GROUND PENETRATING RADAR FOR THE ESTIMATION OF BULKED ROOT MASS

A Dissertation

by

BRODY L TEARE

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Dirk B. Hays
Committee Members,	Xinyuan Wu
	Mark Everett
	Nithya Rajan
	Sorin Popescu
Head of Department,	Dirk B. Hays

December 2021

Major Subject: Molecular and Environmental Plant Sciences

ABSTRACT

Plant root phenotyping is a difficult task because of the opaque nature of soil, leading knowledge of root traits to lag behind above ground traits. While a large multiplicity of methods for observing roots have been described, each is limited in some way. This difficulty in observing plant roots causes difficulty for researchers and breeders working with root crops, such as cassava. Cassava is a crop with a long production cycle and low reproduction rates. As a root crop, assessing root properties such as time to bulking and marketable mass is a major goal for breeders. It has been proposed that GPR could be used to non-destructively estimate cassava root mass. Previous publication has demonstrated the potential for such a tool, but further work is required before GPR can be useful to breeders.

In this body of work we explore the realm of root phenotyping and address the issue of rapid and non-destructive estimation of root mass by GPR. We demonstrate the ability to correlate features of GPR data with bulked root mass in controlled field conditions, and demonstrate the effect of soil water on GPR data in regards to root mass. We proceed to expand that work from controlled conditions into an actual cassava field. There we were able to build upon the predictive model derived in the earlier chapter, confirming that GPR data relate to bulked root mass in unoptimized field conditions. Lastly, we evaluate the ability of predictive models to generalize across locations and genotypes. While strong Pearson correlation up to 0.91 between predicted and observed root mass is achieved, the presented models require calibration by location.

DEDICATION

To my parents who believed in me, my wife who put up with me, and my children, who may one day understand why I did this.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Dirk Hays (advisor) and Professor Nithya Rajan of the Department of Soil and Crop Sciences, Professors Sorin Popescu and X. Ben Wu of the Department of Ecosystem Science and Management, and Professor Mark Everett of the Department of Geology and Geophysics.

Significant help in data collection, processing, analysis, and document editing was received from the Dirk Hays Lab, namely, Henry Ruiz-Guzman, Matthew Wolfe, Afolabi Agbona, Tyler Adams, and Iliyana Dobreva. The field work of growing, maintaining, and harvesting cassava plants used in Chapters 4 & 5 was done by workers at the CGIAR centers CIAT and IITA in Colombia and Nigeria, respectively, and they deserve recognition for their high caliber work.

Other work was performed by the student independently.

Funding Sources

The work was funded in major part through a grant from the National Science Foundation under Grant Number 1543957—BREAD PHENO: High Throughput Phenotyping Early Stage Root Bulking in Cassava using Ground Penetrating Radar.

The work was also made possible by Texas A&M University through a graduate teaching assistantship from the Department of Crops and Soil Sciences.

The content of this work is solely the responsibility of the author and does not necessarily represent the opinions of the funding agencies.

NOMENCLATURE

DAG	Days After Germination
ERT	Electrical Resistance Tomography
EIT	Electrical Impedance Tomography
GWC	Gravimetric Water Content
Voxel	3D analog to a pixel
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
EM	Electromagnetic
C.I.	Confidence Interval
S.E.	Standard Error
M.A.E.	Median Absolute Error

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

List of Tables

Page

1. INTRODUCTION - NON-DESTRUCTIVE BULKED ROOT MASS ESTIMATION USING GROUND PENETRATING RADA

1.1. Background

The continuing increase in world population, coupled with ongoing climate change, is putting pressure on food supplies. This is especially true in developing countries where subsistence farming is still common [1]. Root and tuber crops are the second most important crop group behind cereals, and of these, potato (*Solanum tuberosum*) and cassava (*Manihot esculenta* Cranz) represent the two most important globally [2]. As with all crops, extended time in the fields increases risks. This was well demonstrated in 2019 when farmers in the American Midwest had to leave potatoes in the field due to flooding and Idaho farmers lost potato crops to early cold. Meanwhile, cassava is the security food crop of choice in parts of Africa, and climate change is increasing risks for subsistence and commercial farmers there [3]. These problems could be alleviated by early cultivars which require less time in the field. Further, growers desire short-term and early varieties [4, 5]

Potatoes and cassava share some characteristics which make breeding for earliness difficult. The primary difficulty is simply that you cannot see the root or tuber growing beneath the soil, and therefore cannot determine when it has bulked. Both crops exhibit severe inbreeding depression, are highly heterozygous, and sibling genotypes can vary widely, coupled with low vegetative propagation rate, means that genotypic replications in breeding populations are very small [6, 7]. The consequence of this is that breeders

cannot afford to dig up trials early to discover whether or not they have bulked. It is possible that breeders could have early varieties and never know it because they have waited the standard time to harvest their trials. To more effectively discover early varieties, breeders need a non-destructive tool or method which will allow them to estimate bulky root and tuber mass without having to dig up the plant. Additionally, because breeding nurseries can be very large, the method must allow for high-throughput and timely results. Several methods have been proposed for measuring roots, including x-ray, magnetic resonance imaging, electrical resistance tomography, and ground penetrating radar (ARPA-E, 2019). Of these, ground penetrating radar (GPR) is the closest to being field deployable, affordable, and high throughput.

Our long term objective is to assist in the release of novel early bulking cassava through GPR. The **primary objective** of this proposal is to increase our understanding of GPR as a high-throughput and non-destructive tool for selecting early bulking cassava and potato genotypes. We **hypothesize** that data collected from ground penetrating radar can be used to estimate the mass of bulked roots and tubers. To this end, we have formed a partnership with IDS Georadar, the International Center for Tropical Agriculture (CIAT), the International Institute of Tropical Agriculture (IITA), and the J.R. Simplot Company. IDS is a GPR manufacturer and can provide hardware and knowledge. CIAT and IITA are both international research centers which specialize in cassava and can grow and harvest cassava trials, as well as provide expertise on the crop. Simplot is a major potato producer in the US which is willing to grow and harvest potato trials. This collaboration between a hardware manufacturer, crop specialists, and agronomists and

remote sensing experts at Texas A&M University will allow us to test our hypothesis and achieve our goal through the following specific objectives:

1) **Define novel GPR data processing methods for the estimation of plot level root mass in bulked cassava plants with ground penetrating radar as a high-throughput process.**

   We previously demonstrated the ability to correlate GPR returns with cassava root mass [8]. This works serves as both a validation and the extension of that work. Their methodology was meticulous and time consuming. To be a viable tool for breeders and producers, methods must support high-throughput data acquisition and processing. The foundation of this study will be to develop and prove GPR as a high-throughput option for estimating root mass. This will be done using multiple cassava phenotypes that represent the range of architecture and bulking habits to define the novel methods needed to correlate GPR returns to the harvest mass of the roots, and used to build an algorithm for predicting bulked root mass.

2) **Define novel GPR base algorithms needed to estimate cassava root bulking rate.**

   A single snapshot of bulked root mass is of limited usefulness to breeders while an estimation of bulking rate and an indication of when bulking begins is much more useful. The next obstacle, therefore, is to show that GPR can be used to estimate bulking rate. To minimize travel costs, several plots of cassava will planted with staggered dates, and harvested at a single date. This will simulate a multi-date data acquisition as each plot will be at a different physiological stage

of development. We hypothesize that the roots will demonstrate a corresponding range of sizes. Using the same field and processing methods developed in the previous step, we will estimate the root mass and derive the bulking rate. Comparing the estimation to the harvest data will be a test of the robustness of the prediction algorithm.

3) **Define the GPR data processing methods tuber mass estimation algorithms needed for potato breeding.**

   To show the widespread applicability of GPR as an agricultural tool, we will mimic the experiments in objective 1 using potatoes. We hypothesize that a technology which can measure bulked roots in cassava could also measure bulked tubers in potato. It is entirely reasonable, however, that the application will require adjustments to both the collection methods and the processing. Using the established methods, we will evaluate the results, and if necessary, adjust the methods. Demonstrating the viability of the technology in potato will increase the visibility and improve chances for further funding and research in mainstream crops.

Our proof of concept using GPR to measure root mass proves the viability and applicability of GPR for measuring bulked roots. We expect that our contribution using a novel air-launched GPR antenna will lead to a commercially available tool. This tool in the hands of producers will provide decision support, reducing costs and improving results. The greater impact, however, will be when plant breeders, using this technology, discover early varieties, associate genetics, and produce cultivars which improve the

lives of millions. The widespread introduction of early root crops will increase productivity and reduce producers' vulnerability to climate change.

1.2. Literature Review

1.2.1. Cassava (Manihot esculenta)

Cassava is a tropical plant grown for its starchy bulked roots. More than 800 million people depend on it as a staple food, and it is an important source of starch for even more [9]. It is commonly grown by subsistence farmers because it is hardy and harvest can be delayed until needed [10]. Cassava is the fourth most important basic food crop in the tropics and the global harvest has increased by more than 25% since 1999 and has nearly doubled in some regions [6, 11].

A perennial shrub which commonly grows from one or two main stems, cassava is cultivated as an annual. The leaves can be consumed by humans and are high in protein (20-25% dry weight), though it is more common for the foliage to be fed to cattle. The leaves, along with the stems and roots, contain cyanogenic glucosides that can be poisonous if not processed properly. The root remains the most economically and socially important part of the plant [6]. Development of improved varieties has greatly benefited growers across the globe; however, many farmers still grow varieties that take up to 2 years to produce a harvest [10, 11] Thus, there is a distinct need and desire to develop varieties which bulk early [4, 6, 8].

Cassava breeding has languished behind other important crops. Rigorous breeding programs did not begin until the creation of the CGIAR centers in the 1970's. Cassava is diclinous and monoecious, meaning flowers are either male or female and both occur on

a single plant. Female flowers bloom before male flowers, but self-pollination can still occur. Despite this, cassava is an outcrossing species and exhibits severe inbreeding depression. It's theorized that varieties could be developed which take to inbreeding, as was done with corn, though it could take considerable time to do so [11]. Therefore, every variety is a hybrid and cultivars are propagated using stem cuttings. Plants grown from true seed do not exhibit the same root bulking as those grown from stem cuttings, and therefore are not good predictors of yield. This means two growing seasons are needed before yield selections can be made. Each plant produces enough material to yield 3 to 12 planting stakes. This is a very low multiplication rate and is a key feature of cassava breeding [6]. With such low numbers in breeding trials, breeders cannot afford to dig up plants early to check for bulking.

1.2.2. Potato (Solanum tuberosum)

Potato is an herbaceous crop grown for its enlarged, starchy tubers, which range from spherical to ovoid in shape. It originates in the South American Andes and is grown mostly in temperate climates, but is increasingly grown in more tropical climates [12]. Potato is now the fourth most important food crop worldwide [12, 13]. It has been primarily grown by developed nations, but is becoming more popular in developing nations as a security crop [12, 14]. Potato is the fourth most important food crop globally, with more than 200 million hectares grown annually [13]. The value of potato harvest in the United States is estimated at more than $4 Billion in 2017 [15]

Potato varieties can be classified based on the time to maturity, ranging from very early varieties to very late varieties (65 – 130 days). Varieties with longer maturity generally

out-yield early types, and the earliest types are harvested young, sold as New potatoes [16]. In developed countries, consumption is shifting from fresh potatoes to processed products. This shift to processing puts greater emphasis on harvesting the crop when the tubers are the correct size [14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16][14, 16](Camire et al., 2009; Furrer et al., 2018)(Camire et al., 2009; Furrer et al., 2018)(Camire et al., 2009; Furrer et al., 2018).

Cultivated potatoes are generally tetraploid and highly heterozygous. Because of significant inbreeding depression, each seed produces a unique plant and variability has been kept high in breeding populations [13]. Fortunately, potato tubers can be used to clonally propagate varieties. This has been the major mode of cultivar creation for centuries [17]. Consequently, potato breeders face unique challenges in assessing the wide variety of phenotypes produced by even a single cross. As with cassava, breeders cannot afford to continually monitor tuber progress on all their trials.

1.2.3. Ground Penetrating Radar

GPR is primarily a geophysical tool used for exploring the subsurface of the earth ranging from several centimeters to kilometers [18]. It functions by emitting a very short pulse of electromagnetic energy (EM) and then “listening” for reflections. The instrument measures the amplitude and the travel time of the reflections. Reflections are caused by changes in the dielectric permittivity of the material the wave is traveling through (i.e. the ground and buried objects). Dielectric permittivity describes the ability of a material to store EM energy as dipole movement [19]. Besides reflection, the

emitted signal can also be scattered, transmitted, refracted, and absorbed [18]. Only a small portion of scattered waves are measured, and refracted waves may or may not be captured by the receiving antenna. Transmitted waves are lost unless they are later reflected or scattered back to the antenna. Absorbed energy is lost [18, 20].

The detection threshold and penetration depth are both related directly to the frequency of the EM pulse. The two are related in opposing manners so that as one increases, the other decreases. Increasing frequency improves detection threshold, meaning that smaller objects can be detected and large objects have better definition [21]. In general, the smallest object for which a discreet return can be received has a diameter of at least ¼ of the EM pulse wavelength. The wavelength is defined by the frequency of the pulse and the characteristics of the medium through which it is traveling, primarily the relative permittivity. Thus, the minimum detection threshold changes with soil water content. Increased soil moisture results in better detection (shorter wavelength), but less penetration (depth). In each application, these considerations must be balanced [18].

GPR has established uses in geoscience, civil engineering, archeology, glaciology, and military [22]. GPR has yet to make a widespread debut in agriculture, though it has seen use in the study of forest ecosystems and other environmental applications [23, 24].

1.3. Approach

1) **Define novel GPR data processing methods for the estimation of plot level root mass in bulked cassava plants with ground penetrating radar as a high-throughput process.**

GPR is a good candidate for field measurements because it has relatively low power consumption, has conformation adaptability, and can acquire data rapidly. However,

most off the shelf systems are designed for use on relatively smooth surfaces. Physical adaptation of the GPR is required to move it through agricultural fields quickly. This is the first hurdle which must be overcome. Using hardware which allows rapid prototyping and flexibility, we will utilize an iterative design method to build a GPR platform appropriate to high throughput use in the field.

As this first objective is an extension of the earlier work done by Delgado et al, we will take their work as a template and measure bulked cassava in several varieties. The main purpose is to show that this can be done at a large scale and quickly. Delgado et al measured multiple transects around the plants, we will also, but instead of running multiple passes, we will use an antenna array which allows multiple channel acquisition. This means that with every pass, multiple parallel transects will be measured, greatly reducing the amount of time spent in the field. Additionally, rather than using a ground coupled antenna, we will use an air launched antenna, removing the need to provide a smooth surface to work on. An air launched antenna will move through the field easier, be less prone to damage, and increase the throughput. It will introduce some difficulty, though. Using an air launched antenna will reduce the precision of the measurements and require surface detection and flattening. These challenges are not insurmountable and can likely be solved through automated processing.

Field Design. -The field design will consist of four cassava varieties with diverging phenotypes to provide maximum variability. Each variety will be planted in plots, replicated four times. Because we are not analyzing the effect of the field on the plants, we don’t need to concern ourselves with plot layout for normalizing variance. Put

another way, we don't care where the variance comes from, only that we capture it to build a robust correlation between the signal and the root mass. This allows a simplified field design with cassava varieties grouped for simpler harvest and data labeling.

Software/Processing -We will build a GPR processing software which is designed for easy, rapid batch processing. This software will utilize open source Python GPR libraries, algorithms developed from the literature, and novel processing and analysis techniques developed by our team. This software will allow a wide variety of pre and post processing techniques, as well as quantitative analysis metrics. Utilizing his powerful tool, we can find the best way to process and analyze our unique data. Some of the tools we will implement include signal dewow, band pass filter, surface detection, Hilbert transformation, and Kirchoff migration.

We will follow Delgados steps for processing data and attempt to correlate the results with the harvested root mass. We will also use different and possibly novel techniques to extract features which correspond to the bulked roots. These quantitative features will be used to build a predictive algorithm which can be used to estimate bulked root mass. We will withhold some harvest data from the correlation to validate the predictive algorithm.

Anticipated problems and possible solutions –While it is reasonable to expect Delgados methods to work for this process, it would not be surprising if it doesn't. Using an air launched antenna may introduce significant error in the data, such as GPR returns from above ground objects, movement of the antenna relative to the soil surface, and variation in the look angle of the antenna. These problems may be overcome in

processing. It is possible to remove objects above the surface and to correct for variation in antenna height by detecting the surface and flattening it in processing.

2) **Define novel GPR base algorithms needed to estimate cassava root bulking rate.**

Field Design -Having validated the air launched multi-array GPR system as a valid method for estimating bulked cassava root mass and built a predictive algorithm, we will apply the same technology to a field with plants of varying age. Again using 4 varieties and 4 replications, we will plant new material every month for 7 months. All of the plants will be scanned and harvested at the same time. The staggered planting will provide root sizes from a very young age where bulking is not expected up to a mature age where full bulking is expected. This is a simulation of multiple scans across the growing period of cassava.

Processing -We will utilize the software and processing workflow developed for Objective 1, applying the predictive algorithm to the scan features. Applied across the range of plant maturity, we will discover the lower threshold for bulked root detection, and show the ability of GPR to differentiate between root sizes. This will also work as further validation of the predictive algorithm. It is possible that the algorithm will not be valid for both datasets because of variations between them such as soil moisture. If this is the case, we will correlate the extracted features to the harvest data independent of the previous algorithm.

3) **Define the GPR data processing methods tuber mass estimation algorithms needed for potato breeding.**

Field Design -Plots of 4 different potato varieties representing a range of tuber sizes and shapes will be planted in three replications. These will also be stagger planted across 3 months for a total of 36 plots by 20 hills each. These will be scanned in the same manner as the cassava, using the same hardware, by pushing the cart alongside the row, with the antenna array angled towards the base of the plants. Additionally, we will scan using a bi-static type antenna array. The bi-static array will straddle the row and capture reflections at a different angle. This may provide a more complete picture of the tuber zone, but until we try it, it is difficult to say if the information will be appreciably different.

Processing -The GPR data will be processed using our software package and the workflow developed for cassava. As mentioned before, potato tubers grow in a different orientation than cassava roots. They're also shaped differently. We therefore expect the workflow to need adjustment. The primary correlation metric will be tuber mass at the plot level. Because tuber diameter is also very important to producers, we will attempt to find features which correlate to the mean tuber diameter, or the weighted range of diameters.

Anticipated problems and possible solutions –Potatoes exhibit different growth habits than cassava. Where cassava roots extend from the plant like a star, potato tubers grow in bunches directly under the plant. The bunched nature of the tubers may make them appear as a single large object in the GPR data. This may require different processing than cassava and may render the multiple transects of the antenna array redundant. In

this case we will investigate methods of combining transects to improve the data quality and may need to correlate to a single scan line instead of interpolating between several.

1.4. Impacts and Future Directions

The primary goal of this research is to provide a tool to root crop breeders which will allow them to make useful estimates of their crops bulking rate and status without having to damage the crop. The impacts of such a tool are wide ranging. Initially, we hope to see cassava varieties developed which mature in a fraction of the time current popular cultivars take. This would improve the livelihood of cassava farmers by reducing the risk of having a crop in the field for a long time, increasing the productivity of their land, and increase their income. The corresponding increase in cassava production would make it more accessible to non-farmers and represent a real increase in their standard of living by reducing food costs. Beyond the impact to cassava growers, this technology has the potential to impact all root and tuber crops, as demonstrated in the third objective with potatoes. Besides breeders, this can be used as a tool by producers to help them decide when to harvest, or for logistics to better predict yield and arrange the necessary labor and equipment for handling it.

