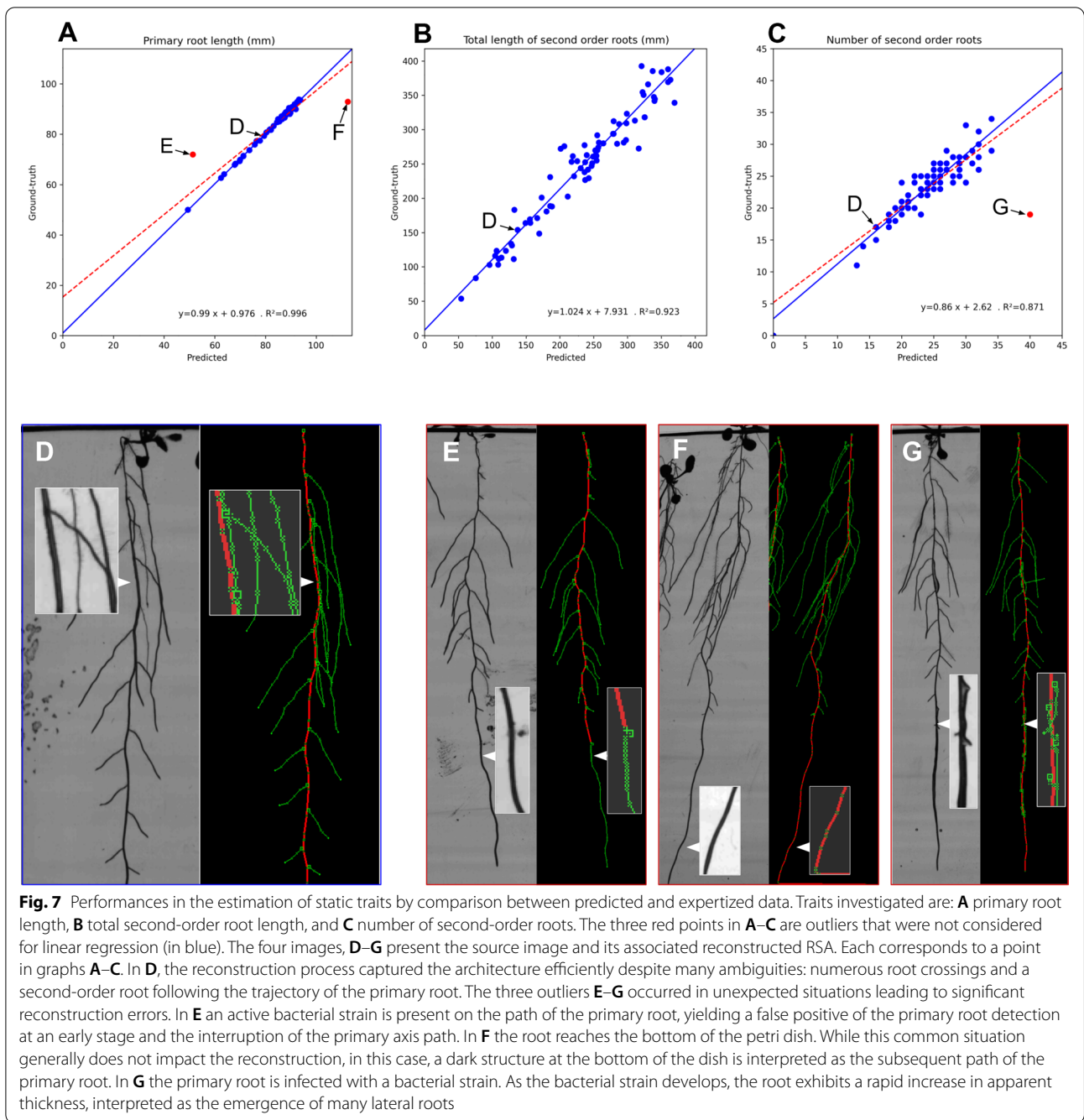
spatiotemporal pathway: each vertex of  $F$  represents a growing region of root pixels between two observations; thus, it is not associated with any exact coordinates, neither in time nor space. To build a precise spatiotemporal model, we need to know the precise position of each root tip at each observation time point and then interpolate this position between the time points to follow the roots' actual pathway, whatever their speed and tortuosity.

The position of the root tip at each time point can be estimated from the edges of the forest  $F$ , recalling that each edge  $e \in F$  corresponds to a transition between two consecutive growth areas of a root (see Fig. 6). Thus, edges of  $F$  correspond to the successive positions of root tips at each time point. The spatiotemporal model is then represented by the dual graph  $I = (V_I, E_I)$  of  $F = (V_F, E_F)$  where a vertex encodes the spatiotemporal position of a root tip and an edge encodes the root segment between two positions (see Fig. 6). In this dual graph (Fig. 6, right), a vertex  $w \in I$  corresponds to an edge in  $F$ . Its temporal coordinate is well-defined; it is exactly an observation time point.

Conversely, the spatial coordinates of  $w \in I$  have to be defined.  $w$  corresponds to an adjacency relationship between two distinct connected components in  $S$  (see the blue line and the blue arrow in Fig. 6, centre). We define the geometry of this adjacency in the image space by extracting the set of pixel edges and corners separating the two adjacent components. We choose the spatial coordinates of  $w \in I$  to be the coordinates of the central-most pixel corner or pixel edge over all the corresponding frontiers (see the dark blue circle in Fig. 6, centre). We process all the edges of  $F$  that way to compute the

corresponding vertices of  $I$ . Afterward, we create for each root an additional vertex  $w \in I$  to identify the root tip. We estimate its position by identifying the pixel of the last component that is the most distant (geodesically) from the nearest identified vertex  $w \in I$  (for an illustration, see the lower white circle in Fig. 6, right). Once all the vertices of  $I$  are computed this way, we connect them with edges corresponding to the vertices of  $F$ . The geometry of the edges is defined by straight lines connecting successive positions of root tips. These edges constitute an oversimplified version of the roots pathway, which do not precisely follow the roots pathway, especially for primary roots, which are growing very fast. We refine  $I$  to ensure geometrical fidelity of the RSA model, allowing accurate measurements (especially root length). For each edge  $\varepsilon = (w_1, w_2) \in I$ , we compute the shortest path between  $w_1$  and  $w_2$  in the corresponding connected component  $CC \in S$ . We use the Dijkstra algorithm and compute the shortest path of 8-connected pixels in  $CC$ . We define the edge cost using the standard inter-pixel euclidean distance, weighted by a scalar function derived from the distance map of the traversed pixels to the contour of the  $CC$ . This morphological feature forces the path to avoid understeering curves; it fosters representative RSA models.