From here, the proof of concept, the idea needs further development. Our lab plans to continue the development of the software, and to release it to researchers, or license it to companies. Ideally, we would like to create a web based application which can automatically process GPR data, so that users need not have expertise in data processing. The hardware also needs further work. The antenna should be optimized, selecting the best radiation patterns and the best number of antennas and channels. This work is

beyond the scope of our lab and must be picked up by a company with the expertise and resources to design and test GPR antennas. Because of our collaboration, it seems likely that IDS Georadar will do this.

2. FROM LAB TO FIELD, CURRENT AND EXPERIMENTAL METHODS IN PLANT ROOT PHENOTYPING

2.1. Introduction

Scientists have long recognized the importance of plant roots in agricultural systems, but the difficulty of studying roots as compared with the relative ease of studying aboveground plant systems has kept many researchers and breeders focused on the surface portion. Our awareness of the complexity and significance of root systems has coincided with slowing advances in yield from breeding; thus, a greater focus on roots has emerged. Calls for concentrated efforts to phenotype plant roots have launched a decades-long campaign to improve root study abilities, increase general knowledge of root systems, and accelerate breeding gains to combat climate change and support a growing population [25–27].

Despite this enthusiasm, studying roots remains a significant challenge for several reasons. Beyond the obvious opacity of most growing mediums, the soil environment is hugely complex and varies vastly, spatially and temporally. To overcome this obstacle, researchers often modify the growing medium in ways ranging from slight intrusions (minirhizotrons) to complete destruction (excavation) to using a completely synthetic medium (agar gel). While popular, the power of these studies to convey useful information on root traits depends on trait heritability and the ability to make reasonable correlations based on field conditions [28].

Several recent reviews have addressed the advances in root phenotyping, but each is limited either in scope or depth [27, 29, 30]. It does not appear that a comprehensive

review of root phenotyping methods has been published for some time. This review aims to consider all of the major and experimental means of examining roots in lab, greenhouse, and field settings. A treatment of the methods and applications is also included, giving special attention to recent publications.

2.2. Clear Medium Methods

Use of a clear growing medium solves the largest problem in studying root systems: the inability to see roots. Researchers have found various methods for growing plants in a manner that exposes the roots to visual observation and measurement. Here we attempt to present these methods in order of increasing complexity, share some recent examples from the literature, and discuss considerations.

2.2.1. Filter Paper

Plant seedlings may be grown on simple filter paper, which is prepared with water and occasionally nutrients. Filter paper is also known as germination paper or, when slightly modified, germination pouches or pouch and wick [31–33]. The method is straightforward: paper is soaked in water and a seed is placed on top, allowing the seed to draw water from the paper. This can be modified by placing the paper inside a transparent bag. The seedling can then be observed visually up to about 10 days after germination (DAG) [31]. In addition to visual observation, imaging of the roots is relatively straight forward by means of a digital camera or a flatbed scanner.

Falk et al. modified germination pouches by standing them upright and suspending them (much like a hanging file-folder system), allowing a high density arrangement within a growth chamber [33]. The seedlings were easily transported to an imaging platform,

which maintained consistent lighting and focal placement for each image. Each germination pouch included a barcode label, which the imaging software was able to read for automatic labeling of the image files, reducing the workload as well as the opportunity for labeling errors. Similarly, Hund et al. also kept the pouches upright for growth [31]. In order to image the roots, pouches were held at a fixed distance from a traditional flatbed scanner which was rotated onto its side, holding the seeds in place on the paper with paperclips.

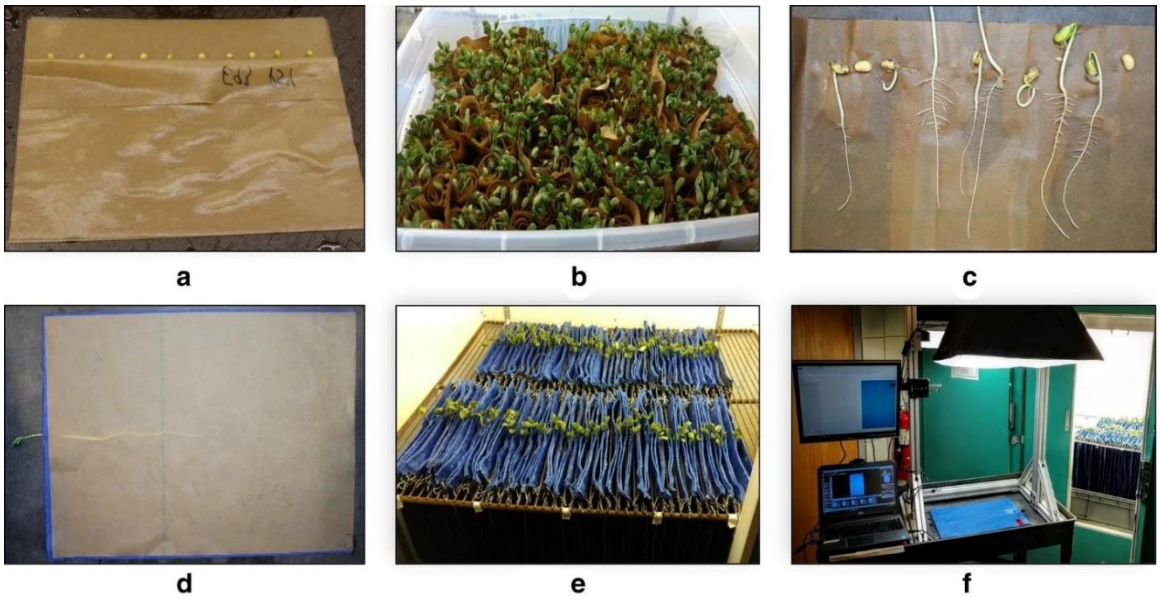

Figure 2.1 Filter paper root phenotping platform. ***a****.10 seeds rolled into germination paper.* ***b****. Plants germinate in a growth chamber.* ***c****. Seedlings are subsampled for transplanting onto blue germination paper.* ***d****. Single seedlings are covered with moistened germination paper and sandwiched to make an experimental unit.* ***e****. Each seedling is bound with binder clips and hung file-folder style in a growth chamber, the bottoms submerged in water.* ***f****. Individual seedlings can be easily removed and imaged in a consistent fashion for automatic data processing and filing. Figure is from Falk et al. and used under CC-BY 4.0 License* [33]

2.2.2. Petri Dish

Petri dishes filled with agar or a similar gel medium have been used in laboratories for

decades. They are frequently used to study the roots of young plants due to their straightforward nature. In the simplest case, seeds are germinated, then placed in a dish to grow, with roots available to observe at any time. Dishes may need to be tipped vertically so that plants correctly experience gravity, thus keeping roots on the surface [34]. French et al. used petri dishes to test their novel root image analysis software, which they used to automatically find root tips, trace the roots, measure curvature, and detect the onset of gravitropic response [34]. In that study, imaging was done using a standard DSLR camera mounted on a tripod.

The petri dish method has been adapted and modified many times to overcome issues or facilitate workflows. Bengough et al. created thin rectangular chambers of plastic with holes at the top to allow gas exchange and growth of the shoot [35]. In recognition of the danger of anoxic conditions, a small air gap was left in the center plane, where the roots primarily grew. A thin layer of gel was added to each large face of the chamber to provide moisture and to prevent condensation on the plastic. This design also facilitated imaging on a flatbed scanner, which at the time would have had superior resolution to digital cameras. Nagel et al. also modified square petri dishes so that the shoot was able to grow outside the plate [36]. Compared to plants grown the more traditional way – with shoots inside the dish – the researchers found significant differences, including longer primary roots, greater root area, and larger shoot biomass. This modification of the petri dish also allowed for the use of an automated imaging platform which the authors constructed. The differences compared to more common petri methods illustrate the hugely complex nature of studying roots, and the effects of environment.

Futsaether and Oxaal did away with the gel completely, instead using a thin rectangular chamber with a 1.2mm thick space filled with glass beads, allowing aerated nutrient solution to be circulated to the roots while the shoot grew in open space [37]. This method has the great advantage of easily controlled nutrient availability. Additionally, it allows for variation in both particle size and distribution in the glass beads, as well as high color contrast for root imaging.

The methods so far have allowed two-dimensional study of roots. For three-dimensional studies, the methods have been extended to deeper containers. Fang et al. studied root response to phosphorus in rice and soybean in 3D [38]. To do this, they grew the plants in clear plastic cylinders 12 cm in diameter, 20 cm tall, and filled with a clear gel. Roots were imaged by rotating the cylinders in front of a 3D laser scanner, then reconstructing them on a computer. Clark et al. extended the cylinder method for use in a high-throughput phenotyping pipeline [39]. The growth cylinders were placed in a larger square container of water, which provided correction for light refraction caused by the round growth chamber; this made it possible to use a standard DSLR camera and light box. Further, a turntable was constructed to rotate the sample in a controlled, consistent fashion, with the authors being able to capture 40 images through 360° of rotation in just 4 minutes, all with a resolution of 50 μm per pixel.

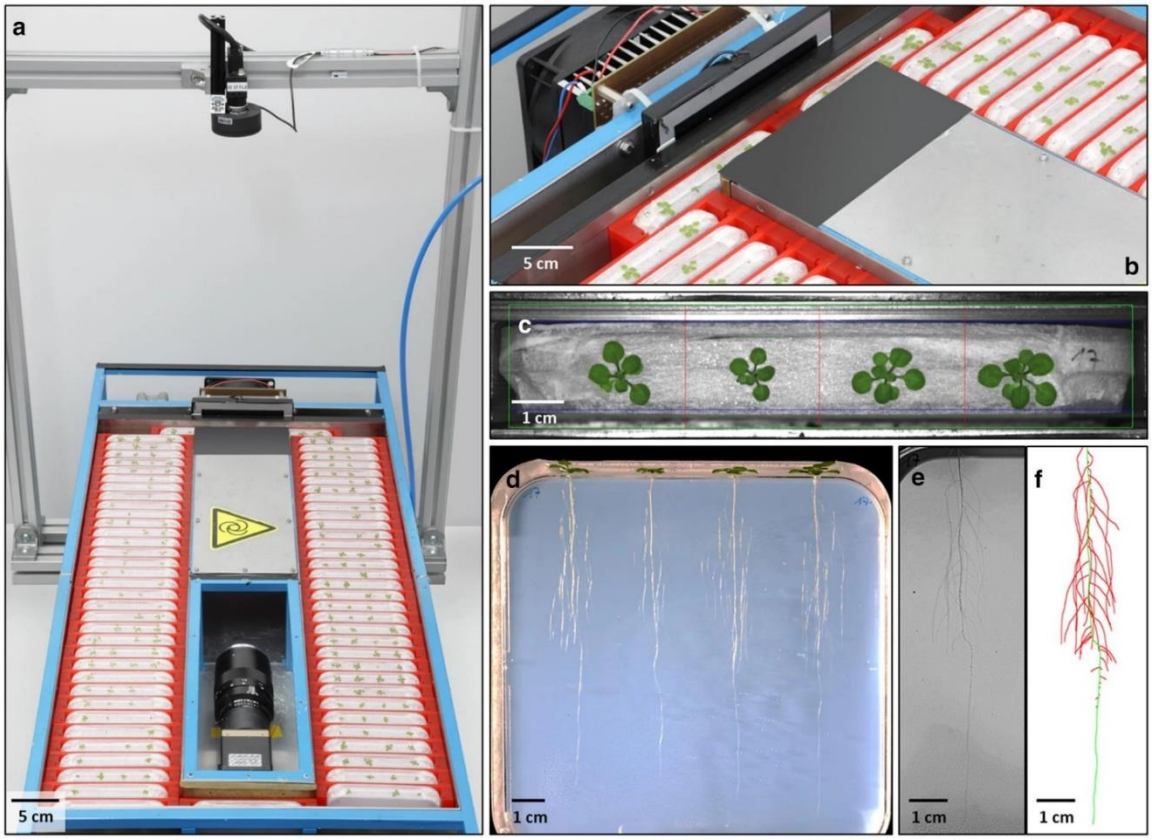

Figure 2.2 The GrowScreen-Agar platform. ***a & b****. The platform is mechanized and able to automatically rotate each sample through the imaging chamber at regular intervals.* ***c & d****. The modified petri dish is held vertically, so the roots experience gravity correctly. In this image, the shoots are kept inside the petri dish.* ***e & f****. The standardized imaging allows for computer detection and tracing of the roots. Figure is from Nagel et al. and used under CC-BY 4.0 License* [36].

2.2.3. Considerations

These types of methods allow high-resolution and short-duration study of root systems. In the experiments summarized in this review, plant ages were generally less than 10 days at the termination of the study, though they ranged up to 20 DAG in some cases [33, 37, 39, 40]. The age of plants is often limited by the size of their container – once roots encounter the edge of a container, their structure is undeniably altered. Bengough et al. concluded that such methods are useful for determining genetic potential and

discovering variation, while Gregory et al. noted their suitability for studying root exudates [35, 41]. Gruber et al. demonstrated the usefulness of gel for studying the effect of nutrient deficiencies on roots, while also noting the need for care in choosing the appropriate gel due to most types including some level of nutrients [42]. Kuijken et al. expounded on the ability to quickly calculate the heritability of root traits from lab studies, but also cautioned that traits discovered by non-field methods need to be confirmed in field conditions [28]. Passioura warned of the possibility for hypoxic conditions in all types of pot studies, regardless of medium [43].

Despite the artificial root environment, similarities to soil-grown roots have repeatedly been found. Liao et al. found that bean plants grown both on filter paper and in sand filled pots produce similar root angles [44]. Clark et al. found that rice plants grown in gellan gum had significantly shorter average lateral root length, but similar crown root numbers to potted plants grown in sand. Other differences (or similarities) depended on whether or not the sand was aerated [39]. Hargreaves et al. compared the roots of barley plants grown in gel and soil, observing similar angles and root numbers but greater lengths in the gel, hypothesizing that it was caused by lower resistance and nutrient availability [40]. On the other hand, Richard et al. compared growth pouches to clear pots for studying root traits in wheat seedlings. They calculated greater heritability in the clear pots than in the pouches, and found significant differences between root angles and length. They were also able to associate the seminal root angle at 5 DAG, and the seminal root number at 11 DAG with drought tolerance [45]. These findings agreed with previous research done in 2D gel chambers [46].

2.3. Field Methods – Classical

When studying roots in the field, there are few substitutes for a good shovel – excavation and coring techniques continue to be standard methods and are widely applied [47]. Excavation gives a complete look at a root system, but the cost and labor can be quite high. Whole plant excavation requires care to avoid losing roots during digging or washing and may result in changes to root architecture when the roots are removed from the supporting matrix. Several alternatives have been described to combat these drawbacks. In this section, common methods for observing roots in the field are discussed.

2.3.1. Trench

The trench method attempts to preserve root architecture by leaving the plant in the soil while exposing the roots on a 2D plane, allowing for the mapping of roots. In concept, the method is very simple – a transect adjacent to a plant is chosen, and a trench is dug along that line to allow access to the roots. In practice, care must be taken to smooth the observation plane with a sharp knife or gentle water sprayer [48]. In this method the primary measurement is root intersections, or simply root counts. Use of a grid overlay can aid the process; alternatively, a plastic sheet can be used to trace roots and mark intersections. By taking small core samples within the trench profile, a calibration can be made to relate root counts to root length density. 3D sampling is possible by removing subsequent layers from the soil wall. Trenching also allows soil features such as horizons and pore structure to be studied; however, Van Noordwijk et al. noted that mixing of the soil horizons may make the site unsuitable for future studies [49]. Despite

this, trenching was shown to be comparable to core methods for estimating root traits, and can be completed in a fraction of the time [50].

Nemoto et al. used trenches to compare the root systems of 255 rice varieties [51]. Using a 60 cm deep trench after harvest they were able to correlate an abundance of roots in deep layers with drought resistance. To further reduce labor, visual scoring was used to estimate root abundance, rather than counting individual roots. De Azevedo et al. sought to establish root length density as a standard metric for root assessment. Utilizing the trench method, they demonstrated use of the metric in sugarcane, and were able to derive root length density from root intersection counts, comparing them to results from the core method [52]. They found good agreement between the methods, and also noted that trenching requires just 10% of the man hours needed for core sampling.

2.3.2. Pinboard

The pinboard method can be considered an extension of the trench method, whereby a grid of pins attached to a board is inserted into a trench wall, and the soil monolith surrounding the pins is removed intact. Once removed, the board holding the pins and soil is laid flat and the soil is washed away. The root structure remains in place so angles and other spatial relationships are preserved [49]. The pinboard method has been popular in previous decades [53–56], and while it is still cited as a possible method [57, 58], no research published in the last five years was found in this review. Instead, it seems most researchers prefer the root-box pinboard method described by Kono et al. [59]. While the root-box pinboard method is not necessarily a field method, it will be reviewed here as a modification of the well-established trench-pinboard method.

The root-box pinboard method (alternatively, rhizobox) utilizes a large soil container with one transparent side. Kono et al. used a 25 cm long x 2 cm wide x 40 cm deep box, but the size varies by author. The box is either constructed with a grid of pins piercing it, or is made such that pins can be inserted later. The clear window allows repeated observations to be made while the plant is growing. When the growth period is over, the soil is washed out and the roots are held in place by the pins. The rhizobox may be considerably less labor intense than the trench-pinboard method and allows for control over environmental conditions if performed in a greenhouse. Further instructions were given by Delory et al. [60].

Kano-Nakata et al. was able to associate increased root mass for rice under drought conditions using this method [61]. Singh et al. studied young sorghum in pinboard boxes by harvesting subsamples at different times. They were able to associate the first nodal root flush with the 4-5 leaf stage, and therefore suggest rhizoboxes as an appropriate standard for studying root traits in sorghum, preferable to the lab methods commonly used in maize and wheat. Thangthong et al. used 120 cm deep boxes to examine drought response in different peanut varieties, finding that the variety with the deepest roots under drought was best able to maintain shoot mass [62]. In 2020, Miyazaki and Arita used root-box pinboards to study the rooting habits of upland rice variety NERICA 4 [63]. Not only did they find very strong correlation ($r=0.976$) between total root length and number of root branches, but they also showed that subsurface irrigation can affect the distribution of roots, with deeper irrigation correlating to deeper root mass.

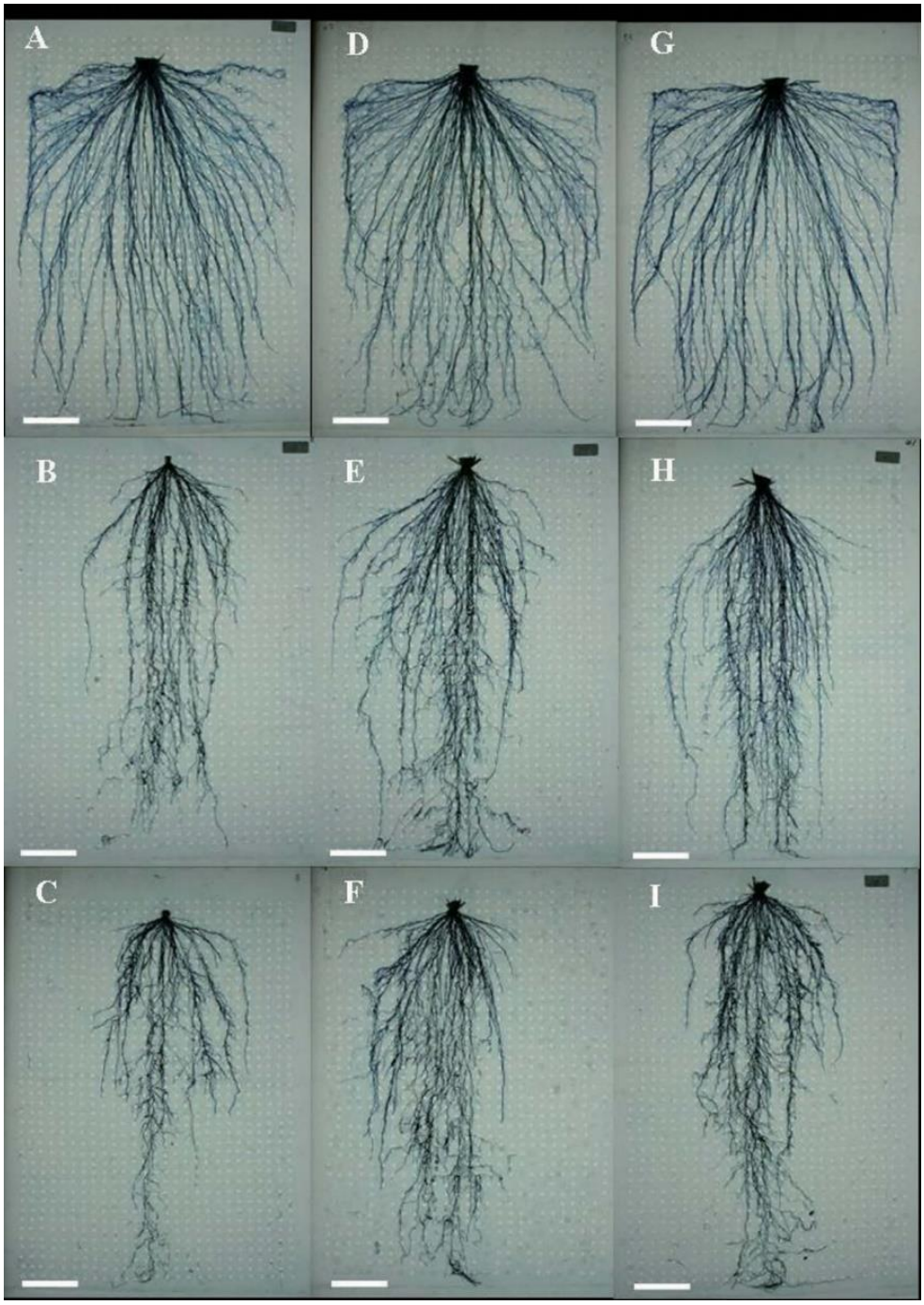

Figure 2.3 An example of how root structure may be captured by using pinboards. These are root systems of rice varieties grown in rhizoboxes with varying degrees of water conditions. ***A,D,G*** *were waterlogged,* ***B,E,F*** *were maintained at 25% GWC, and* ***C,F,I*** *at 20% GWC. Figure is from Kano-Nakata et al. and used under CC-BY 4.0 License* [61]

2.3.3. Rhizotron

Rhizotrons are sub-surface installations that allow continuous observations of roots in the field. In the most simple implementation, a trench is dug and a clear pane is pressed against the soil wall, so roots can be observed as they intersect the pane and continue to grow [64]. More permanent installations may include concrete structures built underground to maintain long term observations and facilitate multiple seasons and studies [65]. Some facilities have been constructed that employ multiple chambers, soils sensors, sampling devices, and weighing scales. These are sometimes called rhizo-lysimiters, and can be quite expensive to build [66, 67]. Rhizotrons permit 2D observations, but some installations include horizontally installed plastic tubes which give extra observation windows. While this type of installation is a powerful tool for root observation, they are not common because of the cost [65].

Overtime the usage of the term 'rhizotron' has shifted to mean any transparent surface applied to a soil monolith – most commonly in large boxes or tubes [68–70] – and are interchangeable with rhizobox. Apparently, the term is still used to refer to underground windows in forestry [71, 72]. Rhizotron boxes and tubes are often left in the field from which the soil is taken, and the soil is carefully packed to the same density, or taken as a monolith.

Boldt-Burisch and Naeth used rhizotrons to investigate the effect of arbuscular mycorrhizal fungi on Bird's-foot Trefoil (*Lotus corniculatus)* and Chee Reed grass (*Calamagrostis epigeios)* root growth [73]. Some treatments were inoculated with fungi and others weren't, but all included clay fragments. The researchers examined the

general growth of roots, but also compared the amount of roots in the clay fragments. They found that treatments with mychorrhizae not only had greater root mass overall, but significantly more root mass in the clay. This led them to conclude that mycorrhizal symbiosis increased root spatial distribution and led to better access to heterogeneously distributed nutrients. Rasmussen et al. used 4 m tall rhizotrons to study the deep roots of Chicory (*Cichorium intybus* L.) [74]. Combining drought treatments and the deep placement of deuterium labeled water, they observed no increase in deep water uptake during water stress. They concluded that the deep roots of Chicory allow a competitive advantage over shallow rooted plants, but were not adapted to compensate for reduced shallow water uptake under drought conditions. Montagnoli et al. used an inverted rhizotron (a clear box buried in soil, using digital image capture to maintain a slim profile) to study the effect of biochar amendment on grapevine root growth [75]. They were able to observe an earlier flush of roots in the spring, and reduced fibrous roots in the summer in amended plots. They concluded that the biochar improved nutrient availability and water holding capacity, thus reducing the need of the plants to forage for water in the summer.

Figure 2.4 An example of a small rhizotron installation. Mohamed et al compared the efficacy of 4 different image acquisition platforms in rhizotrons. ***a,*** *flatbed scanner,* ***b*** *handheld scanner,* ***c*** *smartphone scanner,* ***d*** *time lapse camera. Figure is from Mohamed et al. and used under CC-BY 4.0 License* [72].

2.3.4. Minirhizotron

A Minirhizotron is a clear tube inserted into the soil near a plant, allowing root observations with minimal interference. Taylor et al. report the use of minirhizotrons as early as 1937, and they continue to be used in the present day [65, 76]. While significant drawbacks have been described [65, 77], the method is very popular due to its low cost and ease of use [64]. Early studies employed mirrors and flashlights to observe roots contacting the tube, followed later by the use of fiberscopes and cameras [65, 78].

Several recent studies have focused on the improvement of techniques for both imaging and image processing/root analysis. In hardware, advances have been made using hyperspectral or high-resolution cameras, as well as in automation of the capture process [79–81]. In image processing, the focus centers around automated means of identifying and analyze analyzing roots, often with machine learning techniques [82–84].

As noted above, significant drawbacks exist to using minirhizotrons, though solutions have been proposed. Bragg et al. noted that vertical placement of tubes causes preferential water paths and root growth along the tube, and recommended that tubes should be inserted at a 45° angle to the horizon [78]. Taylor et al. cautioned that minirhizotrons are prone to under-sampling errors, and compared the volume of soil observed with other methods: based on a 5 cm diameter, soil cores explore more than 6X the soil volume [64, 77]. Regardless, minirhizotrons are considered practical means to estimate root branching, longevity, diameter, spatial distribution, length, depth, and rate of growth and decay [64, 78, 85, 86].

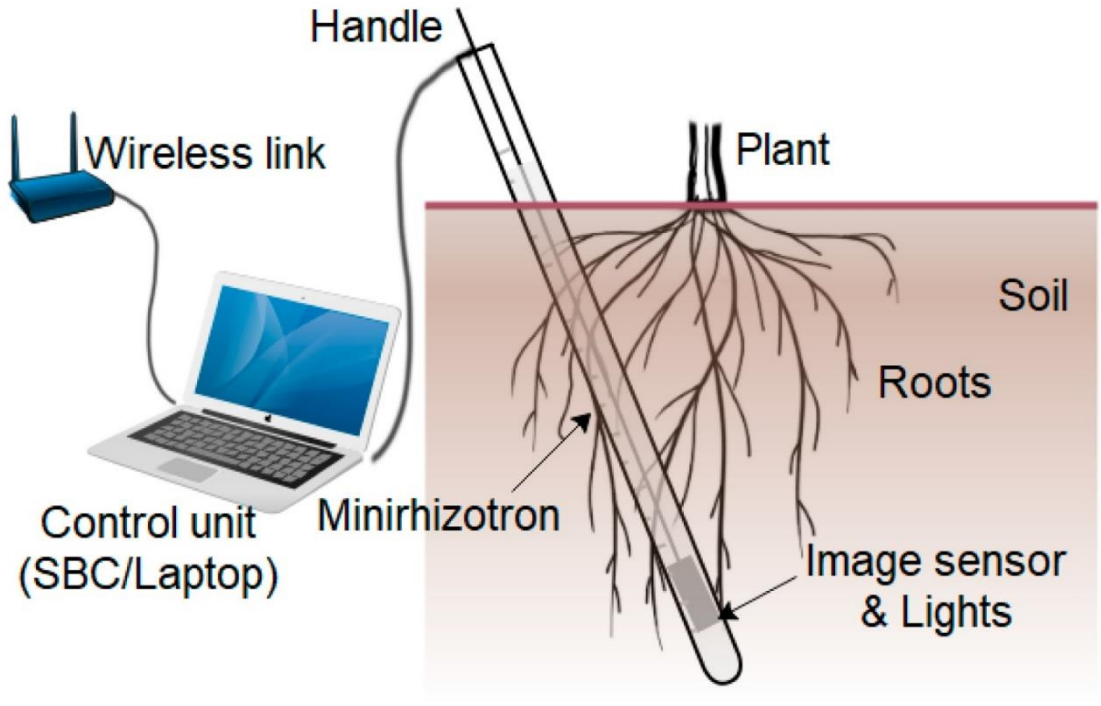

Figure 2.5 The basic components of a minirhizotron – A clear tube inserted into the soil near the subject plant, light source, and imaging device. Figure is from Rahman et al. and used under CC-BY 4.0 License [80].

In 2017, Herrera et al. used a combination of rhizoboxes and minirhizotrons to study the relationship between cover crops and spring wheat in relation to root decomposition [87]. Continuous observation over 3 years of rotations concluded that cover crop species had no effect on the decomposition of spring wheat root, while higher decomposition was associated with high root production and carbon/nitrogen ratio. Haarhoff et al. observed the roots of maize planted at different densities up to 40,000 plants ha^{-1} for two seasons [76]. While they observed non-significant higher shallow root density for lower planting densities, they were able to positively correlate accumulated root length density with grain yield, and concluded that planting density had not affected root growth. Postic et al. used a semi-automated linear-scan minirhizotron to study the relationship between

root depth and yield in wheat [88]. They included soil cores, which were used to calibrate the minirhizotron data from root length surface density to root biomass [64]. Observing 4 varieties in irrigated and rain fed plots, they were able to significantly correlate root length surface density with grain yield, regardless of irrigation treatment. They also reported an association between deep root mass and grain yield in rain fed plots, but not irrigated plots.

2.3.5. Core Sampling

Core sampling is a standard technique and is frequently used as a reference when comparing other methods [49]. Performed by using an auger or sharpened tube, it can be done by hand or by motor: The popular Giddings probe is an example of a vehicle mounted hydraulic sampler [60]. Soil cores are returned to the laboratory where they are washed over mesh screens to capture the roots, which can be measured or imaged. Core sampling can be hampered by stony soil or an abundance of woody roots. Do Rossario et al. recommend core diameters of at least 7 cm, and that cores be stratified every 10 cm of depth [49]. It may also be desirable to perform preliminary sampling to determine the optimal placement for higher density sampling to confidently capture the whole range of variation [49, 89]. Common metrics derived from soil cores include root length density, root mass, distribution, and root diameter.

Fiorini et al. used soil coring to study the effect of no-till management on the depth of roots and soil carbon [90]. They converted a previously conventional till field to no-till for 3 years, and compared the results to a neighboring tilled field. They found that no-till increased the root length and root carbon in the top 5 cm, but not the deeper layers,

which is somewhat contrary to other studies [91]. Muhandiram et al. investigated whether improved forage grass varieties provided greater resilience to soil compaction [92]. Comparing three *Festulolium* hybrid varieties with perennial ryegrass and tall fescue, they compacted field portions after establishing baseline data. Forage yield was initially impacted for all varieties, but improved over time in the hybrids. Using soil cores to measure root traits, they found the hybrid varieties produced higher root mass under compacted conditions. The authors concluded that improved root structure provides resilience to machinery-derived soil compaction.

2.3.6. Core-break

The core-break method is essentially an observation on a plane, like the rhizotron and trench methods. A soil core collected in the field, as described above, is broken in segments generally 5 or 10 cm long, and the number of root segments visible on each side of the break are counted [93]. By washing a subsample of cores to measure root length, the root counts can be converted to root length density, as can be done in the trench method [48, 94]. However, researchers have warned that the relationship between root count and root length density is crop, plant age, and soil specific, so the calibrations cannot be widely applied [95]. Van Noordwijk et al. noted that while core-break counts result in greater variance than washing cores, they take significantly less labor to perform, and therefore more samples can and should be taken to compensate [48]. Indeed, efficiency of sampling time may be the most attractive feature of this method. Wasson et al. studied the problem of variability in the core-break method, and tested a solution using Bayesian hierarchical nonlinear mixed modeling [96]. Using their model,

they were able to distinguish wheat genotypes by their rooting density, and were also able to calculate heritability of root traits with high confidence. Li et al. observed 1,184 soil cores that were 4.2 cm in diameter and 1.8 m deep, in a study of three unrelated wheat populations [97]. With more than 600 total genotypes, the researchers monitored canopy temperature using aerial infrared imaging. Then, working in just the warmest and coolest plot canopies (population tails), they collected 4 soil cores per plot about 2 months after harvest. Plants with the coolest canopies had significantly deeper roots for two populations, ranging from 6.2 cm in one population to 8.1 cm deeper in the other. The authors concluded that canopy temperature could be used as a proxy for rooting depth in breeding selection under water stress. They also noted the time savings of the core-break method. Bai et al. used core-break to investigate whether seedling root screens performed in the lab would correlate to field data in rain-fed winter wheat [98]. Seedlings were first screened on germination paper for root angle and thickness, then screened again for the ability to penetrate a wax layer in a sand column – all traits thought to be indicative of deep rooting behavior. Comparing to root counts and rooting depth in the field, they found a positive relationship between root diameter and yield, but no relationship between the seedling screens and root depth. They also noted an apparent limiting relationship between shallow roots and deep roots, where greater numbers of shallow roots were associated with fewer deep roots.

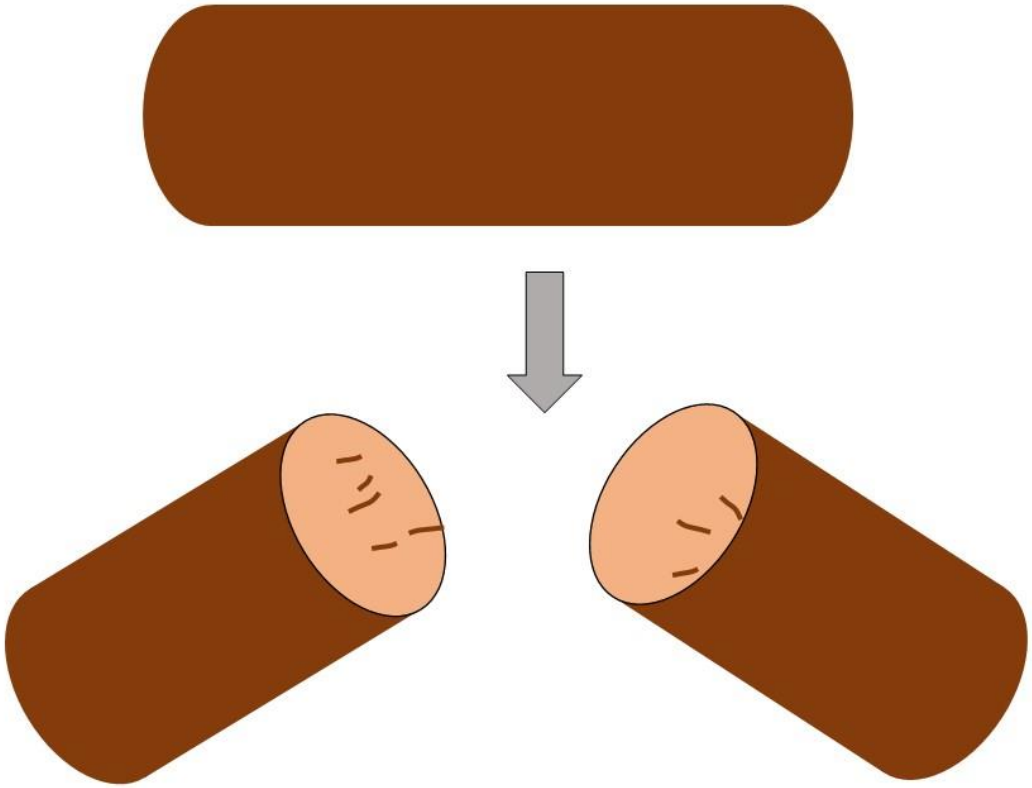

Figure 2.6 Core break method is performed by removing an intact soil core, then breaking it in the approximate middle, or desired depth. The number of visible roots is counted on each side of the break and added together for the total root count at that depth. Figure adapted from unpublished work by Matt Wolfe, 2021.

2.3.7. Ingrowth Cores

Ingrowth cores are an adaptation of the coring method that allows repeated observations at the same position. They are especially good for comparing the relative growth of roots across time, and are often used in longer lived species such as perennials and woody plants. Ingrowth cores consist of a mesh bag inserted into a borehole and filled with root-free soil. Roots are allowed to grow into the core for some time, often for 2-3 weeks, then the cores are removed to capture the change in root growth during that time

[49, 99]. A tube of similar diameter can be used to insert the bag and keep it expanded during filling, and a rope tied to the deep end can be used to withdraw the core. The repeated filling and sampling increases the labor requirements compared to core washing, and while it is important to match field conditions as closely as possible, decreases in soil density and differences in soil moisture were not found to significantly affect the root growth [99]. Ingrowth cores are also better than subsequent coring for estimating belowground net primary productivity in grasslands [100]. In-growth cores are preferable to simple coring primarily when assessing root traits over time.

Reinsch et al. studied changes in soil organic carbon in a forage maize-grassland rotation [101]. Comparing the carbon isotopes before and after conversion from grassland to maize, and using ingrowth cores to measure belowground net primary productivity as a function of root abundance, they questioned whether using a no-till planting system would maintain the high soil carbon of the grassland compared to intense tillage, even though maize root productivity may decrease in no-till. They discovered that not only does the no-till better conserve the soil carbon, but maize root production was not reduced by no-till as compared to conventional tillage. They concluded that using no-till in combination with a 2-or-3-year grassland phase will maintain the high soil organic carbon benefit of grasslands, while also allowing the forage benefit of maize. Lei et al. questioned whether historical land use affected the root production in bioenergy fields, as carbon sequestration is sometimes promoted as a benefit of such crops [102]. Using ingrowth cores, they monitored grasslands and fields converted to maize, switchgrass, and restored prairie for 8 years. They did not find any significant effect of land use

history, but did find all three crops to be inferior to grassland in root productivity. Additionally, they observed maize to have the lowest root mass production, despite it being favored among bioenergy producers.