Finally, we reduce the number of points in the pixel path by decimating it using the Douglas-Peucker algorithm [6]. The temporal coordinate of the vertices is estimated by a linear interpolation computed from the nearest two bounding vertices whose temporal coordinates are known.



*Upstream extrapolation of primary roots to the seed* Leaves sometimes fall into the medium during the time course, inducing false detection of new organs. We address this problem by artificially removing the falling leaves from the images using a series of morphological operations. However, this results in a loss of information about the primary and lateral root starts, causing the pipeline to detect the root system initiation

point below the target point (the seed). The pipeline corrects this defect during this post-processing operation by computing a pixel path of minimum cost from the target height to the detected initiation point using the Dijkstra algorithm. To that end, we define the cost of pixel-to-pixel traversal as the product of the inter-pixel distance (1 for pixel sharing an edge or  $\sqrt{2}$  for pixel sharing a corner) by the average intensity of the two pixels.



**Downstream extrapolation of the stopped laterals** After computation, we check every root end, and investigate the one ending before the last observation time, probably corresponding to lateral roots stopping their growth before the sequence ending. Such an event is infrequent, thus suspect. If these early-stopping roots  $r_{stopped}$  are not incident to any other root, we consider that  $r_{stopped}$  have stopped its growth. In that case, we keep it in the model if it has a continuous growth during at least three time points, or we reject it instead. In the other case where  $r_{stopped}$  is incident to another earlier root  $r_{first}$ , we assume that  $r_{stopped}$  continued its growth hidden below or parallel to  $r_{first}$ . In this situation, we extrapolate the growth of  $r_{stopped}$  to its estimated final position in the time-sequence and add these new points to the RSA model.

### Pipeline validation

We built an annotated dataset for validation purposes. In this section, we detail the procedure and the dataset involved in these experiments.

### Validation methodology

Assessing the pipeline's performance involves confronting the pipeline's outputs with corresponding data annotated by an expert to assess the quality of the analysis and identify drawbacks. We reused and modified the Smartroot software library [16] to develop an ImageJ plugin to edit the pipeline's output. Our plugin provides a simple GUI to navigate an RSA model in 2D+t and apply elementary editing operations: adding, modifying, and removing a root or a vertex. Each edition operation takes place in 2D+t. After each operation, the topology is checked automatically to correct the graph if needed. The expert stops editing the RSA model when he estimates that both the topology and geometry are valid at each time point of the sequence of observations. In our dataset, the mean number of operations to correct the RSA of a single plant was  $15.4 \pm 6.3$  operations per plant (mean  $\pm$  standard deviation).

### Validation dataset

The entire dataset was acquired within a seven days period. It contained  $N_{tot} = 1000$  plants, divided into 200 petri dishes loaded into HIRROS. Experiments took place over 160 h, with one observation every 8 h (each plant/dish is observed 21 times): the total image dataset contains 21.000 observations of a plant at a timestep. We selected a representative subset of 36 Petri dishes ( $N_{test} = 180$  plants) to be manually annotated along the 21 time points (3780 plant.timestep) to perform validation tests. The expert identifies which plant can be considered an outlier during the annotation process. Most of the time, outliers are bacteria-contaminated plants that

stopped growing or died during the time course. These plants are removed from the dataset and are not included in the validation dataset.

The 36 dishes cover the genotypic and environmental variations. The first 18 Petri dishes are used as a training set to train the model and fine-tune the pipeline parameters. The 18 remaining dishes are used as a validation set. The metrics presented in the results section are computed on this validation set, by comparing the phenes measured on the estimated architecture with those on the annotated architecture. Consequently, the results presented hereafter are computed on a test set of  $18 \times 5 \times 21 = 1890$  plant.timestep.

### Phene measurements

The annotation process consists on editing 2D+t RSML until they fit perfectly with the image time-series. Afterwards, these RSML can be processed by automatic routines to extract static and dynamic phenes. Then, predicted versus annotated RSML are compared to measure the performance of the pipeline for static and dynamic phenes predictions. More specifically, the length of a root at a given time  $t$  is obtained by summing the length of the individual segments constituting its root central line, and having an apparition time lower or equal to  $t$ .

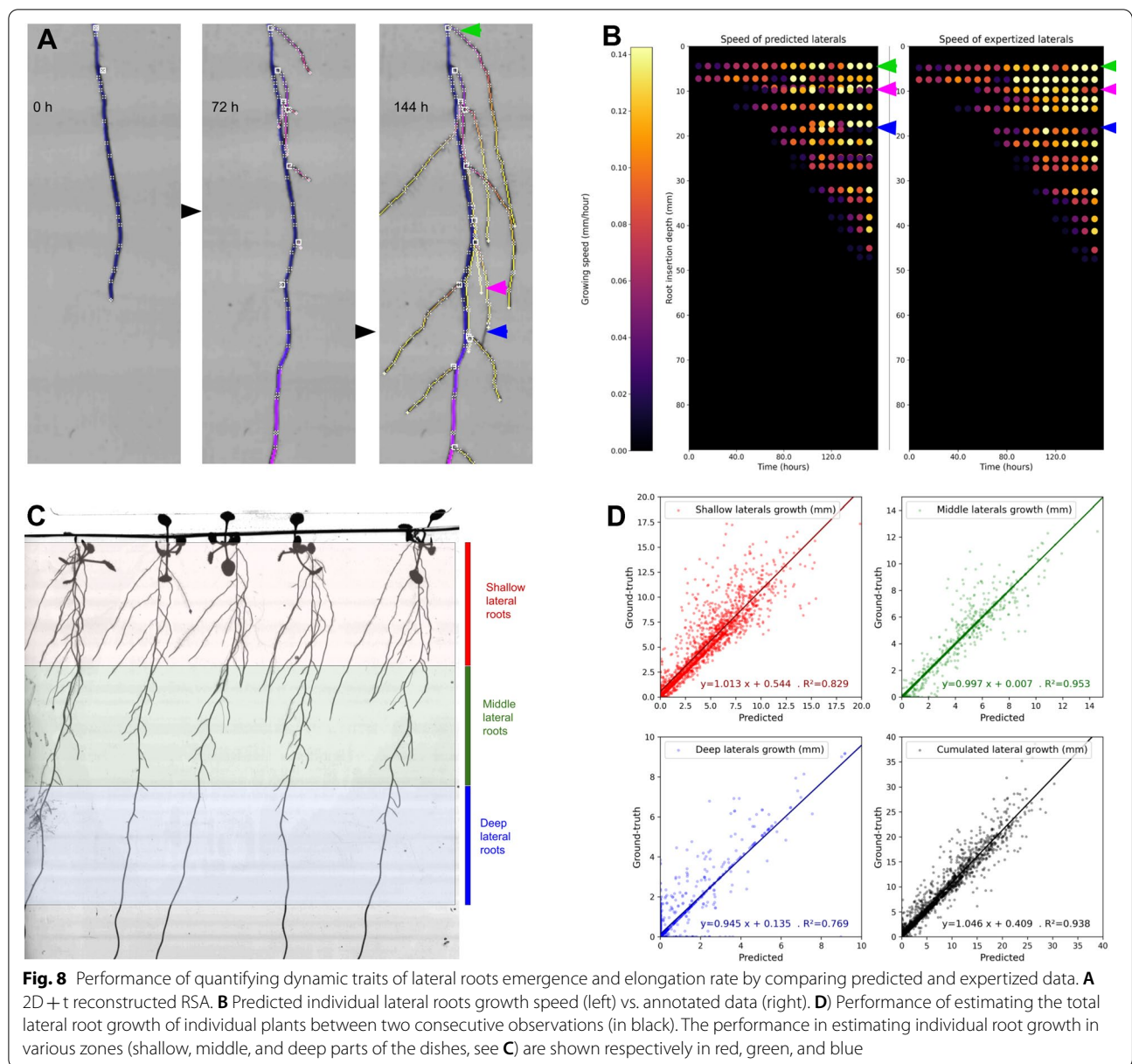
### Results

The performance of the pipeline is measured using several metrics to estimate the geometrical and topological fidelity of the reconstruction.

### Performance of static trait estimation

We first assess the performance of the pipeline on the estimation of static traits by using global traits frequently used in the literature: total predicted number of organs and predicted cumulated length for primary and second-order roots. Our dataset is composed of complex root architectures with numerous organ crossings. We compute these traits at the last time point when the architecture reaches its maximum complexity. At this stage, we evaluate the capabilities of our method to disentangle the numerous organs when processing complex root systems.

The three graphs in Fig. 7 illustrate the accuracy of the measurements, especially on organ length estimation (Fig. 7A and B), with a correlation coefficient of 0.996 and 0.923 for primary and lateral roots, respectively. Moreover, regression lines between predicted and ground-truth values have small intercepts and slope coefficients near 1 (i.e., 0.99 for primaries and 1.02 for second-order roots, see Fig. 7A and B). Our method, applied on a complex dataset, also exhibits state-of-the-art performances [35, 36] on second-order root counting (Fig. 7C), with a correlation coefficient of 0.818. These performances make



our method suited for collecting the most common phenes required in high-throughput experiments. Figure 7E–G highlight three outliers (~3% of the dataset) due to developing bacterial strains. These non-plant elements were identified as roots and severely impacted the prediction of the pipeline.

**Performance of dynamic traits estimation**

The pipeline reconstructs RSA over time, allowing to quantify each root’s elongation rate and its evolution throughout the time series, providing access to dynamic traits. In this section, we evaluate the pipeline performance of dynamic traits quantification. We measured the

total additional length between successive time points and the additional length corresponding to laterals. We further investigate the source of errors by estimating the additional length of laterals within three classes: roots attached to the primary’s upper part, middle part, and the bottom part (Fig. 8).

Figure 8A represents the successive observation of a root system overlaid with the corresponding reconstructed architecture. Figure 8B represents the 2D + t reconstruction of a unique plant from data automatically obtained by the pipeline (left panel) or from expert analysis (right panel). Each horizontal dot-line represents a lateral root, and the colour of each dot represents the

elongation rate of the individual root between two successive time points. This figure visually illustrates the high accuracy of the analysis. Most organs are well estimated (green arrow), with minor differences occurring when the track of a root is lost and prolonged artificially (pink arrow) or when a root is cut into two single organs (blue arrow). The cumulated lateral root growth is estimated between two successive observations with high accuracy ( $R^2 = 0.938$ ). Regression lines between predicted and ground truth values have small intercepts and a slope coefficient near 1 (i.e., 1.05). The most challenging part to estimate is the deeper compartment of lateral roots, inserted in the deepest third of the petri dish. Roots growth from this compartment is estimated with a lower precision ( $R^2 = 0.769$ , regression line slope = 0.95). These roots appear at the end of the time course, are smaller, and the length of the growing units is closer to the pixel size. Thus, their growing speed cannot be estimated accurately. The shallowest third of the petri dish (Fig. 8C, in red) suffers from other misestimation biases induced by reconstruction errors when the leaves fall into the root system ( $R^2 = 0.829$ , slope = 1.01). Finally, roots in the middle zone do not suffer from these issues and exhibit a high accuracy ( $R^2 = 0.953$ , slope = 1.00). To our knowledge, there are no phenotyping studies that validate such dynamic traits, while these measurements are of high interest to plant scientists.