2.3.8. Shovelomics

Shovelomics is a relatively new method proposed by Trachsel et al. in 2011 [103]. Looking for a solution to high-throughput root phenotyping in the field, they removed the crowns of maize plants with accompanying roots in a 20 cm radius to 25 cm deep, washed the soil away with water, then visually scored the roots. Scoring occurred for 10 traits, including root numbers, angles, and branching. They compared these scores with manual measurements and found a significant positive correlation for all of them, with strong correlation for at least half. They did suggest that angle scoring could be improved by using a protractor as a reference. Reporting an average time of just 2 minutes to score each crown, they recommend this method as a fast and reliable way to phenotype roots in the field. A few years later, Bucksch et al. modified the method by placing the crowns on a black piece of wood and imaging the crown with a digital camera [104]. By creating a groove in the wood to help standardize the placement of the crowns, they report low variability caused by user error, and rapid assessment by automatic computer classification of images. The method was again modified by Colombi et al. by setting up a light tent in the field to control lighting and further aid automatic image classification [105]. Additionally, they split the maize crown in half to expose the interior plane, allowing additional traits to be measured. Their MatLab based software, Root Estimator for Shovelomics Traits (REST), was then able to automatically

compute 33 traits. These adaptations allowed for higher throughput while maintaining precision.

Abiven et al. studied the effect of biochar application to maize fields in Zambia [106]. They found a 45% increase in yield, but also large increases in root traits as measured by shovelomics, in line with results reported by Colombi et al. Root mass was approximately doubled under biochar, and branching and fine roots were increased by 70±56%. They noted that biochar likely affected the root architecture by significantly altering the soil chemistry and physical properties such as density and porosity.

Arifuzzaman et al. used shovelomics to assess 216 rapeseed genotypes for root traits [107]. For this study, they used the method as originally proposed, plus the suggested protractor. Combined with genotypic data, they were able to map 31 QTLs associated with root architecture, including soil level taproot diameter, root angle, and primary root branching. Kengkanna et al. report that cassava has a unique root system because it is planted from stem cuttings, and that little has been done to develop effective root phenotyping methods for this crop [108]. Using the open source and collaborative DIRT (Digital Imaging of Root Traits), they applied shovelomics and measured additional traits specific to cassava, such as the number of basal roots and nodal roots. They were able to observe both genotypic differences in root traits and response to drought and recommend that cassava breeders use shovelomics to define drought-adapted root traits and make selections.

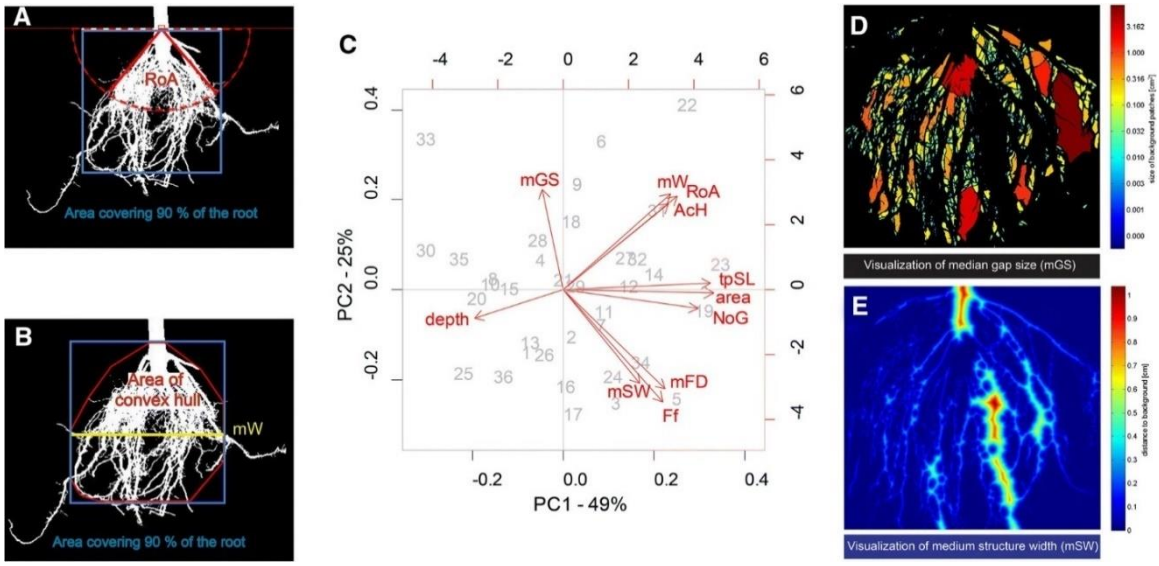

Figure 2.7 An example of the image analysis possible using shovelomics with the REST software as published by Colombi et al [105].***A & B*** *Segmented binary image of a maize crown. The root opening angle (RoA), maximum width (mW), and the area of the convex hull have been calculated.* ***C*** *Biplot of th principal component analysis for several root traits, please see original publication for a full description.* ***D & E*** *Color coded display of the median gap size between roots and the medium structure width or distance from the root structure to the background. Figure is from Le Marié et al. used under CC-BY 4.0 License* [109].

2.4. Non-Invasive Methods

The methods reviewed thus far have all been invasive techniques, even if minimally. In this last section we consider non-invasive root measurements – observation methods which do not disturb the growing plant. While some of them have been used for some time, others are relatively new, and each could be considered experimental. Improvement of these methods is ongoing; some were dismissed in the past but have seen renewed interest due to advances in technology, while others have been recently developed. Several of them require knowledge outside the general scope of plant science; therefore, a comprehensive description of the method is not suited to this work but can instead be found in the references.

2.4.1. Capacitance

Capacitance is a measure of the electric charge stored between two surfaces separated by a dielectric insulator. The insulator permits an electric field to pass, but not a current, causing a buildup of charge when one side has a potential (voltage) applied. Chloupek first proposed that a root system could be estimated by measuring the capacitance of the root-soil interface in 1972 [110]. Extending the work, he used an electrode attached to the base of the plant stem, and another inserted into the soil 25 cm away, to measure the capacitance of potato, clover, sunflower, and mustard plants [111]. The correlation coefficient between capacitance and fresh root mass was 0.7. Kendall et al. attempted to verify Chloupek's work in 1982 by measuring the root capacitance of hydroponic red clover grown in environmental chambers, as well as alfalfa grown in the field, at different stages of development [112]. Using clamps instead of needles, they found partial validation of previous findings – correlation to root dry mass decreased with the age of plants, and capacitance did not change proportionally when roots were severed, though opposing data have been presented more recently [113]. Dalton investigated the method and illuminated several important aspects [114]. They proposed an electrical model to explain the capacitance of roots, hypothesizing that internal xylem tissue and ionic solution forms a conductor while the root tissue forms a dielectric insulator from the soil or rooting medium, suggesting this could be modeled as a series of parallel resistor-capacitor circuits. This circuit is in series with the electrode-stem and electrode-soil circuits, which are measured jointly with the root circuit, thus explaining some of the error in the method. Dalton measured the capacitance of hydroponic tomato roots

and found a correlation of 0.77 to root dry mass. They also monitored root capacitance over time, finding a decrease in capacitance after 50 days, which they posit is an effect of suberization in the roots. Further study found a large effect of soil moisture and stem electrode placement – increasing soil moisture increases the capacitance measurement, and higher stem placement reduces capacitance.

Beem et al. observed that the correlation was more dependent on soil properties than the genotype, and suggested capacitance is a relative measurement unless a calibration is made [115]. In their study, they were able to rank maize genotypes by root capacitance and found it matched ranking by root mass. Psarras et al. attempted to use root capacitance for studying apple seedlings under drought [116]. They calculated an R^2 value of 0.73 in the final harvest but had low significance in previous measurements, which they attribute to the low water content of the soil. Ozier-Lafontaine and Bajazet measured the complex impedance of the stem-root-soil-electrode system over many frequencies [117]. This allowed them to decompose the measurements by apparent component capacitance, resulting in a numerical model with up to 0.97 explained variation. Ellis et al. also measured complex impedance using a 4-electrode method [118]. They were able to determine the stem and soil were primarily resistive circuits, and therefore the bulk of capacitance must be caused by the roots. They also found 4-probe measurements to be comparable to 2-probe when the soil is wet.

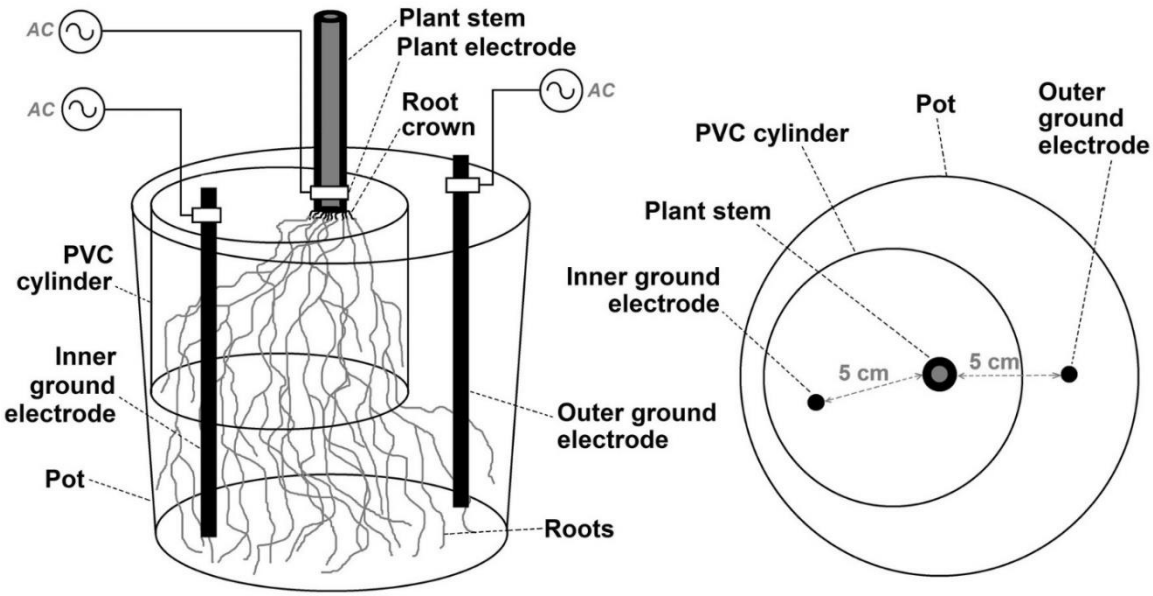

Figure 2.8 A conceptual diagram of root capacitance measurement in a pot. Electrodes are placed at the base of the stem and in the soil. The split-pot design was used here to test whether current was flowing primarily through the soil solution or the roots. Figure is from Cseresnyés et al. and used under CC-BY 4.0 License [113].

Several studies have confirmed the relationship between root system size and capacitance, while a few present dissenting evidence [118–124]. However, authors have successfully used the capacitance method to map QTLs, evaluate drought response, and even observe root association with arbuscular mycorrhizal fungi [125–127]

While it is well understood that root capacitance should be used as a relative measure within fields, it remains possible that correction factors and calibrations could allow absolute measurements [128]. Users need to be considerate of electrode placement, soil moisture, and equipment frequency and voltage. We wish to note the link between capacitance and impedance because some authors use them somewhat interchangeably. Impedance is the resistance to current flow in an alternating current circuit. Because capacitors resist changes in voltage, they can add to the impedance of a circuit. By

measuring the complex impedance, it is possible to derive capacitance for circuits without inductance, such as the plant-root-soil circuit. Additionally, both can usually be measured by the same piece of equipment and electrical connections.

2.4.2. Electrical Resistivity Tomography

Electrical resistivity tomography (ERT) is a common geophysical method for studying near surface features. Resistivity is the electrical property of materials which resists the flow of current. It can be measured by injecting a current with two electrodes and sensing the charge potential between two other (nearby) electrodes. ERT is performed by placing several electrodes in the ground, either in a line or in a grid, then current injection and potential measurements are made at many combinations and levels of spacing. As electrodes are spread further apart, the current moves deeper in the soil, which enables detection of the depth parameter. These measurements are combined using tomography, producing either a 2D or 3D map of the subsurface resistivity depending on the dimensionality of the electrode array [21]. Because roots affect the bulk resistivity of soil, it is possible to detect them using ERT; however, publications are limited [129, 130].

While earlier authors have noted the possibility of mapping coarse roots in trees, Amato et al. appear to be the first to attempt the quantification of tree roots [131]. Using a 48 electrode array, they captured 2-dimensional tomographs under a group of alder trees. The 0.25m spacing and number of electrodes resulted in a tomographic image with 1,264 cells (like pixels) along a 12 m transect to a depth of 1 m. Soil cores of 7.5 cm diameter were taken along the same transect at 0.1 m intervals, washed, and analyzed to map root

mass per unit volume of dry soil – root mass density. The dataset was divided, using one portion to fit a regression model and the other portion to test the regression model for prediction. Coefficients of determination (R^2) were 0.93 and 0.97 between cell resistance and root mass density. In addition, this method also provided a qualitative description of root location. Despite these strong correlations, the authors caution the relationship may not hold for low root density, and that soil texture, moisture, temperature, and the presence of other confounding factors such as stones, must be considered. Later, Amato et al. extended their work by using ERT in alfalfa [132]. Significantly, this seems to be the first attempt at mapping or quantifying herbaceous roots with ERT. Using two soil types in three rhizo-boxes, they planted two boxes with alfalfa and left the third unplanted as a control. Bore-hole electrodes (electrodes placed on a rod and inserted into holes in the soil) were employed. Four arrays were used, each with 18 electrodes (total 72) spaced 1 cm apart, inserted to a depth of 19 cm. This high-density electrode placement resulted in 3096 resistivity values in a cube 20 cm to a side. Soil cores of 2.5 cm diameter were collected to a depth of 20 cm, and in a grid of 5 cm spacing. The cores were washed and roots were collected and measured to calculate root length density and root mass density. Correlating paired points of the resistivity map and the soil cores, the authors found significant correlations of R^2=0.57 and 0.37 for root mass density and root length density. Therefore, they concluded that 3D ERT is useful in investigating root mass, but again caution that low root densities may be confounded by other soil properties such as soil water and texture.

More recently, Corona-Lopez et al. used electrical impedance tomography (EIT) to differentiate between healthy and club-root infected rapeseed plants [133]. As noted previously, impedance is the resistance in an alternating current circuit and therefore is similar to ERT, though excitation frequency and capacitive components also become a consideration in EIT. The authors built cylindrical pots of 18 cm diameter by 13 cm deep, with two rings of 16 integrated electrodes evenly spaced around the circumference. The pots were filled with compost, some being planted and inoculated with *Plasmodiophora brassicae* while others were kept as uninfected control and unplanted control. Differences in root volume and density were detectable and significant, leading them to conclude that EIT could be used in other studies, including in much larger pots as the method is easily scaled. Rather than trying to estimate root mass, Rao et al. used a high density 2D electrode array to monitor soil water depletion in a drought study [134]. They used 123 electrodes placed on the soil surface and in bore-holes along a 7.2 m transect, resulting in 3,084 electrode combinations. They were able to distinguish differences in the depth and width of water depletion between the three cover crops plus one crop mix. Notably, the differences were more pronounced and easier to distinguish under drought conditions. The authors conclude that ERT is appropriate for estimating water depletion by roots, but that it should be measured over time to show changes in resistivity instead of trying to use absolute resistivity.

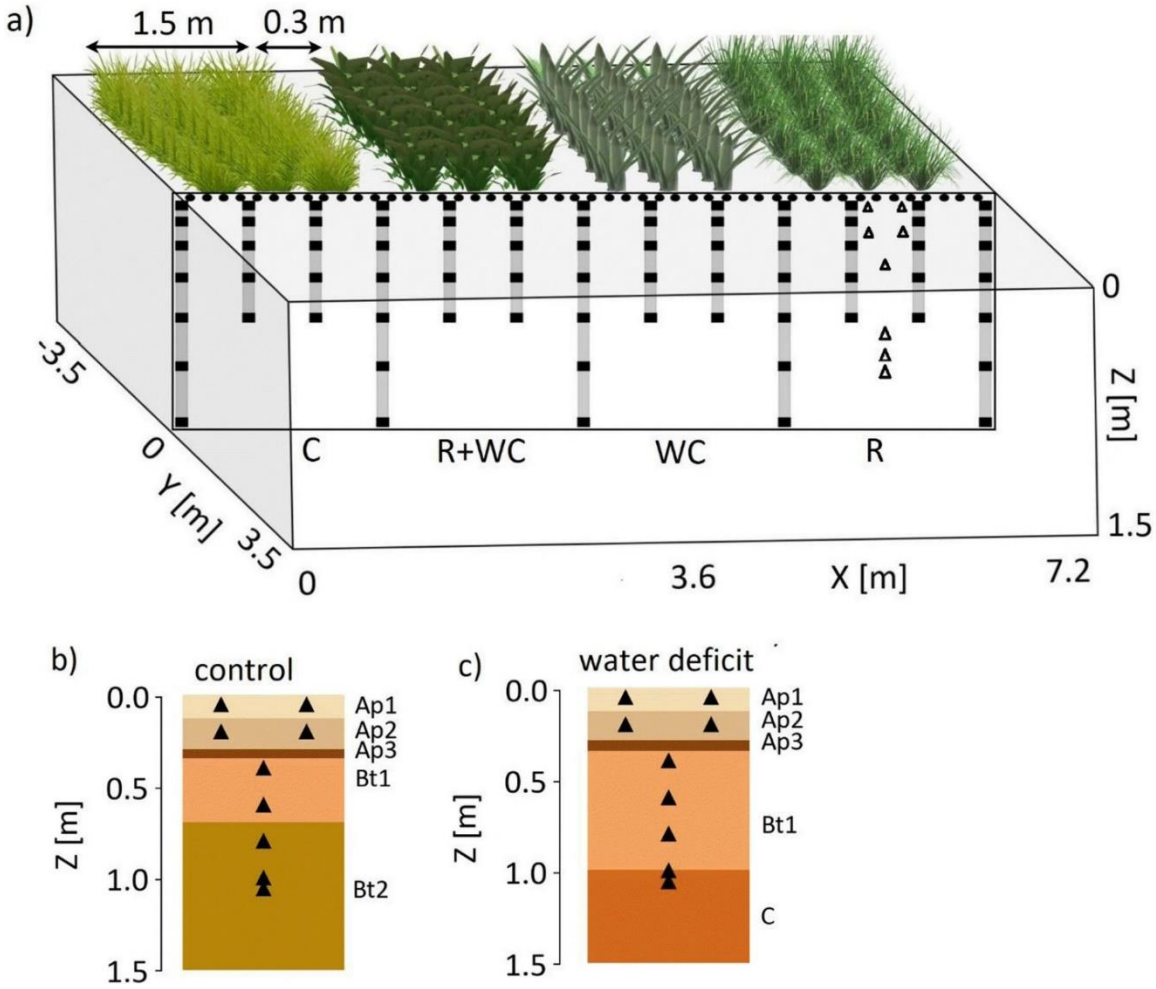

Figure 2.9 ***a)*** *A schematic view of an individual field block showing the placement of ERT electrodes used for 2D imaging. Black dots represent surface electrodes while black squares represent borehole ring electrodes.* ***c & d)*** *Soil horizons and placement of TDR probes (black triangles) for water content measurement in the control and water deficit plots. Figure is from Rao et al. and used under CC-BY 4.0 License* [134].

2.4.3. X-Ray Tomography

X-ray imaging is familiar to most people from its use in medicine. Most commonly, this is done in a single plane, producing a 2D image of internal structures. Less commonly, the imaging plane is rotated several times, or the subject is rotated relative to the imaging plane. Using tomographic methods of back-projection, a 3D model is produced,

referred to as a CT scan (X-ray Computed Tomography). X-rays are photons, and the image is a map of the attenuation (or absorption by dense material), thus X-ray imaging is limited by attenuation – subjects that are too thick block out all the light and prevent an image capture. This is perhaps one of the main limitations of X-ray tomography for studying plant roots as soil quickly attenuates the photons, making sample diameter and image resolution two of the primary considerations. Nevertheless, several scientists have demonstrated the use of X-ray for imaging and analyzing root systems, the greatest advantage of which may be the entirely non-destructive nature and potential for temporal repetition.

Heeraman et al. reported one of the earliest uses of X-ray tomography in imaging plant roots in 1997 [135]. Repurposing an industrial scanner, they scanned 14-day-old bean plants grown in 5 cm diameter tubes. The scans took 4.1 hours per pot and resulted in a 3D image 5 cm diameter by 0.8 cm deep with a resolution of 0.16 mm per voxel. Compared to destructive sampling, the X-ray tended to overestimate total root length. Perret et al. later used a large volume scanner to image 21-day-old chickpea roots [136]. One goal of the study was to increase the sample volume, and they were able to image a cylinder 14 cm in diameter by 23 cm tall at a voxel resolution of 0.275 mm. They further developed software to detect and analyze the roots for number of laterals, volume, length, and others. Their method tended to underestimate root length compared to root washing, which they attributed to the scan resolution. In their 2012 review of X-ray CT, Mooney et al. noted that continually improving X-ray technology was achieving resolutions as small as 0.5 μm, or volumes up to 30 cm in diameter, while also

reiterating the tradeoff between resolution and sample volume [137]. They also found that advances in computing power have led to greater accuracy of root assessment with improved root tracing algorithms. Indeed, Mairhofer et al. presented their root tracing software RooTrak in 2013, followed by Pfeifer et al. in 2015, who repurposed commercial software and claimed a 50% reduction in computing time. Gao et al. introduced their software Rootine in 2019 [138–140]. Rootine is adapted from methods used in medicine to trace blood vessels, and is a shape-based segmentation approach. The authors claim a further 50% reduction in computing time, to just 3 min 40 s for a 500 x 500 x 500 voxel model. Additionally, they report more complete detection, with only 1% underestimation of root length as compared to root washing.

Koebernick et al. used X-ray to study the influence of root hairs on soil porosity in 2017 [141]. Growing barley mutants with and without root hairs, they imaged volumes of 0.42 cm diameter by 8 cm tall, with a resolution of just 0.005 mm. With such fine resolution, they were able to see an increase in pore space directly adjacent to the root hairs, as well as the resultant decrease in pore space 1 mm away. Additionally, they were able to estimate the pore size distribution and connectivity. Later, Helliwell et al. reported similar results for pea, tomato, and wheat [142]. In 2019, Kirk et al. used X-ray CT in flooded rice pots, along with soil gas analysis, to study the exchange of CO_2 produced by root respiration [143]. Their 0.04mm resolution was fine enough to map aerenchyma in the roots and gas bubbles in the soil. Building models from their observations, they showed that rice roots are capable of venting CO_2 at much greater rates than previously thought.

2.4.4. Magnetic Resonance Imaging

Like X-ray scans, magnetic resonance imaging (MRI) is familiar from its common use in medicine. MRI is based on the same principal as nuclear magnetic resonance (NMR), but was rebranded to avoid negative associations with the word “nuclear” when it was adapted for medical imaging [144]. MRI is performed by creating a strong magnetic field around the subject, which causes the alignment of paramagnetic atoms within the field, including hydrogen atoms, which are abundant in all living things. Within the strong magnetic field, secondary magnets create weaker directional field pulses which ‘tip’ the spin of the atoms for a short time, causing the atoms to emit a field that can be detected by radio-frequency receiver coils. Through the localization of the pulsed fields and array of receiver coils, the origin of emitted energy can be calculated and an image produced. The length of time each atom emits energy is called the relaxation time, and is affected by neighboring atoms and molecules, allowing for the differentiation between materials in the image [145]. Using multiple detector arrays results in a 3D image. In 2015, Metzner et al. compared root imaging by MRI and X-ray CT and concluded that MRI was more suitable for large pots, as it doesn’t suffer from the attenuation problem [146]. They also found better agreement to root washing for root length because it was easier to discriminate roots in MRI than X-ray, though the previously discussed advances in automated segmenting may change that result now.

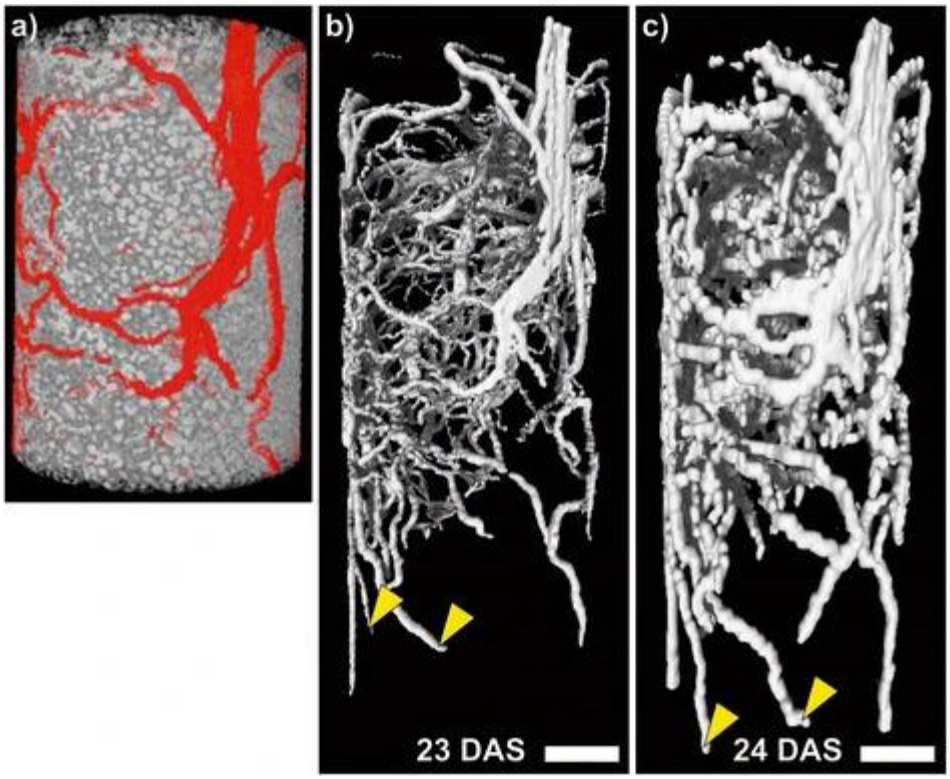

Figure 2.10 A bean plant grown in a 34 mm diameter by 200 mm tall pot and imaged by X-ray CT and MRI at 23 and 24 days after sowing. ***A)*** *The upper portion of the soil column, imaged by CT with a voxel size of 0.028 x 0.028 x 0.028 mm^3, the roots highlited in red.* ***B)*** *The segmented root system of the same plant imaged by CT with voxel size 0.056 x 0.056 x 0.056 mm^3.* ***C)*** *The same subset as* ***(b)*** *imaged one day later by MRI with a voxel size of 0.333 x 0.333 x 1.000 mm^3, the larger dimension being vertical. The scale bar is 10mm. Figure is from Metzner et al. and used under CC-BY 4.0 License* [146].

The use of MRI in plants is nearly as old as the technique, with Rogers et al. publishing one of the earliest studies in 1987 [147]. Recognizing the potential of ferromagnetic particles in soil to interfere with imaging, they tested 30 soils and 8 artificial potting media, and imaged the roots of bean plants. They reported significant image degradation when magnet fractions were over 4% by weight, and that non-ferrous natural soils tended to perform better than artificial substrates such as peat, even though the artificial media tended to have very low magnetic fraction. Rather significantly, they discovered that soil water was invisible to their scans because of the low relaxation time of bound

water, a phenomenon which was not repeated in the artificial media. Recently, Bagnall et al. demonstrated the potential of a low-field MRI device, which produces a primary magnetic field at a small fraction of the strength of standard MRI facilities [148]. They tested natural agricultural soils and were able to image plant roots in soil with >10% clay content, with magnetic fraction >4%, at resolutions of 0.63mm. While the capture times were extended because of the low power, and hence weaker response, the study represents an important increase in MRI capabilities for root studies.

A 2016 study by van Dusschoten et al. using MRI found significant correlation of root mass estimation compared to root washing, which improved to 90% when small roots below the detection threshold were omitted [149]. They achieved a resolution of 0.73 mm in an 11 cm pot and were able to estimate fresh root mass based on signal intensity, as well as root length, number of root tips, number of lateral roots, root angle, and root distribution, among others. They also imaged the rhizobium nodules of a young bean plant (32 days). They noted the possibility of studying the root systems of tuberous or storage root plants because of the large pot size. In 2020, Perelman et al. demonstrated the ability to map sodium ions (Na^+) in soil and around roots [150]. They grew tomatoes in pots with saline soils, and imaged the roots and Na^+ concentrations after 6 weeks. They found an association between high transpiration in the tomato plant with greater accumulation of Na^+ in the taproot and distal Na^+ depletion in the surrounding soil. The low transpiration treatment showed lower Na^+ movement and less accumulation.

2.4.5. Ground Penetrating Radar

Ground penetrating radar (GPR) is a geophysical method for studying subsurface features. It is commonly used in geoscience, civil engineering, architecture, and archeology [20]. GPR works by emitting a pulse of electromagnetic energy (EM) into the ground, where it is either transmitted, absorbed, scattered, or reflected. Reflection is caused by changes in dielectric permittivity. In agricultural soils, the primary drivers of dielectric are soil texture and water content, with water content having the greater influence [21]. In general, GPR systems consist of system electronics which generate and process the signal, a transmitting and receiving antenna or antenna array, and a computer-based control and capture software [18]. Each EM pulse returns a 1D signal trace called an A-scan. As the transect is measured, A-scans are stacked together to create a 2D image called a B-scan. If multiple lines are scanned side-by-side, or an array is used, the B-scans may be stacked into a 3D image, or C-scan [20]. The resolution depends on the ability of the system to distinguish closely spaced signal peaks, and will be determined by the wavelength and the sampling rate of the system, shorter wavelengths and faster sampling rate giving finer resolution. The smallest detectable object, or detection threshold, is determined by the signal frequency – the rule of thumb being ¼ wavelength. However, a tradeoff exists between frequency and attenuation, with higher frequencies being more quickly attenuated so that deep observations are not possible [21]. Unlike CT and MRI data, GPR data are in the time domain, and the images are not representative of spatial relations, making GPR data difficult to interpret visually. In many cases, GPR data are used qualitatively for the location of objects rather

than for quantitatively measuring objects [19].

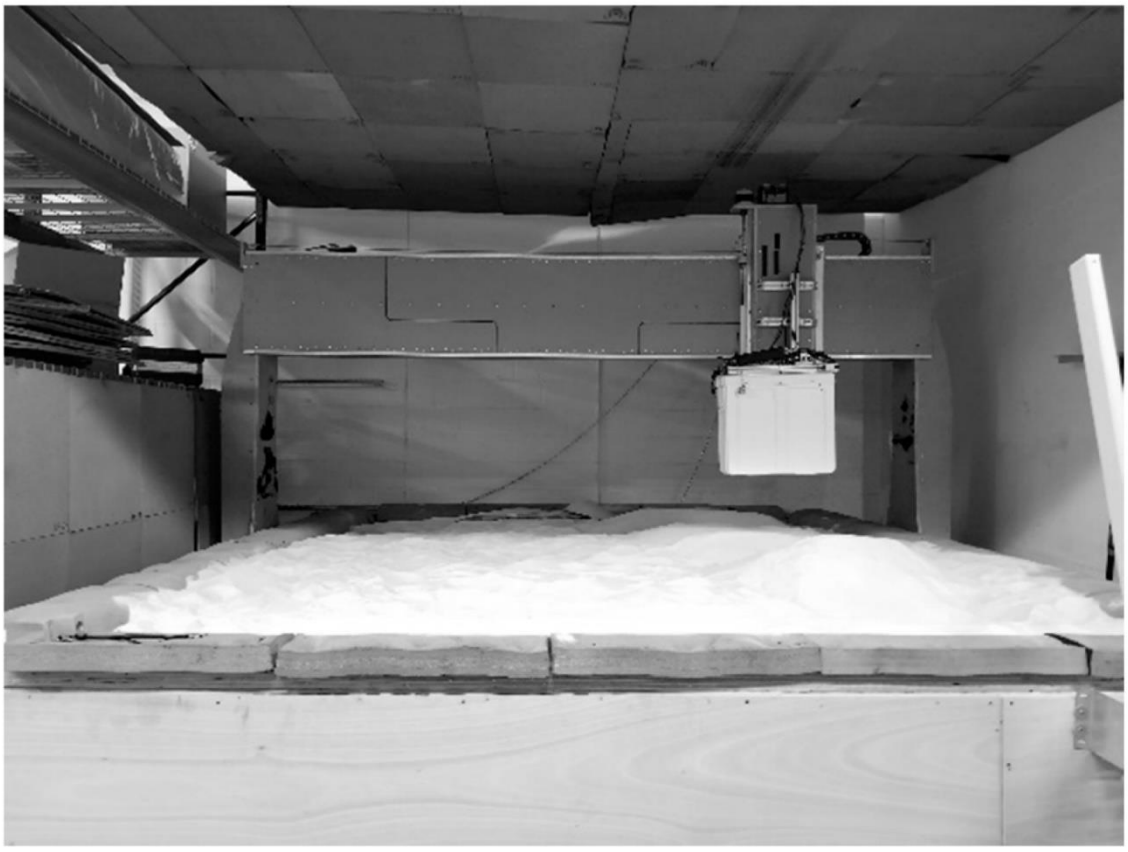

Figure 2.11 An example of an air-launched (not contacting the ground) GPR antenna. This one is attached to a computer controlled gantry over a sandbox and used in controlled experimentation to develop methods of root measurement. Figure is from Delgado et al. and used under CC-BY 4.0 License [151].

The use of GPR for detecting plant roots is recent, with Butnor et al. being one of the earliest in 2001 [152]. Their research possibly launched the use of GPR for locating and quantifying tree roots, with at least 117 citations at the time of this review. In this study, a 1.5 GHz antenna was used to detect tree roots as small as 0.5 cm. They correlated GPR reflections to excavated root mass and found correlation coefficients ranging from 0.49 to 0.55. The use of GPR in agriculture and row crops is significantly newer, with only 5

apparent publications by the end of 2020. The first peer-reviewed article was published in 2008 by Konstantinovic et al. [153]. They used an ultra-wideband antenna to detect and estimate the mass of sugar beets in the field. Mounting horn antennas to field machinery, they scanned rows of matured sugar beets. Using amplitude threshold methods, they were able to detect better than 90% of the beets. Then, using the detected position of the beats, they measured the amount of reflected energy at that point, and correlated it to fresh root mass, measuring correlation coefficients from 0.60 to 0.70 (discreet numbers not reported). Curiously, it seems there was no follow up to this research, and the next published article is nearly a decade later, in 2017. Delgado et al. investigated the use of a small GPR system for detecting the onset of root bulking in cassava [8]. Cassava, being a tall bushy plant, is not suited to field equipment such was used by Konstantinovic et al. Therefore, Delgado et al. used a ground coupled system, which they passed in a tight grid on all sides of the plants, one plant at a time, for a total of 60 plants. The cassava was planted at staggered dates, giving a large range of bulked root sizes at the time of observation. Using amplitude threshold methods, they correlated pixel counts to root fresh mass, with R^2 ranging from 0.51 to 0.77 within genotypes and 0.63 across genotypes. Then, comparing predicted mass to actual mass within each age class, they demonstrated the ability to detect the onset of root bulking. Liu et al. used threshold analysis and intensity averaging to estimate root mass in a field crop, though they are the first to report significant results with fine root mass [154]. Working with wheat and energy cane, they found significant correlations, with R^2 ranging from 0.12 to

0.54 at different locations. Contrary to general GPR wisdom, they found better performance in clay soils, possibly because the roots were more concentrated.

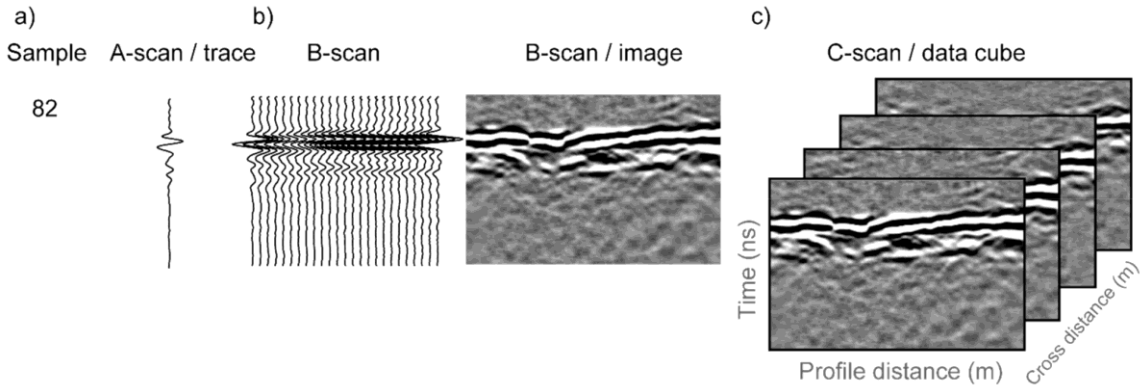

Figure 2.12 An example of GPR data in the time domain. ***A)*** *A single sample or amplitude value from the collected signal trace or A-scan.* ***B)*** *A collection of A-scans along transect create a B-scan, and can be represented as an image, the greyscale representing amplitude values.* ***C)*** *The collection of parallel transects is called a C-scan, and can be interpolated into 3D. In this example, bright and dark bands approximately 1/3 of the way down result from the strong reflection at the air/soil interface. Figure is from Dobreva et al. and used under CC-BY 4.0 License* [155].

2.5. Conclusions and Future Directions

The last few decades have seen an explosion in root research, advances in methods, and challenges to previous assumptions. However, the need still exists for a rapid, affordable method to assess roots in the field, without modifications to their growing environment. Lab methods have great power to measure numerous root traits, from angles to exudates, but are unable to effectively convey the effects of environment and are generally limited to very young plants. Additionally, the challenge remains to connect seedling observations to mature plants in the field. Pot methods can sometimes approach field conditions and offer observations of features such as lateral root development, but are still limited by the amount of space required to set up such experiments in greenhouses

or environmental chambers. Field observations remain the gold standard, and while some new methods have been developed, invasive or destructive observations are still required in the majority of studies. The abundance of field methods suggests researchers are still looking for the ideal solution which balances the need for descriptive results with resource and budget constraints.

Some field methods have been developed or modified to address one of the widespread problem of balancing time and cost with high quality and representative data. Namely, minirhizotrons have become less expensive, and image capture technology, especially automation, has reduced the time required to collect data, allowing for increased samples. Shovelomics has similarly addressed the issue of time and cost, creating a high-throughput method to analyze root crowns. Finally, image capture, recognition, and automated processing techniques have reduced the time requirement of nearly every method, a trend which will no doubt continue. Despite these innovations, significant drawbacks remain, including the intrusive and/or destructive nature of field methods. Non-invasive methods (mainly relying on electrical properties) have shown promise, but tend to be limited by resolution, cost, feature discrimination, or some combination of the above. X-ray imaging was once seemed poised to provide a comprehensive solution to phenotyping roots even in very large plants, with predictions of increasing sample sizes allowing large pots, and reduced cost allowing wider access; however, it seems the technology favored increasingly detailed scans with smaller sample sizes, and while the price has decreased over the decades, X-ray machines are still out of reach for most researchers. MRI has continued to improve resolution without the sample size limitation,

though scan time, throughput, and access are still a significant limitation. ERT, while promising, may not be able to produce the precision many researchers are looking for. The level of expertise required to perform and process ERT studies also places it outside the reach of many researchers. Capacitance and GPR measurements both suffer from the inability to distinguish root characteristics, and at the moment appear only able to estimate root mass, though it's possible this could change with more investment. However, both GPR and capacitance could be considered or adapted to a high-throughput method.