**Table 1** Comparison of primary and lateral root length of two populations with different GxE conditions

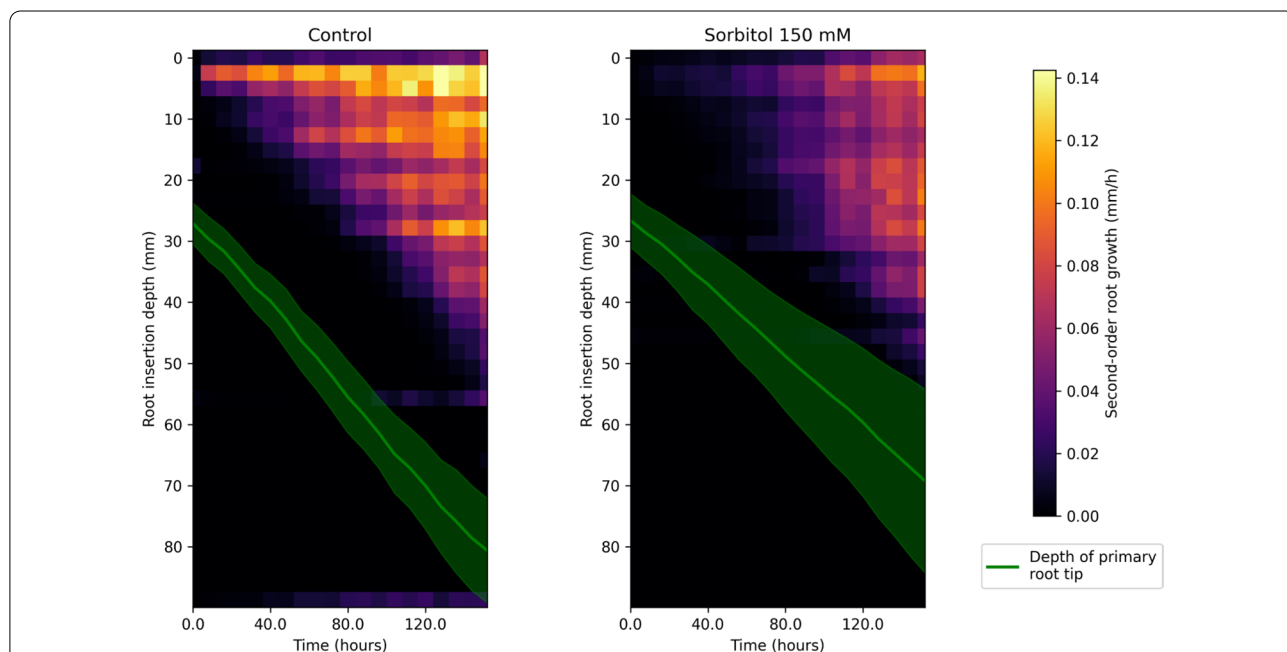
	Primary root length	Lateral root length
Control (mean ± std over N = 30)	80.6 ± 8.5	184.7 ± 36.8
Sorbitol, 150 mM (mean ± std over N = 30)	69.2 ± 15.0	89.7 ± 41.3

**Characterization of spatiotemporal phenotypes of root systems**

In this section, we investigate the resulting spatiotemporal traits extracted to unravel the studied root systems’ characteristics and assess our pipeline’s interest for GxE studies.

**Topological and geometrical characterization of root system dynamics**

Root system topology is widely complex, requiring the information to be simplified before being analysed. We derived an approach used for meristem organ characterization to produce a heatmap of the growing areas of the root systems. In order to evaluate this approach, we compared the reconstructed graph of plants grown for six days on a control medium or subjected to a 150 mM sorbitol osmotic stress (Fig. 9). Terminal RSA analysis showed that osmotic stress reduces primary and lateral



**Fig. 9** Elongation rates of second-order roots as a function of their insertion depth on the primary root. Each graph results from statistics computed over N = 30 specimens of the same genotype grown under control conditions or subjected to osmotic stress (150 mM sorbitol). Primary root tip depth is displayed in green (mean ± std-dev)

root length by 16 and 50%, respectively (Table 1). Kinetic reconstruction clearly showed that the primary root elongation rate was reduced under osmotic stress since the application of treatment (green line) but appeared more variable between plants than under control conditions (dark green cone). The elongation rate of lateral roots under osmotic stress was significantly reduced during the first part of the experiment (p-value < 0.001 under Mann–Whitney test, see Additional file 1: Fig. S1), while this statistical distinction progressively vanishes during the last days of the experiment. In addition, osmotic stress dramatically delayed the development of older lateral roots in the 0–15 mm depth interval and significantly reduced the elongation rate of these roots (colder color on the heat map). On the contrary, lateral roots on the 30–50 mm depth interval appeared less affected by osmotic stress with no significant delay in development and little difference in the elongation rate, suggesting different sensitivity or acclimation to osmotic stress. Finally, osmotic stress also altered the topology of the root system by reducing the distance from the last visible lateral root to the primary root tip (distance between the green line and the last visible line of the heat map). Altogether, these results open promising and unexpected perspectives to analyse the dynamics of the RSA response to osmotic stress.