Despite the large number or root observation methods and the recent advances, it's clear that an ideal method for phenotyping roots in field conditions does not yet exist. The next decade will undoubtedly continue to see advances in root studies as new methods are developed to achieve fast, non-destructive sampling of belowground plant material. Researchers can build upon these previous studies to develop new root imaging systems that suit a variety of crops needed to meet rising food demands, as well as contribute to a more sustainable agricultural industry. In the meantime, a large variety of methods exist, each of which is well suited to specific study parameters and research goals. There is no "one size fits all" method for observing plant roots, and each scientist must carefully consider the strengths, weaknesses, and applicability of the current methods when planning their experiments.

3. EFFECT OF SOIL WATER ON GPR ESTIMATION OF BULKED ROOTS, METHODS AND SUGGESTIONS

3.1. Introduction

Recent publications have shown the importance of plant root systems in breeding for resilience and climate change [25–27, 156]. In addition to the importance of structural root systems, many crops are grown for their bulked roots or tubers, such as potato and cassava — two globally important crops. Cassava (*Manihot esculenta*) is a tropical plant grown for its starchy bulked roots. More than 800 million people depend on it as a staple food, and even more rely on it as an important source of starch [9]. It is commonly grown by subsistence farmers because it is hardy and harvest can be delayed until needed [10]. Cassava is the world's fourth most important basic food crop and the global harvest has increased by more than 25% since 1999, nearly doubling in some regions [6, 11]. Varieties commonly require 12 months or more before harvest maturity is reached, making early root maturity a primary factor in variety selection by farmers, and a major goal for breeders [6, 10, 157].

While many advances have been made in methods to study fibrous roots, few have been made to assist breeders of bulked root crops like cassava [27–29, 47]. Yield monitoring depends on point sampling or post-harvest metrics [158–160]. A relatively new method called shovelomics was successfully applied to cassava with some modifications, but it remains a completely destructive method with all associated disadvantages [108]. Cassava breeding is hampered by very low multiplication rates, and early trials may

have only 3 clones to evaluate, making it untenable to destructively sample the roots until it is certain they have bulked [6]. This highlights the continued need for non-destructive methods to assay root crops.

In 2017, Delgado et al. reported on the use of ground penetrating radar (GPR) for estimating cassava root mass and detecting the onset of root bulking [8]. A later 2019 study by Delgado et al. furthered the work by performing high density scans on buried roots in a climate controlled sand box [151]. This study was able to approximate 3D models of the buried roots. GPR has also been used to estimate the root mass of sugar beet, wheat, and peanut [155, 161, 162]. GPR is a geophysical tool for detecting belowground features such as fault lines or buried utilities [21]. GPR works by emitting a pulse of electromagnetic energy into the ground, where it is either transmitted, absorbed, scattered, or reflected. Reflections are caused by changes in dielectric permittivity, a measure of how strongly molecules can be polarized. In agricultural soils, the primary drivers of dielectric are soil texture and water content, with water content having the greater influence by an order of magnitude. Therefore, the high water content of roots has the potential to reflect GPR signals.

In general, GPR systems consist of system electronics which generate and process the signal, a transmitting and receiving antenna or antenna array, and computer-based control and capture software. GPR data are in the time domain, and the images are not representative of spatial relations, making GPR data difficult to interpret visually. This is because the GPR is directly measuring time-of-flight for the signal, the speed of which is controlled by the dielectric of the medium, such that the velocity (V) can be estimated by

the ratio of the speed of light in a vacuum (C) to the square root of the dielectric (ε) (see Eq. 1). In many cases, GPR data are used qualitatively for the location of objects rather than quantitatively [19]. Soil moisture has the effect of attenuating GPR signals; consequently, it is standard practice to prefer dry soil for GPR measurements [20]. However, in an earlier experiment which attempted to estimate fine roots with GPR, Liu et al. suggest that wet soil may improve data quality [154].

$$V = \frac{C}{\sqrt{\varepsilon}}$$

Equation 1 The velocity (V) of an electromagnetic wave is dependent on the dielectric (ε) of the medium in which it travels, and is relative to the speed of light in a vacuum (C).

Delgado et al. (2017) used a small radar system which was carefully passed along the soil surface in a grid around each plant, and then measured the depth of each root to allow a supervised processing method. While this method serves as proof of concept, it is unsuitable for high-volume phenotypic evaluations common in the early-stage testing in plant breeding. GPR systems commonly use antennas which are in contact with the ground, called ground coupled. In agricultural systems, because of the soft, uneven soil, and the likelihood of standing plant mass, an air-launched antenna (antenna not ground coupled) is more appropriate. However, lifting the antenna from the ground has the potential to add observational error to the data. Air-launched antennas tend to suffer from increased ground clutter, which is the bright reflection caused by the interface of the air and soil, and any material at or near that interface, such as plant mass. Additionally, variations in the orientation of the antenna, whether caused by wobbling of the cart or other factors, can introduce variations in the measurement. Lastly, all GPR

systems potentially receive interference from outside sources, such as cell phone signals or other nearby instrumentation. These often can be minimized during the collection by stacking, which is the rapid and automatic collection and averaging of several GPR pulses.

In this paper, we describe a field ready GPR system which is more suited to high volume applications in estimating bulked root mass by using an air launched antenna array. We describe preliminary experiments in controlled field settings using a model root crop, daikon radish, and novel data processing methods for extracting quantifiable data from GPR scans. Importantly, the effect of soil water was explored. We present methods for collection, and suggest some good practices which future researchers should consider. The strength, limitations, and future potential of GPR will discussed.

3.2. Methods

3.2.1. Location

A raised bed of loamy sand was built in the Brazos River Bottoms, near Texas A&M University in College Station, Texas (Figure 3.1). The bed is built of concrete blocks, stacked approximately 2 meters high, and the bed is approximately 3.5 by 22.5 meters long. The soil is sandy loam, transported from nearby farmland, and is kept free of weeds, therefore, the soil is homogenous and without distinct horizons. The soil was broken with a shallow till, leveled, and settled by watering before experiments began. 21 plots were measured out at 80 cm intervals, and marked off with stakes and string, then holes were carefully dug in the center of each plot. The holes were square shaped,

approximately 65 cm on a side, with flat bottoms and measured 15 cm deep from the surface. The number of plots was limited by available space.

Figure 3.1. Study was conducted in large raised beds filled with sandy loam soil. Plots were carefully measured and marked by string.

3.2.2. Root Mass

Daikon radish were purchased from local grocery stores then weighed and labeled individually in random order. Daikon radish were used rather than cassava because of local availability, cassava being unavailable in the required quantities. Daikon radish

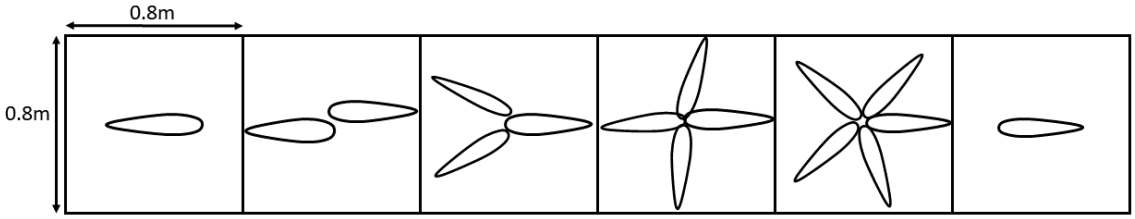

Figure 3.2. Roots were placed horizontally to maximize the angle between adjacent roots, as space allowed. Plots contained between 1 and 5 roots.

were considered an appropriate model root because of similarities in size and shape. The roots were placed in the holes horizontally and arranged to maximize the angle between adjacent roots and mimic the root growth of cassava. The number of roots in each plot was varied between 1 and 5 to increase variation in plot mass and variation in root orientations (Figure 3.2). Enough roots were obtained to fill 19 plots. Per plot root mass ranged from 542 g to 2931 g.

3.2.3. Sensors

Campbell Scientific CS655 soil moisture sensors (Logan, Utah, USA) were placed in the first and last plots, at two depths: 5 cm and 20 cm. The sensors were inserted horizontally into the undisturbed soil on the side of the plots. Moisture levels were recorded before radar scanning. Additionally, 2 flat metal plates were place in the first and second plot at the same depths as the sensors, so that the plates straddled the root zone. These were meant to demarcate the root zone in the GPR data by acting as distinct reflectors.

The radar sensor used was an experimental loaded-vee dipole array, manufactured by IDS Georadar (Pisa, Italy) [163–165]. The array consists of 4 transmitters and 4 receivers in alternating pairs, each spaced 4 cm from adjacent antennas (Figure 3.3). The antennas are wideband with a center frequency of 1.8 GHz. The radar captures 512 samples over 18 ns, and pulses every 1 cm, as measured by an encoder wheel. Channel configurations paired every antenna with its directly adjacent neighbor, giving a total of 7 channels, each offset by 4 cm. The sampling time increases with the number of channels, and in this case prevented automatic stacking as that would result in lost data

due to hardware limitations of sampling speed.

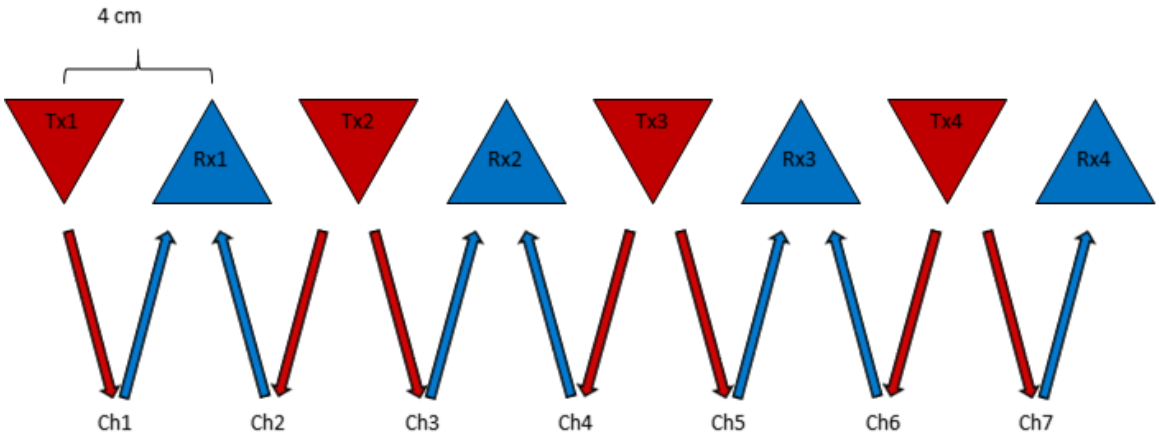

Figure 3.3. The antenna array consisted of 4 transmit and 4 receive antennas, spaced evenly 4 cm apart. 7 scan channels paired each antenna with its neighbors. Tx designates a transmitting antenna, while Rx designates a receiving antenna, Ch indicates a channel pairing.

The array was air launched and mounted on a 4 wheeled cart that straddled the plots and placed the bottom of the antenna enclosure 39 cm from the ground surface (Figure 3.4). The antenna was pointed directly at the ground, and a plastic rod was attached at the center of the enclosure to give a ground indication of the nadir of the radar. The plastic rod indicated the center of the antenna at ground level, and was used as a reference for marking the plots digitally in the capture software.

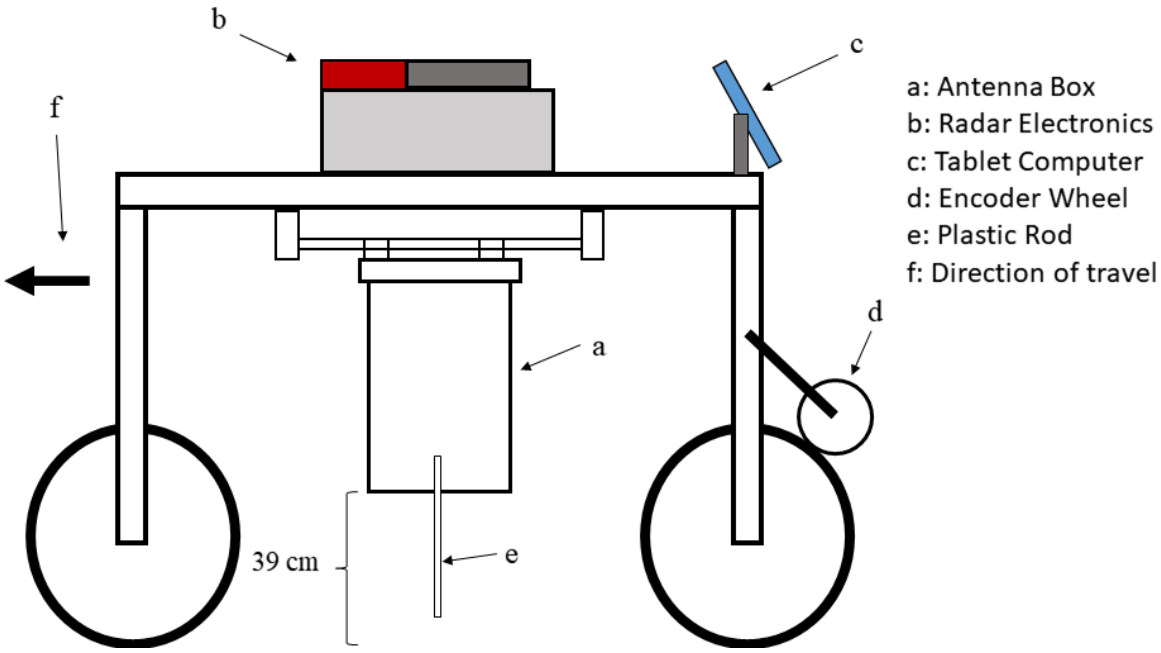

Figure 3.4. The antenna cart had 4 wheels and straddled the plots. The antenna array was hung in the middle with a plastic rod on one side to indicate the position of the center of the array on the ground.

Soil Moisture

The experiment was scanned at 5 different water contents, beginning with a 'dry' state. Two oscillating sprinklers were placed in the field such that the sprays evenly covered the entire expanse without overlap. The sprinklers were measured to provide about 13.5mm per hour by collecting water in two 5-gallon buckets placed in the sprinklers' path. The sprinklers were run for times varying from 2 hours to 8 hours, then the field was allowed to rest between 4 and 48 hours before scanning, allowing surface pools to drain and attempting to let subsurface moisture equalize spatially. Table 3.1 shows the percent volumetric water content (VWC) of each treatment, and the standard deviation across the sensors.

Table 3.1 Average volumetric water content of each treatment, and the standard deviation as percent water, volume/volume basis.

	No Irrigation	Irrigation 1	Irrigation 2	Irrigation 3	Irrigation 4
Average VWC	12.5%	17.3%	16.7%	18.6%	16.7%
Standard Dev.	0.86%	3.68%	3.24%	3.11%	1.7%

3.2.4. Capture

The GPR cart was assembled in the field and given time to equilibrate to ambient temperatures. Before collecting data, several scans were passed across the entire transect to allow the electronics to "warm up". More than 1 meter was allowed between the starting position of the cart and the first plot, and the ending position and the last plot, to ensure the radar captured the entire extent. Using the plastic rod (Figure 3.4 e) and string as an indicator, a digital marker (fiducial) was placed in the data in between each plot. The experiment was scanned 6 times for each treatment.

3.2.5. Data Processing and Analysis

Data was processed using GPR Studio version 1.0 (Crop Phenomics LLC, College Station, TX, USA (cropphenomics.com)), a Python software library developed for the quantitative analysis of GPR data. The software utilizes published data processing libraries and custom-built functions specific to GPR analysis. Analysis was conducted in two stages.

In Stage 1, each scan was separated into plots based on the digital markers placed in the data during field capture. The plots were then filtered to only those containing roots. Data were subset to the approximate root zone, then passed through a Butterworth bandpass filter, removing noise below 0.5 GHz and above 1.05 GHz. The 7 channels

were each standardized to themselves by subtracting the channel mean from each value, and dividing by the channel standard deviation, similar to how a statistical z-score is calculated. This removed offset differences between channels caused by automated signal calibration in the field. Standardized channels were then squared to move all values to the positive domain and minimize the background information which tends to gather about the mean, or 0 in standardized data. Channels were interpolated into a 3D cube using linear interpolation. A horizontal window, or time slice, 5 rows thick was passed from the top of the cube to the bottom, summing the amplitude in each window, which can be considered an indicator of total reflected energy in that window. Several window depths and alternative measurements, such as window variance, were tested – the most effective is reported here. Window values were correlated against observed root mass in each plot, the results filtered for p-value < 0.1, while also controlling for appropriateness in the depth of the window and consistency between observations. The depth with the lowest p-value was stored for further processing. The cumulative energy values for each plot were divided by the length of that plot, in scan columns, resulting in relative energy density. The workflow pipeline is shown in Figure 3.5

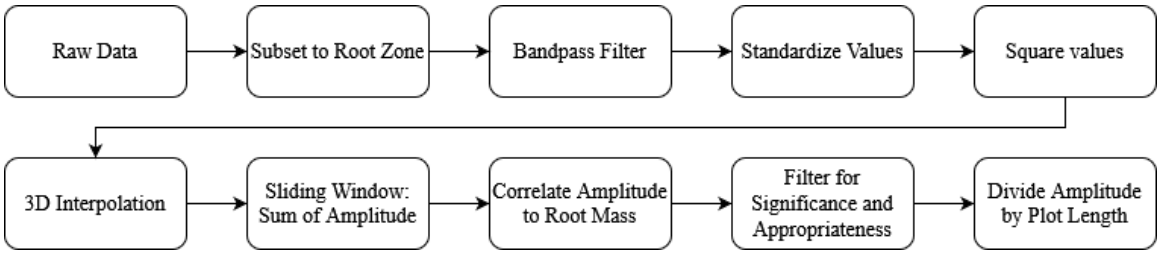

Figure 3.5. Stage 1 of the data processing workflow.

In Stage 2, the results of Stage 1 for repeated observations were averaged together to reduce observational variance. Table 3.2 shows the mean relative energy density per plot with the standard error of the mean (SEM), and the coefficient of variation (CV) for a selection of plots and treatments, demonstrating the observational variance in each repeated scan. Observational variance is an indication of noise in the data, which is a significant source of error here, and is further discussed below. Repeated observations help to remove the error through averaging.

Linear regression analysis was performed between the averaged energy density and the observed root mass.

Table 3.2. Observational variance per plot, the mean relative energy density for some of the plots and treatments, and the coefficient of variation. Plots are arranged in order of increasing root mass to aid interpretation.