#### **Time-lapse visualisation of the growing root system**

In addition to the RSA reconstruction pipeline, we propose a method to visualise the growth of the root system. The resulting movie shows the compilation of a 7-days growth in 1 min. It gathers all the pipeline achievements in a realistic rendering, giving a rapid and accurate overview of the reconstruction of the experiment and facilitating investigations by non-computer specialists.

The movie is built using multiple interpolation techniques. First, the reconstructed architecture is rasterized in the pixel space. The apparition time of each pixel of the root central lines is determined by interpolating the spatiotemporal coordinates of the surrounding vertices in the model. Each pixel of the segmented RSA is then given the apparition time of the nearest pixel of the corresponding central line. After this operation, the binary segmentation of the whole architecture is parameterized by a time function, giving the apparition time of each pixel. This 2D + t segmentation is used to interpolate successive masks of the growing RSA that are applied on the original keyframes  $\{I_i\}_{i=1..N_t}$  of the registered 2D time series.

The root system mask obtained for a time  $T_i < t < T_{i+1}$  is applied to the image  $I_{i+1}$  to build the root architecture part while its complementary is applied

to the image  $I_i$  to gather the background part. The assembling is done by summing these two images.

When  $t$  is close to an observation time  $T_i$  (images near a keyframe), we proceed to a fading by computing the same operations using the subsequent keyframes ( $I_{i+1}$  and  $I_{i+2}$ ). Finally, the individual results are merged with fading factors that vary following a linear ramp to make these transitions unnoticeable.

After interpolating time-lapse images, we compute the position and speed at each time frame from the reconstructed RSA. They are overlaid on the image and displayed as colored tips with a colormap indicating the elongation rate.

The default parameters of the procedure result in a movie of the growing RSA generating one image every 15 min, from an initial time series composed of 21 time-frames, one every 8 h.

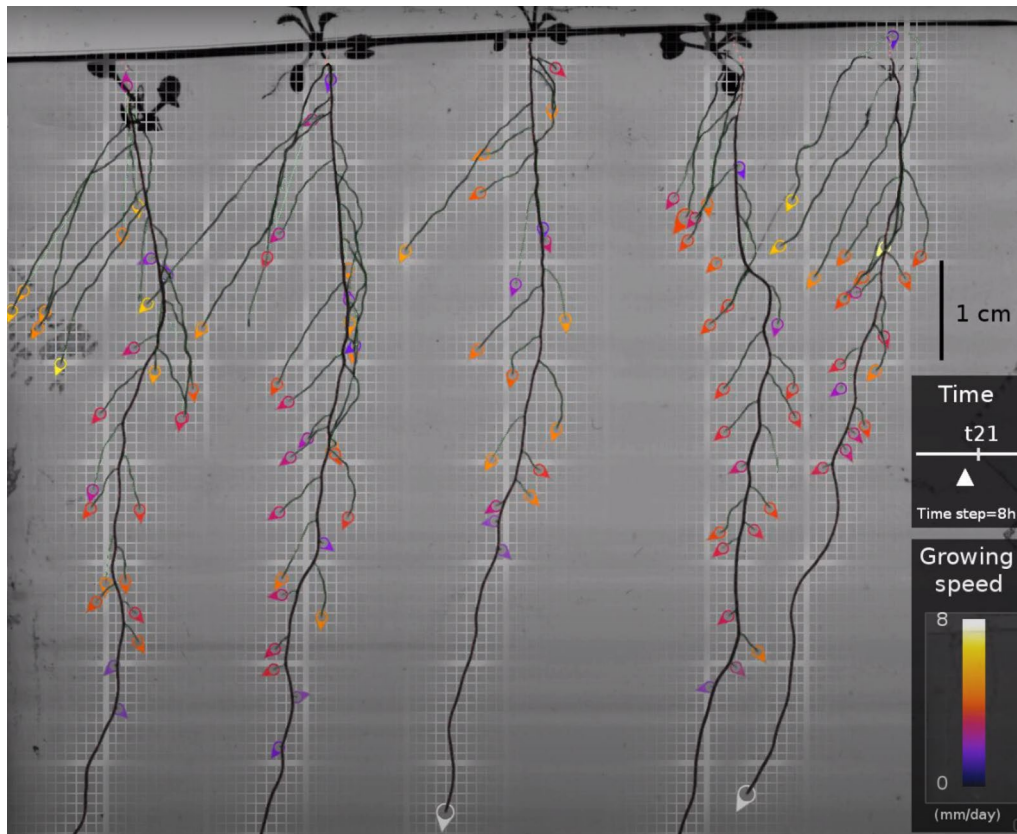
An example movie is available online at this youtube link: <https://www.youtube.com/watch?v=SWEqnOhIOU>. The provided open-source software can generate such a movie from any of our 200 reconstructed time sequences in approximately 1 min (see Fig. 10).

#### **Discussion**

In this contribution, we proposed a novel method for robust acquisition of root system architecture with its topology and development with HIRROS, a high-throughput root phenotyping platform. Our automatic pipeline combines spatial and temporal information to segment structures with their apparition time and a topological tracking algorithm to reconstruct the whole architecture, including primary and second-order roots. Our pipeline is robust to root crossing. We evaluated our method on a dataset of complex root systems and assessed performance in terms of geometric and topologic fidelity of the reconstructed 2D + t architectures. We show the ability of our method to capture spatio-temporal phenes of interest to the community, with a high level of guarantees.

We developed a high-resolution automated imaging system that can handle up to 200 Petri dishes (i.e., up to 1000 plants) and image the plates at high resolution (19  $\mu\text{m}$  per pixel) and high contrast allowing excellent root/background separation. Using a telecentric lens avoids image distortion and provides a huge depth of field combined with a micrometric controlled plate translation. HIRROS combines the high-resolution imaging as proposed by Gaggion [10] and Ohlsson [27], with high-throughput [25], offering a unique opportunity to explore the GxE root adaptive responses to a wide range of abiotic stresses by introducing hydro mineral modifications of the culture medium affecting root system development while keeping the optical characteristics of the medium.





**Fig. 10** Screenshot of the example movie. Accessible at <https://www.youtube.com/watch?v=SWEqnnOhIOU>

Petri dishes that contain transparent agar gelified medium are commonly used for *Arabidopsis thaliana* and model plants RSA analysis. However, they imply several constraints like root crossing, secondary roots juxtapositions, and shoots' movement due to 2D space limitation. These physiological and technical constraints restrict RSA analysis to static or global dynamic parameters [10]. In order to access the dynamic responses of the individual roots, we developed a dedicated image analysis pipeline.

Our proposition differs from most state-of-the-art solutions, where segmentation is performed separately on successive 2D images, which involves a subsequent step to match the successive segmentations or RSA reconstructions [23, 36]. One solution is to use a well-trained neural network to perform root segmentation successfully in each 2D image using the information from the previous images [36]. However, training this type of network requires a large training set of manually annotated images and a tedious task of reannotation in case of transfer learning. We propose a novel method incorporating a registration step to achieve accurate superposition of successive observed images and avoid tree

matching problems. The registration step produces a 2D + t stack of aligned images, which enables us to propose a pixel-wise root detector that performs root segmentation and determination of root pixel apparition time in one go. Our pipeline combines spatial and temporal information, facilitating the segmentation step to such an extent that our mean-shift detector achieves performance comparable to deep neural networks [23, 36].

However, our method has shown some limitations. When unwanted structures appear in the image field, they can be mistaken for growing roots, resulting in errors that must be detected and post-processed during the graph construction step. This point complicates the algorithm, which must adapt to specific circumstances by deploying additional post-processing operations. The pipeline could benefit from a pre-processing step integrating a neural network in charge of predicting the pixel-wise probability of the presence of a root and removing unwanted structures from the input of the mean-shift detector. That way, we could take the best of the two worlds and probably achieve even better performances, although this prospect implies that our pipeline

must be trainable. As such, our pipeline does not require any tedious annotation and training tasks.

Nevertheless, while our pipeline is fully automatic, some parameters need to be calibrated for a new experimental configuration: mean intensity of roots, mean intensity of background, and max acceptable root speed. In a trainable version of the pipeline, these parameters could be derived automatically from a dataset containing some annotated examples. By integrating these improvements, our pipeline could be transferable to other datasets.

Analysis of root system growth in Petri dishes is a unique tool for high-throughput phenotyping. However, observing root systems in 2D involves dealing with multiple root crossovers as root system complexity increases. These situations are difficult to disentangle and can hinder the tracking of individual roots. Many algorithms have been proposed to address this challenge. Most of them are greedy algorithms, processing root after root or time after time [16, 23, 29, 36] without a global view. In this work, we design a global tracking algorithm based on the region adjacency graph extracted from the segmentation of individual root segments. We evaluated the success of this algorithm in challenging cases when laterals hide below primaries or multiple laterals ( $N=3, 4, 5+$ ) intersect in a small space.

The formal framework of our algorithm relies on strong assumptions: there is no third-order ramification, and through time all the growing roots pixels are included in the final image. Our results show that some specimens do not meet these assumptions, which induces reconstruction errors, such as when plants are infected with bacterial strains. Moreover, our algorithm is designed to accurately track one primary root per plant and its second-order roots. It does not include the possibility of multiple primary roots or third-order organs. Further work will focus on adjusting the assumptions to deal with the emergence of third-order organs and to work with other species with different root system architecture patterns. To that end, the algorithm will be adapted to other assumptions regarding maximum order of ramifications, and the possibility to consider multiple root types such as seminal roots. This will allow to fit this pipeline with other architectures.

Similarly, we plan to test the pipeline on other imaging systems, such as paper pouches, rhizotubes [13], or rhizotrons. If the parameterization of the pipeline should be relatively straightforward for paper pouches and rhizotubes that have an homogeneous background, analysis of rhizotron derived images would require adjusting many parameters, especially the relative brightness of the root and the background.

Finally, future experiments will include larger time series with more developed root systems. The image space cluttering increases when the root system grows. The additional complexity between successive timesteps increases, including more root crossings and more failure opportunities for our pipeline. Thus we plan to investigate larger image series in order to identify the limits of our pipeline in terms of image space cluttering.

Performances of the method have been evaluated using standard metrics on a complex architecture dataset. Interestingly, the relative error of the total lateral length does not seem to increase with the development of the root system. This point could be crucial to study more complex architectures. In addition, we validated the pipeline's ability to estimate organs' growth rate. By providing such guarantees, we open the possibility of using these dynamic phenes in GxE studies. We illustrated this point by comparing plant development under various levels of osmotic stresses, highlighting differences in primary and lateral growth patterns. Finally, we developed a visualisation method to construct a simulated time-lapse growth movie. This method opens up the possibility of generating a large amount of realistic synthetic data with a parameterizable timestep and voxel size, leading to the design of a trainable model.

## Conclusion

The presented method provides a unique solution for high-throughput phenotyping of root system architecture. It is robust to root crossing, which is the main difficulty for accurate tracking of the individual organs from 2D image time-series. The reconstructed architecture is provided in RSML format, which integrates the topology and geometry of the root system, parameterized with the appearance time of the root segments.

Further work will focus on validating this pipeline with a larger panel of GxE variations and longer times series, adapting the hypothesis used for the topological tracking, and making the pipeline trainable to transfer it to different species or mounting configurations (box size, rhizotron). The resulting architectures will be used as topological and geometrical models to feed functional and structural root models and further understand root system development.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13007-022-00960-5>.

**Additional file 1: Figure S1.** Successive p-values obtained at each time step with the Mann-Whitney test. This corresponds to the null hypothesis  $H_0$ : "At each timestep, the mean growth of the root systems under osmotic stress and the mean growth of the control root systems are equal." (see Fig. 9 for the original data).