	Dry		Irrigation 1		Irrigation 4	
	Mean	C.V.	Mean	C.V.	Mean	C.V.
Plot 1	7.03±0.31	10.7%	6.77±0.87	31.5%	4.71±0.42	19.8%
Plot 4	7.11±0.29	9.8%	8.21±0.86	25.6%	6.44±1.05	36.3%
Plot 8	8.04±0.23	7.1%	6.17±0.88	34.8%	5.85±0.76	29.0%
Plot 12	7.03±0.30	10.3%	7.68±0.79	25.1%	5.90±0.76	28.7%
Plot 16	8.42±0.40	11.5%	8.90±1.6	44.1%	5.90±0.62	23.4%
Plot 19	8.44±0.33	9.6%	8.26±0.97	28.9%	10.22±1.4	30.7%

The metal plates placed above and below the root zone were not identifiable in the radar data, possibly because of the type of paint applied to them. Therefore, it was not possible to use them to fine tune the root zone as had been planned.

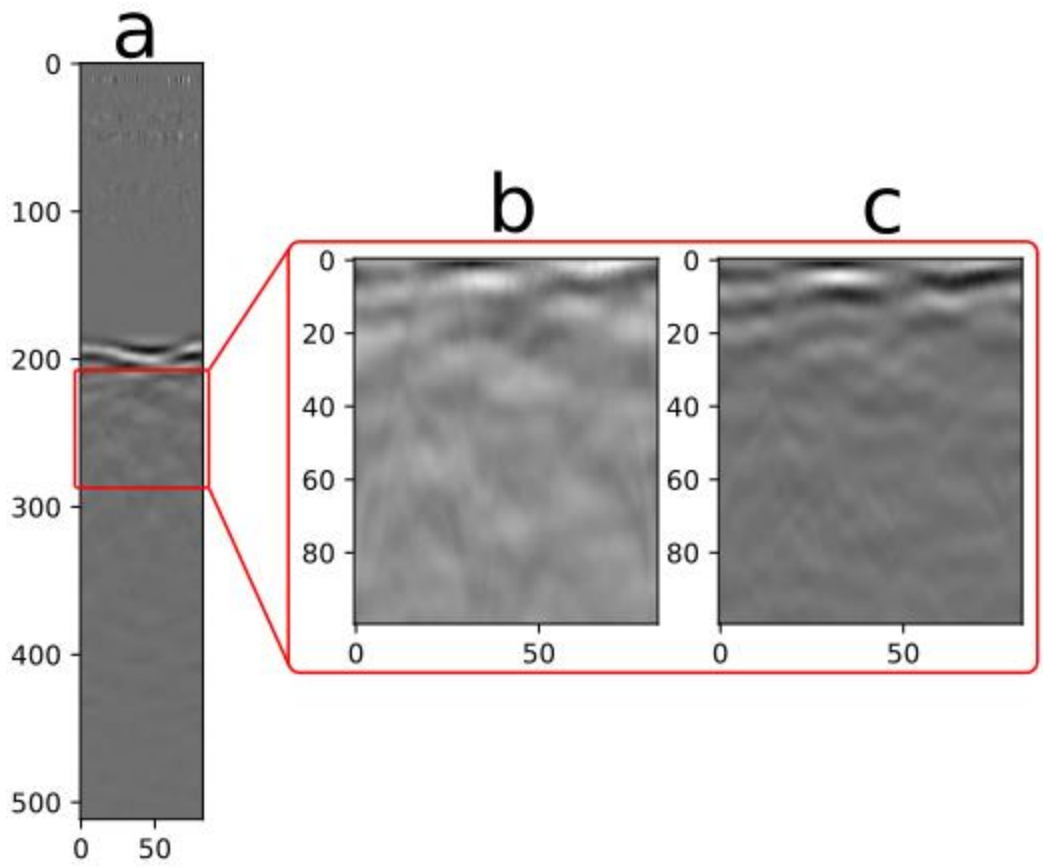

Figure 3.6. Sample radargrams from Stage 1 of the processing workflow. In unprocessed GPR data, both white and black represent high amplitude energy reflections, albeit at opposite phase angles. Low energy is shown as grey. A - Raw data showing a single plot. Notice the surface reflection near line 200. B- Plot cropped to the root zone, corresponds to lines 205 to 305 on left image. C - Standardized plot.

3.3. Results

Our results demonstrate a significant relationship between reflected GPR signal and bulked root biomass. Additionally, the results demonstrate the importance of a homogenous dielectric environment in the soil, independent of water content. Lastly, we show that for relatively shallow agricultural studies, dry soil is not necessarily superior to wet soil for GPR measurement, a finding which supports the hypothesis of Liu et al, 2018 [162].

All treatments were significant at the $p < 0.05$ level after repetition averaging. The final treatment, which had the most homogenous wet soil but not the wettest, showed the strongest correlation with 63% explained variation, followed by the wettest treatment, then the dry treatment, which had the least variation in soil moisture. Larger variation in soil moisture decreased the strength of the correlations (Figure 3.7). Although the number of plots was low, the repeated measurements and multiple treatments reinforces the probability that GPR features are indicative of root mass and are not random. Dividing the sum of energy by the scan length modestly improved the correlations by adjusting for the effect of scan length, indicating that scan length was not a significant contributor to the correlations, but that variations in scan length introduced a small amount of error.

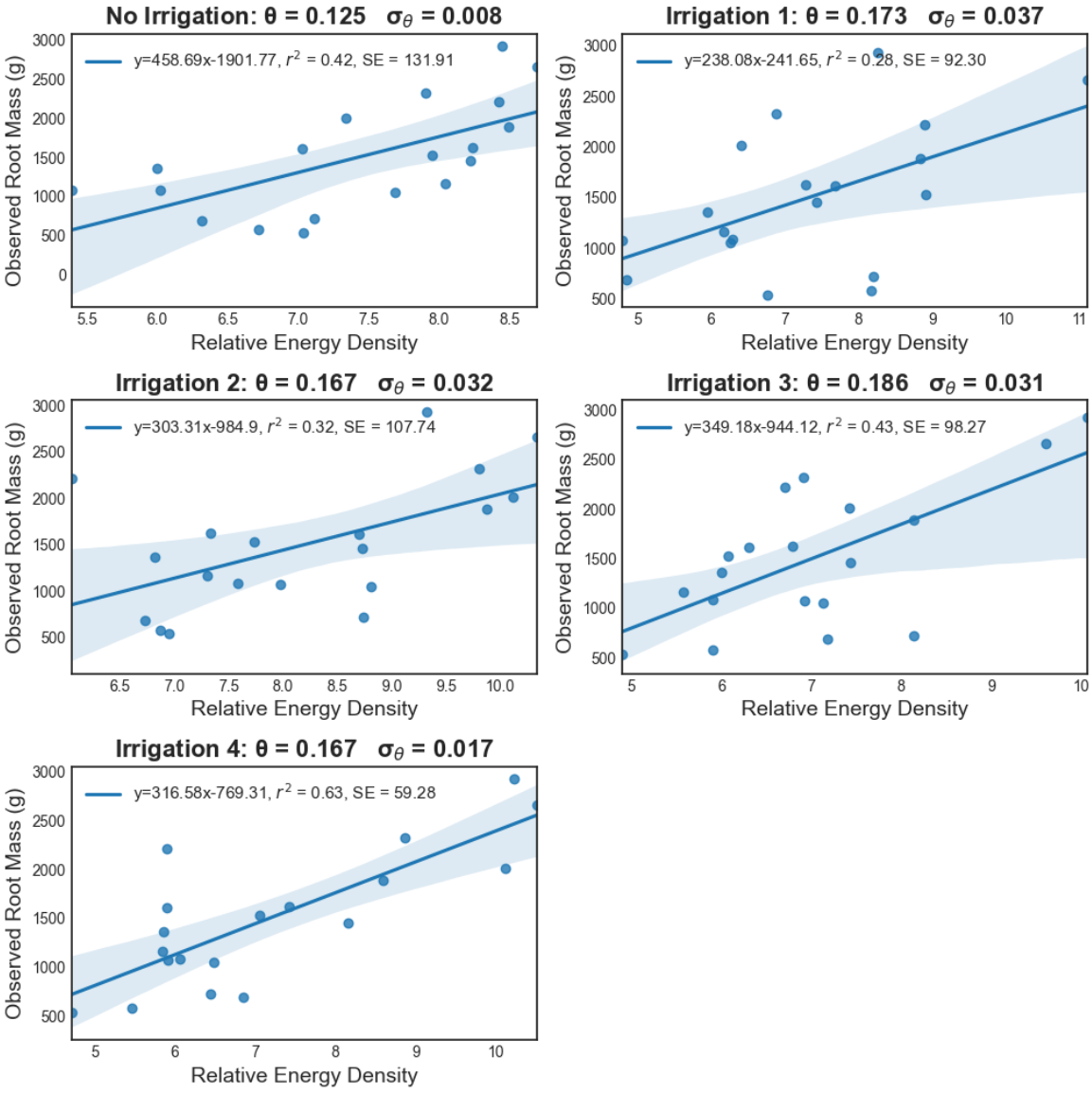

Figure 3.7 Correlation regression for all 5 treatments. Shaded area indicates 95% confidence interval based on 10,000 iterations of bootstrap testing. Note that Θ represents VWC. While the un-irrigated treatment had the lowest variation in VWC, the Irrigation 4 treatment had the strongest correlation, indicating the potential reduction of noise in wet soils vs dry soil. Irrigation 1 had the highest VWC variance, and the weakest correlation.

The window depth of significant correlations varied slightly between treatments, which is expected as an effect of the varying dielectric. The depth of the dry treatment was

greater than the wet treatments, which is unexpected, and will be discussed below. Standard deviation of the plot size, as measured by scan lengths, was 2.86 cm, less than 4% of the target plot length of 80 cm, and was randomly distributed in relation to plot number, observation, and treatment. Standard deviation between observations, within treatments, was not correlated to volumetric water content (VWC), variation of VWC, or plot biomass. The only exception for this is the Irrigation 4 treatment, in which relative energy density deviation was correlated to plot biomass at the $p < 0.05$ level.

Table 3.3 Summary of the mean relative energy density of each treatment, the depth of best correlation, and the regression results.

Treatment	Mean Energy	Energy σ	Depth (rows)	Pearson R	r2	Slope	Intercept	p-value
No Irrigation	7.422	0.954	66	0.645	0.415	458.7	-1901.7	0.0029
Irrigation 1	7.326	1.513	40	0.530	0.281	238.1	-241.6	0.0195
Irrigation 2	8.201	1.262	45	0.564	0.318	303.3	-984.9	0.0119
Irrigation 3	7.007	1.270	40	0.653	0.426	349.2	-944.1	0.0024
Irrigation 4	7.176	1.698	48	0.792	0.626	316.6	-769.3	0.000054

3.4. Discussion

Rapid estimation of bulked root mass is possible with GPR. These results show correlation strength up to 79% using these methods. Further, we have demonstrated that increased VWC can improve the detection of bulked roots, as long as the dielectric is homogeneous across the study. The bulk dielectric of soil is driven primarily by water, and the interfaces of dielectric change cause the reflection of GPR energy. With a sufficiently high signal frequency, soil structure has the potential to introduce noise in GPR data through the minute reflection and scattering of EM energy, driven by soil features such as compaction layers, aggregates, and pores. By increasing the VWC,

some soil pores fill with water, and the dielectric heterogeneities are reduced, leading to less noisy GPR data. Additionally, as the dielectric of the soil increases, the signal velocity reduces, effectively increasing the sampling resolution of the system, as the sampling is a function of time.

These two factors may explain why increased VWC improved the strength of correlation. Indeed, we hypothesize that as VWC variance decreases, the strength of the correlation should increase, possibly maximizing the predictive potential near field capacity. Unfortunately, soils near field capacity are easily compacted, and are difficult to work in. Therefore, some compromise must be found to maximize predictive power of the GPR and minimize the impact and difficulty of field work. This optimal level of soil moisture is most likely dependent on soil texture, and could be expressed as a fraction of field capacity. Further studies in multiple soil types should lead to standardized recommendations of optimum water content for major texture groups.

This study, like others before it, presents a supervised correlation – that is, the depth and mass of roots is known, so it becomes less difficult to determine the optimum depth of radar information to analyze. The window of analysis is relatively narrow – only five rows, or approximately 2 cm of soil depth – and selecting the correct depth without previous knowledge of the root depth is difficult at this time. As research continues, it may become possible to distinguish the zone of highest information density, and researchers are already working towards that goal [155]. In this study, however, the noise was too great to establish the root zone from only GPR data. For this application, noise may be considered as all recorded energy which is not reflected by plant roots. As

discussed earlier, GPR systems record all intercepted energy in the antenna range, regardless of the energy source. Noise may also be generated within the GPR system itself, and there has been some indication that the prototype system used here is not immune to this type of noise. This can be reduced by careful engineering, and through data filtering, if the inherent noise has been characterized. Other sources of noise include reflections and scattering caused by variations in soil structure, stones, clay clods, surface roughness, and above-ground biomass. Because we placed roots in the soil, rather than growing them, aboveground biomass was not an issue in this study, but has been in other data which are not yet published.

As noted, the soil type and water content have a large effect on GPR data. This variation makes it difficult to build a unified correlation between studies, fields, or even dates. As such, GPR remains a relative measure of root mass, suitable for ranking within a single field and date, otherwise requiring a specific calibration at each use. It remains possible that a correction factor could change this. Inclusion of multiple blank plots in the study may provide that correction factor, such that data can be normalized to the feature values of the blank plot, accounting for the soil type and moisture content. Further studies are planned to investigate this possibility. Without locational correction, GPR data may still be used to rank plots for genotype, and rankings may be compared across locations and/or time.

These results demonstrate the effect of soil moisture not just on the ability to pick out roots, but also the effect on the method. As mentioned in the results, the depth of best correlation was deeper for the dry treatment than the irrigated treatments, which was

unexpected. GPR energy is reflected at the interface of dielectric contrast. When the object causing the reflection, such as a root, has sufficient diameter, the reflection may happen at both interfaces on the signal vector – that is, it can reflect from both the top and bottom of the root. In other uses of GPR, the thickness of large objects can be estimated by measuring the distance between the top and bottom reflection. In this study, however, discreet returns were not observed; rather, the total reflected energy for a volume was measured. It is possible the reflected energy in the dry treatment was more intense at the bottom of the root, whereas the irrigated treatments reflected primarily from the upper interface. In all treatments, a nearly continuous range of window depths showed significant correlation to root mass, indicating that information about the root mass was present across a depth corresponding approximately to the root diameters. This also suggests the possibility of some distinct characteristic for that region, such that it may be possible to find that region using machine learning techniques, so that supervised correlation is no longer required, and furthering the usefulness of GPR as a predictive tool.

In 2019, Delgado et al. reported a similar study designed to test commercially available GPR models in bulked root imaging [151]. A C-Thrue GPR system (IDS Georadar) was mounted to a computer controlled gantry and passed over a climate controlled sandbox. Cassava roots of varying sizes were buried at orientations parallel, orthogonal, and 45° to the scan direction. A single antenna pair was passed in transects at 2.5 cm intervals over the sandbox, with signal pulses every 0.2 cm. The GPR data were interpolated to form 3D models of the buried roots and interpolated image dimensions were compared

to physical dimensions. The study illuminates several important factors for the application of GPR to root measurement, namely, the superiority of vertical antenna polarization over horizontal, and the effect of root orientation on measurement accuracy. However, the study differs significantly from the current – the focus was on 3D imaging rather than mass estimation, a single antenna pair was used in a high-density grid rather than an antenna array, the antenna was ground coupled rather than air-launched, and the soil medium was air dry such that no effect of soil water content was studied. Finally, the data collection method was not appropriate for high volume phenotyping.

The application of GPR for the quantification of roots is still in its infancy, and significant research is required before it can be used as a predictive tool. We have shown here that GPR data contain information about root mass, but it is also clear that other factors influence the data, and noise is a problem. Radar data is highly sensitive to processing parameters, such that adding or removing a step readily effects correlation. The presented methods utilized a multi-channel radar to rapidly collect 3D information. To produce the 3D information, the channels were interpolated using simple linear interpolation. Similar to other 3D data, such as LiDAR, care must be taken to align the interpolated entities. In the case of GPR, the primary point of alignment is usually the soil surface, because it is discrete and constant. In this study, the field was level, and the antenna was facing straight down at nadir, resulting in well-aligned channels with consistent positioning of the surface between channels. This is not always possible. In many cases, the antenna array cannot pass directly over the center of the root mass because plants are still present, so the antenna may be angled to point towards the plant

center. Additionally, errors in channel calibration can produce small offsets that change the apparent height of the antenna relative to the ground surface. Finally, uneven ground surface can cause differences between channels. In these cases, the ground surface must be identified in each channel so they can be aligned before interpolation, as described by Dobreva et al [155]. Automated methods of identifying the ground surface would greatly reduce the time required for channel alignment.

Though each application will have its unique problems, there remain several constant considerations which we suggest become standard practice when using GPR to measure roots. Foremost among these is to understand your radar system. Unlike visual tools such as LiDAR, GPR emissions are not highly focused and are generally shaped like an ellipsoid bubble, meaning the energy extends in front of, behind, and to the sides of the antenna. This is why at least 1 m was allowed between the cart and the first study plot, so that initial readings would be outside the plot. This also means care must be taken for transitory reflectors, such as workers, to not enter the volume of sensitivity while collecting the data.

GPR data are highly dependent on the dielectric of the soil; therefore, it is strongly recommended that dielectric measurements always be made at the time of scanning. This can be done in many ways, such as measuring dielectric directly with a probe, measuring water content and converting using the Topp equation, or by burying a reflector at a known depth, which allows a velocity estimation by dividing the known depth below the surface by the difference in signal time from the surface to the reflector [20, 166]. Knowing the dielectric, or the signal velocity (see Eq. 1), allows the conversion of data

from the time domain to the space domain, enabling estimation of depth. Further, some GPR processing techniques require these parameters. Many studies will be interested in the root mass at certain depths, as is currently measured with destructive techniques. This is only possible if the signal velocity is known.

Published methods to date have relied on measuring reflected energy, whether by amplitude threshold and pixel counting, or summations of other features. These techniques are inherently tied to the volume of soil analyzed, meaning that plot size will auto-correlate with feature count. Plot size must therefore be carefully controlled. In this study, plot size was controlled in the field through careful measurement and marking. Other studies have controlled plot size by cropping the data, and others have controlled by conversion to either feature density, root density, or both. We recommend the former whenever possible, as it protects the integrity of the data. However, current root phenotyping methods frequently use root density as a measurement and is acceptable to many researchers [167].

Whichever way the plot length is controlled, the data must be related to the field. Some GPR systems are capable of integrating GPS data into the scan data, while others can utilize digital markers. Some have neither capability, thus plot positions must be derived another way, possibly by placing reflectors at plot ends. Experience dictates caution in the latter method – the reflector must be easily identifiable in the radargram, and reflectors placed on the soil surface are easily lost in the surface reflection. In such cases, an aerial reflector is recommended.

Finally, based on the results of this study, care must be taken to ensure homogeneous dielectric environment at the time of scanning. Depending on the hydraulic conductivity of the soil, several days may be required after an irrigation event.

3.5. Conclusion

The use of GPR technology to quantify root mass in agricultural fields is still very young and is yet to be widely accepted. However, interest is growing as more studies are published showing the potential. With up to 63% explained variability (r^2=0.626, r=0.792), this study confirms previous publications, and demonstrates the feasibility of an air-launched antenna array for rapidly collecting GPR-based root estimations. Further, it is the first study to show improved data quality for wet soil over dry soil. It is novel in demonstrating the importance of dielectric homogeneity for estimating root mass. Though this study was small and should be confirmed with more samples, it serves as a proof of concept that merits further investigation. Importantly, we have demonstrated considerations in the use of agricultural GPR and begun the work of establishing a standardized method.