### Author contributions

RF contributed to method design, development and implementation with the help of CP. AC contributed to data annotation and method evaluation with the help of RF. MM conducted preliminary studies. PN and CP designed the study. RF, CP and PN wrote the manuscript. All the authors have approved the manuscript and have made all required statements and declarations.

### Funding

The authors acknowledge support from RHIZOPOLIS (Agropolis fondation, 2011–2015), EUROOT (FP7-KBBE-2011-1-2-05-289300.27-<http://www.euroot.eu/>), Preciput ANR (CS-0315/PRE-02), OPERA (LabexTULIP AO Inter-labex 2018), H2020 BreedingValues ((grant agreement ID: 101000747), and the Thomas Jefferson Fund. CP and RF have been supported by the MaCS4Plants CIRAD network, initiated from the AGAP Institute and AMAP joint research units.

### Availability of data and materials

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

### Software availability

The architecture reconstruction pipeline is supplied as an ImageJ plugin with online documentation (Plugin page: <https://imagej.net/plugins/rootssystem-tracker> [9]) and as open-source code on GitHub (<https://github.com/Rocsg/RootSystemTracker>). Once the reconstruction is achieved, the architecture can be investigated using the 2D + t model edition plugin, can be used to extract other phenes using RSML reader modules, or can be used to reconstruct a time-lapse movie.

### Declarations

#### Ethics approval and consent to participate

This declaration is not applicable.

#### Competing interests

The authors declare that they have no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

#### Author details

<sup>1</sup>CIRAD, UMR AGAP Institut, 34398 Montpellier, France. <sup>2</sup>UMR AGAP Institut, Univ Montpellier, CIRAD, INRAE, Institut Agro, Montpellier, France. <sup>3</sup>Institute for Plant Sciences of Montpellier (IPSiM), Univ Montpellier, CNRS, INRAE, Institut Agro, Montpellier, France. <sup>4</sup>Inria & LIRMM, Univ Montpellier, CNRS, Montpellier, France.

Received: 18 July 2022 Accepted: 18 November 2022

Published online: 01 December 2022

### References

- Balduzzi M, Binder BM, Bucksch A, Chang C, Hong L, Iyer-Pascuzzi AS, Pradal C, Sparks EE. Reshaping plant biology: qualitative and quantitative descriptors for plant morphology. *Front Plant Sci.* 2017. <https://doi.org/10.3389/fpls.2017.00117>.
- Boudon F, Preuksakarn C, Ferraro P, Diener J, Nacry P, Nikinmaa E, Godin C. Quantitative assessment of automatic reconstructions of branching systems obtained from laser scanning. *Ann Bot.* 2014;114(4):853–62.
- Boursiac Y, Pradal C, Bauge F, Lucas M, Delivrias S, Godin C, Maurel C. Phenotyping and modeling of root hydraulic architecture reveal critical determinants of axial water transport. *Plant Physiol.* 2022;190(2):1289–306. <https://doi.org/10.1093/plphys/kiac281>.
- Delory BM, Li M, Topp CN, Lobet G. archiDART v3.0: a new data analysis pipeline allowing the topological analysis of plant root systems. *F1000Research.* 2018. <https://doi.org/10.12688/f1000research.13541.1>.
- Diener J, Nacry P, Périn C, Diévert A, Draye X, Boudon F, Gojon A, Muller B, Pradal C, Godin, C. An automated image-processing pipeline for high-throughput analysis of root architecture in OpenAlea. In 7th International Conference on Functional-Structural Plant Models. 2013. 85–87.
- Douglas DH, Peucker TK. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica.* 1973;10(2):112–22. <https://doi.org/10.3138/FM57-6770-U75U-7727>.
- Edmonds J. Optimum branchings. *J Res Nat Stand B.* 1967;71(4):233–40.
- Fernandez R, Moisy C. FijiYama: a registration tool for 3D multimodal time-lapse imaging. *Bioinformatics.* 2021;37(10):1482–4. <https://doi.org/10.1093/bioinformatics/btaa846>.
- Fernandez R. RootSystemTracker plugin page <https://imagej.net/plugins/rootssystemtracker>. Accessed 12 Jul 2022
- Gaggion N, Ariel F, Daric V, Lambert É, Legendre S, Roulé T, Camoirano A, Milone DH, Crespi M, Blein T, Ferrante E. ChronoRoot: high-throughput phenotyping by deep segmentation networks reveals novel temporal parameters of plant root system architecture. *GigaScience.* 2021;10(7):giab052. <https://doi.org/10.1093/gigascience/giab052>.
- Griffiths M, Mellor N, Sturrock CJ, Atkinson BS, Johnson J, Mairhofer S, York LM, Atkinson JA, Soltaninejad M, Foulkes JF, Pound MP. X-ray CT reveals 4D root system development and lateral root responses to nitrate in soil. *Plant Phenome J.* 2022;5(1):e20036. <https://doi.org/10.1002/ppj2.20036>.
- Herrero-Huerta M, Meline V, Iyer-Pascuzzi AS, Souza AM, Tuinstra MR, Yang Y. 4D Structural root architecture modeling from digital twins by X-Ray computed tomography. *Plant Methods.* 2021;17(1):1–2.
- Jeuzy C, Adrian M, Baussard C, Bernard C, Bernaud E, Bourion V, Busset H, Cabrera-Bosquet L, Cointault F, Han S, Lamboeuf M, Moreau D, Pivato B, Prudent M, Trouvelot S, Truong HN, Vernoud V, Voisin AS, Wipf D, Salon C. RhizoTubes as a new tool for high throughput imaging of plant root development and architecture: test, comparison with pot grown plants and validation. *Plant Methods.* 2016;12:31. <https://doi.org/10.1186/s13007-016-0131-9>.
- Kuhn HW. The Hungarian method for the assignment problem. *Naval Res Logist Q.* 1955;2(1–2):83–97.
- Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *J Exp Soc Psychol.* 2013;49(4):764–6. <https://doi.org/10.1016/j.jesp.2013.03.013>.
- Lobet G, Pagès L, Draye X. A novel image-analysis toolbox enabling quantitative analysis of root system architecture. *Plant Physiol.* 2011;157(1):29–39. <https://doi.org/10.1104/pp.111.179895>.
- Lobet G, Pound MP, Diener J, Pradal C, Draye X, Godin C, Javaux M, Leitner D, Meunier F, Nacry P, Pridmore TP. Root system markup language: toward a unified root architecture description language. *Plant Physiol.* 2015;167(3):617–27. <https://doi.org/10.1104/pp.114.253625>.
- Luo W, Xing J, Milan A, Zhang X, Liu W, Kim TK. Multiple object tracking: a literature review. *Artif Intell.* 2021;293:103448. <https://doi.org/10.1016/j.artint.2020.103448>.
- Lynch JP. Harnessing root architecture to address global challenges. *Plant J.* 2022;109(2):415–31. <https://doi.org/10.1111/tpj.15560>.
- Mairhofer S, Zappala S, Tracy SR, Sturrock C, Bennett M, Mooney SJ, Pridmore T. RooTrak: automated recovery of three-dimensional plant root architecture in soil from X-ray microcomputed tomography images using visual tracking. *Plant Physiol.* 2012;158(2):561–9. <https://doi.org/10.1104/pp.111.186221>.
- Mairhofer S, Johnson J, Sturrock CJ, Bennett MJ, Mooney SJ, Pridmore T. Visual tracking for the recovery of multiple interacting plant root systems from X-ray  $\mu$ CT images. *Mach Vision Appl.* 2016;27(5):721–34. <https://doi.org/10.1007/s00138-015-0733-7>.
- Maurel C, Nacry P. Root architecture and hydraulics converge for acclimation to changing water availability. *Nat Plants.* 2020;6(7):744–9. <https://doi.org/10.1038/s41477-020-0684-5>.
- Möller B, Schreck B, Posch S. Analysis of Arabidopsis Root Images—Studies on CNNs and Skeleton-Based Root Topology. In Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021. 1294–1302
- Nacry P. Root Phenotyping platform webpage. 2022. <https://www1.montpellier.inra.fr/wp-inra/bpmp/en/platform/root-phenotyping-platform>. Accessed 12 Jul 2022
- Nagel KA, Lenz H, Kastenholz B, Gilmer F, Aversch A, Putz A, Heinz K, Fischbach A, Scharr H, Fiorani F, Walter A. The platform GrowScreen-Agar enables identification of phenotypic diversity in root and shoot growth traits of agar grown plants. *Plant Methods.* 2020;16(1):1–7. <https://doi.org/10.1186/s13007-020-00631-3>.
- Ndour A, Vadez V, Pradal C, Lucas M. Virtual plants need water too: functional-structural root system models in the context of drought

- tolerance breeding. *Front Plant Sci.* 2017;8:1577. <https://doi.org/10.3389/fpls.2017.01577>.
27. Ohlsson JA, Leong JX, Elander PH, Dauphinee AN, Ballhaus F, Johansson J, Lommel M, Hofmann G, Betnér S, Sandgren M, Schumacher K. SPIRO—the automated Petri plate imaging platform designed by biologists, for biologists. *bioRxiv.* 2021. <https://doi.org/10.1101/2021.03.15.435343>.
  28. Ourselin S, Roche A, Prima S, Ayache N. Block matching: a general framework to improve robustness of rigid registration of medical images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention 2000.* (pp. 557–566). Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-540-40899-4\\_57](https://doi.org/10.1007/978-3-540-40899-4_57)
  29. Pound MP, French AP, Atkinson JA, Wells DM, Bennett MJ, Pridmore T. RootNav: navigating images of complex root architectures. *Plant Physiol.* 2013;162(4):1802–14. <https://doi.org/10.1104/pp.113.221531>.
  30. Rellán-Álvarez R, Lobet G, Dinneny JR. Environmental control of root system biology. *Annual review of plant biology.* 2016.
  31. Ruffel S, Krouk G, Ristova D, Shasha D, Birnbaum KD, Coruzzi GM. Nitrogen economics of root foraging: transitive closure of the nitrate–cytokinin relay and distinct systemic signaling for N supply vs. demand. *Proceedings of the National Academy of Sciences.* 2011;108(45):18524–18529. <https://doi.org/10.1073/pnas.110868410>.
  32. Smith AG, Han E, Petersen J, Olsen NA, Giese C, Athmann M, Dresbøll DB, Thorup-Kristensen K. RootPainter: deep learning segmentation of biological images with corrective annotation. *New Phytol.* 2022;236(2):774–91. <https://doi.org/10.1111/nph.18387>.
  33. Symonova O, Topp CN, Edelsbrunner H. DynamicRoots: a software platform for the reconstruction and analysis of growing plant roots. *PLoS ONE.* 2015;10(6):e0127657. <https://doi.org/10.1371/journal.pone.0127657>.
  34. Takahashi H, Pradal C. Root phenotyping: important and minimum information required for root modeling in crop plants. *Breed Sci.* 2021;71(1):109–16. <https://doi.org/10.1270/jsbbs.20126>.
  35. Yasrab R, Atkinson JA, Wells DM, French AP, Pridmore TP, Pound MP. RootNav 2.0: deep learning for automatic navigation of complex plant root architectures. *GigaScience.* 2019;8(11):giz123. <https://doi.org/10.1093/gigascience/giz123>.
  36. Yasrab R, Zhang J, Smyth P, Pound MP. Predicting plant growth from time-series data using deep learning. *Remote Sensing.* 2021;13(3):331. <https://doi.org/10.3390/rs13030331>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