4. NON-DESTRUCTIVE GPR ESTIMATION OF CASSAVA ROOT MASS

4.1. Introduction and Background

More than 800 million people depend on Cassava (*Manihot esculenta Cranz*) as a staple food crop and it is the 3rd most important source of calories in the tropics [9, 11]. Throughout the tropics, cassava is frequently grown by subsistence farmers as a drought resistant crop, and is now increasingly grown as a cash crop for its unique culinary starch properties [10, 168]. The adoption of improved varieties by smallholder cassava producers in Nigeria has led to a measurable reduction in poverty levels [169]. Once considered an orphaned crop with little investment, cassava is now a major interest for several national and international crop improvement centers, and has shown a 150% increase in yield since the 1970's [170]. Despite all this, phenotyping tools to monitor or predict yield before harvest are seriously lacking, and the lack is hampering breeding efforts [6, 108]. The deficit of tools exists because cassava is a root crop. Phenotyping and yield prediction tools abound for crops with aboveground yield, but root and tuber crops have been left in the dark [171–173]. Attempts have been made to predict belowground yield from above ground parameters with varying degrees of success [174–176]. Other methods remain destructive, limiting usefulness for breeders and researchers who are often restricted by cassava's vegetative propagation to small plant numbers [108, 170]. Yield prediction and monitoring are important aspects of cassava breeding for several reasons: many varieties of cassava require 12 months or more to produce a full harvest, making selection of early genotypes a major goal for

breeders and because the reproduction rate of cassava is very low, destructive monitoring for early bulking is unavailable [157, 170]. Due to these constraints, a rapid non-destructive method to estimate bulked root mass continues to exist.

Ground penetrating radar (GPR) is a geophysical tool that uses electromagnetic reflections to explore below ground features and has been proposed as a solution for estimating bulked roots in agricultural fields. This application of GPR is still in its infancy but represents a growing field - GPR has been significantly correlated to root mass in cassava, wheat, energy cane, sugar beet, and peanut [8, 153–155]. With more research, it promises to be an effective tool for non-destructive monitoring of root and tuber crops, including cassava [170].

GPR works by emitting a pulse of electromagnetic energy into the ground, where it is either transmitted, absorbed, scattered, or reflected. Reflected and backscattered energy is recorded along with the time from the pulse. Reflections are caused by changes in dielectric permittivity, a measure of how strongly molecules can be polarized. In agricultural soils, the primary drivers of dielectric are soil texture and water content, with water content having the greater influence by an order of magnitude [19]. Therefore, the high water content of roots has the potential to reflect GPR signals. The reflections caused by roots are measured as the compound waveform caused by the constructive and destructive interference of all other reflections simultaneously being received by the GPR antenna, making the extraction of the root information non-trivial. As GPR is a mature technology in other fields, methods for processing GPR data exist, but tend to focus on a qualitative analysis, such as the location of belowground features

or objects. Quantitative analysis of GPR reflections is less common. For a further explanation of how GPR works, we refer the reader to the referenced publications on that topic [20–22, 177].

In 2021, Teare et al. presented a foundation on which to build common methodology for agricultural GPR, discussed some of the limitations, and suggested some directions for future work [178]. In this paper, we follow the data processing method they proposed and build upon it using GPR data collected from a cassava field trial in Cali, Colombia. We demonstrate significant correlation between extracted GPR features and bulked cassava root mass in four contrasting genotypes.

4.2. Methods

4.2.1. Location and Planting

The cassava field was located on the campus of the International Center for Tropical Agriculture (CIAT) near Cali, Colombia. The soil texture borders between loam and clay loam, with a gradient running north-south, the north having slightly greater clay content and the south having greater sand content. An aerial image of the field captured 3 months before harvest is shown in Figure 4.1.

Four cassava varieties were planted in four replications, with the replications running along the soil texture gradient. The varieties were GM3893-65 from the CIAT breeding germplasm, CM523-7, MPER-183, and HMC-1, commercial varieties popular in Colombia. Each plot or replicate was made of 5 rows, with 9 plants per row. Planting was done at the end of December 2017 using fresh stems. The field was kept weed free, irrigated, and fertilized as needed until GPR scanning in January 2019.

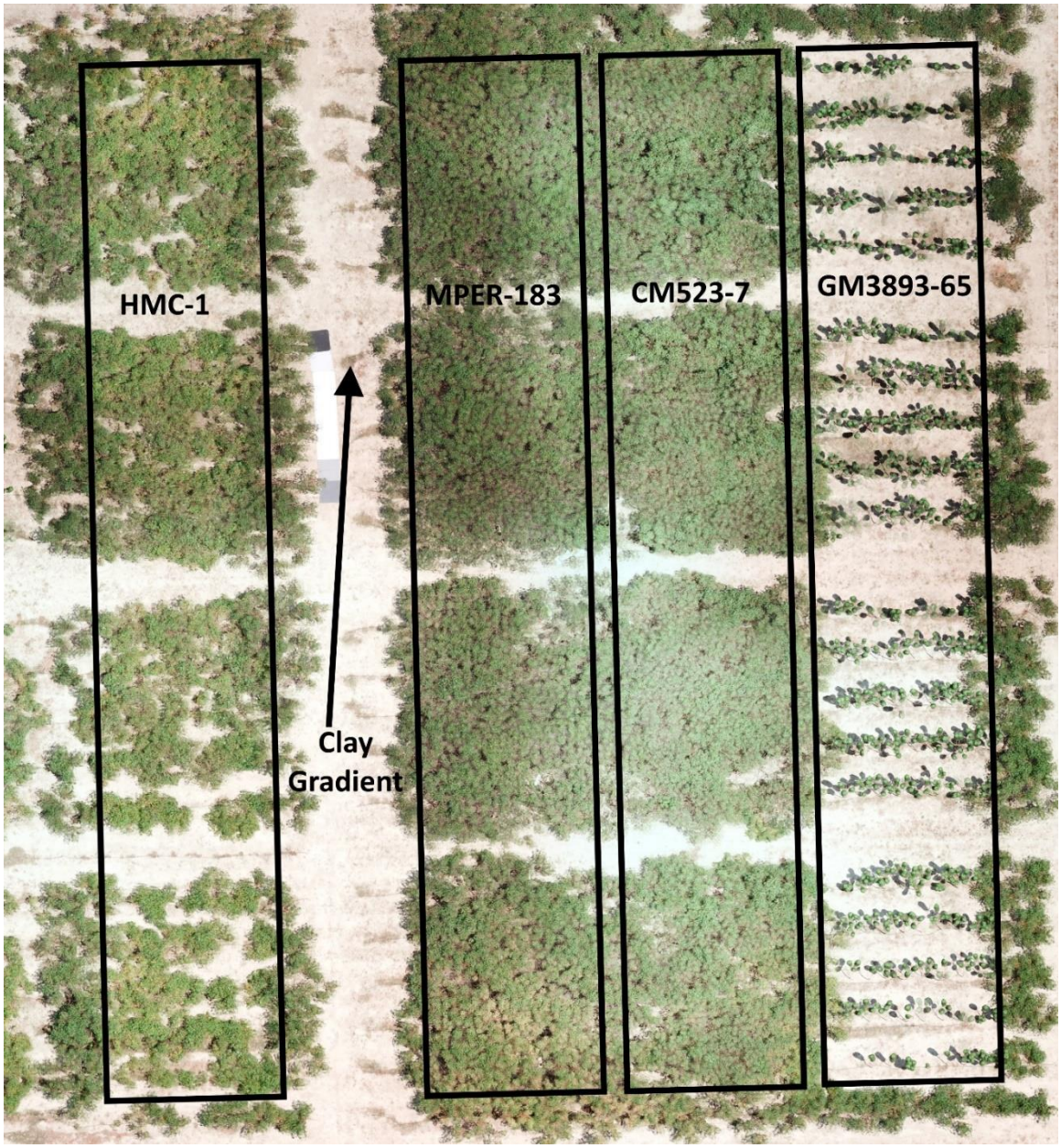

Figure 4.1 Replications ran across the major soil gradient. The difference in branching habits between genotypes is apparent in this aerial photo. Photo was captured 3 months before harvest.

4.2.2. GPR Sensor

The radar sensor used was an experimental loaded-vee dipole array, manufactured by IDS Georadar (Pisa, Italy) and has been previously described [163–165]. The array

consists of 4 transmitters and 4 receivers in alternating pairs, each spaced 4 cm from adjacent antennas (Figure 4.2). The antennas are wideband with a center frequency of 1.8 GHz. The radar records 512 samples over 18 ns, and pulses every 1 cm, as measured by an encoder wheel. Channel configurations paired every transmit antenna with an adjacent receive antenna, giving a total of 4 channels, each offset by 8 cm, as shown in Figure 4.2.

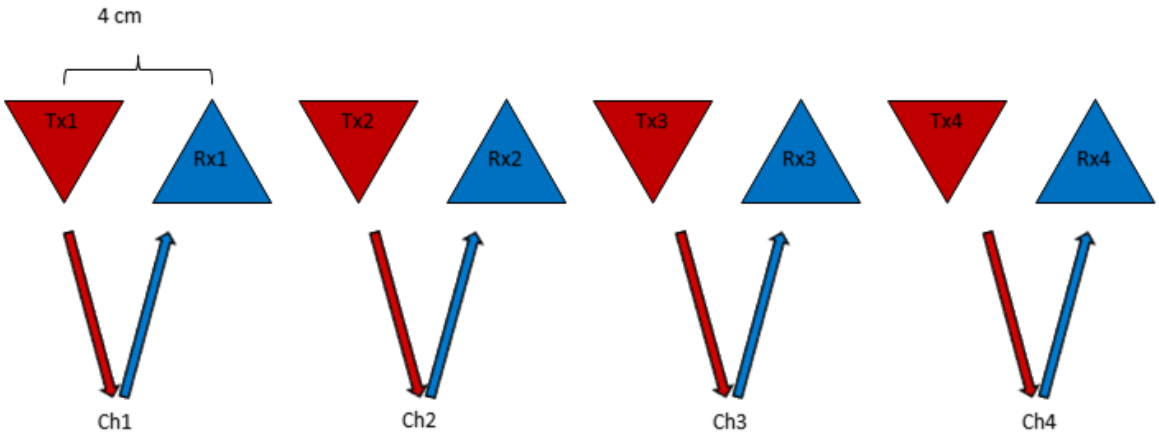

Figure 4.2 Arrangement of GPR antennas and the configuration of the data channels. Red triangles are transmitting antennas and blue triangles are receiving antennas.

The array was air launched and mounted on a 2 wheeled cart that moved along each plant row (Figure 4.3). The antenna was angled towards the center of the plant row, such that the look angle was 30 degrees off nadir. The highly branching genotypes required some branches be pulled back to allow the cart to pass.

Figure 4.3 The GPR cart ran on two wheels and carried the antenna array approximately 35 cm above the ground, allowing the array to be angled towards the center of the plant. The bycicle-like configuration allowed the cart access between plant rows. Some branches required lifting to allow passage.

Data Collection

The GPR cart was assembled at CIAT headquarters and transferred to the field, giving the electronics time to equilibrate to the temperature and humidity. Preliminary scans were collected on the border rows to allow the electronics to "warm up" before collecting experimental data. GPR capture was started with the front wheel adjacent to the first plant so that the forward looking nature of the antenna could capture the entire plot. Digital markers (fiducials) were placed in the data to mark the beginning and end of

each plot. All five rows were scanned in each plot. The field was only scanned one time because of time constraints.

Plants were harvested immediately after scanning was completed to avoid any change in root mass. Experienced field workers harvested the roots according to the standard method employed at CIAT. The fresh root mass of each plant was weighed in the field using a portable scale.

4.2.3. Data Processing and Analysis

Data were processed using GPR Studio version 1.0 (Crop Phenomics LLC, College Station, TX, USA (cropphenomics.com)), a Python software library developed for the quantitative analysis of GPR data. The software utilizes existing data processing libraries combined with custom-built functions specific to GPR analysis. Analysis was done in Python using the SciPy and Scikit-learn libraries. The workflow is summarized in Figure 4.4 below.

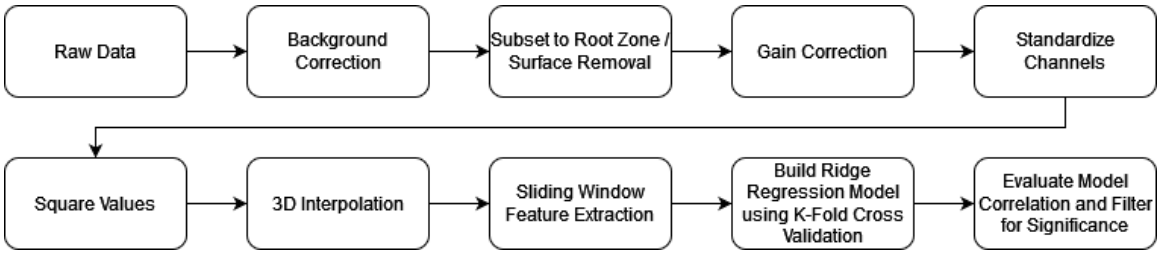

Figure 4.4 Data were processed and analyzed in a single flow using Python libraries.

Each scan was separated into plot rows using the digital markers placed during the data acquisition, then organized into a database according to genotype. Two of the channels exhibited significant noise and were discarded. Background correction was performed by mean subtraction filter, which corrected for banding noise inherent to this GPR system.

Using an automated surface detection algorithm, the data were subset to the approximate root zone, removing the ground surface reflection. Gain correction was applied to compensate for energy loss with signal depth and prevent bias in the model due to rooting depth. Each channel was then standardized to itself by subtracting the channel mean from each value, and dividing by the channel standard deviation, then squared to move all values to the positive domain and minimize the background information, as described in Teare et al. [178]. Standardization was only applied within channels, not between channels.

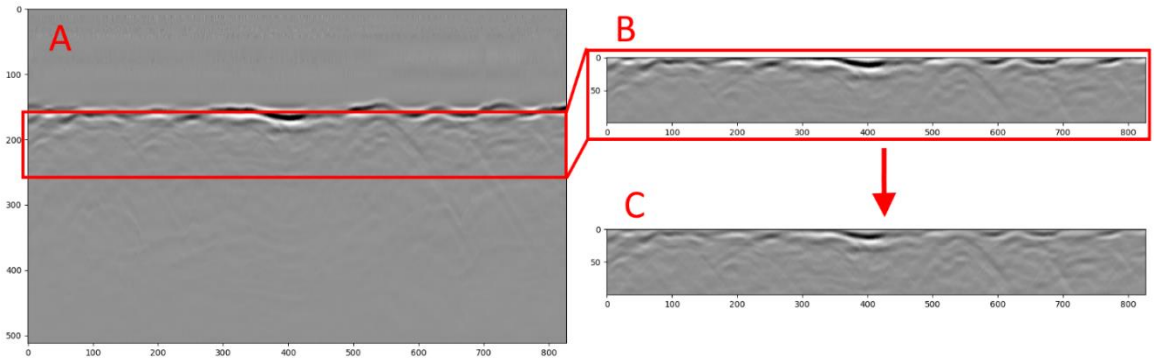

Figure 4.5A - A single channel of raw GPR data. 5B - Automatic surface detection was used to remove the surface reflection and subset the data to the approximate root zone. 5C - Signal gain was applied to offset the effect of energy loss with depth.

After standardization, the channels were interpolated into a 3D data cube via linear interpolation. A sliding window was passed from the top of the cube to the bottom, as shown in Figure 4.6. The window was 5 samples (image rows) thick, representing 0.175 ns of signal capture, or between 1 and 2 cm of soil depth depending on soil dielectric. At each depth, four features, or predictors described below and summarized in Table 4.1, were calculated within the window. These features carry information about the reflected energy and the below ground cause of the reflection. Feature 1 is the sum of all signal

amplitude within the window, and can be considered an indicator of the total reflected energy. Feature 2 is the standard deviation of signal amplitude, and can be considered a measure of the heterogeneity of the subsurface at that window location. Feature 3 is the mean of all signal amplitude, divided by the standard deviation, which is a measure of noise caused by the background soil matrix. Feature 4 is the inverse of the maximum amplitude, a parameter which may help control for outliers. These features are collected at either the genotype or field level and scaled between 0 and 1 within that group.

Table 4.1 Summary table of the GPR features extracted from the sliding window.

Feature 1	Feature 2	Feature 3	Feature 4
Sum of amplitudes	Standard deviation of amplitudes	Amplitude mean divided by the standard deviation	Inverse of the maximum amplitude

The scaled features are passed into a ridge regression model with the observed root masses, and the model is built using 10-fold cross-validation repeated 4 times, then the model performance is evaluated. Ridge regression was used to protect against bias caused by collinearity between predictors. The model is built and evaluated at each depth of the sliding window. Those depths which produced a correlation with p-value less than 0.01 are saved with the model inputs and outputs. For each genotype, the depth which produced the strongest correlation is selected as most representative of the true root depth then saved for further analysis. This process will be discussed further below.

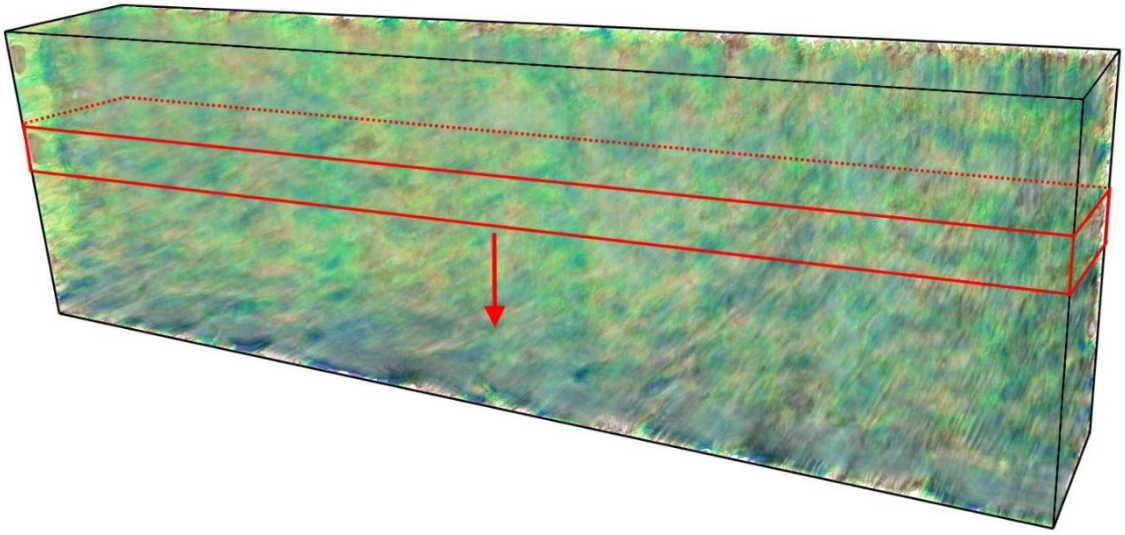

Figure 4.6 A sliding window 5 rows thick was passed from the top of the interpolated GPR data cube to the bottom, calculating indepent feature values at each depth.

Data were analyzed at both the genotype level and the field level. When considering the field level, two methods were used. The first evaluated the entire field as a single body, causing each genotype to use the same sliding window depth. The second method used the optimal depth previously calculated at the genotype level, accounting for genotypic differences in rooting depth.

4.3. Results

The genotypic effect on harvest mass is easily evident in the harvest data, with MPER-183 producing the greatest root mass, and GM3892-65 the lowest. The distribution of harvest mass for each genotype overlaps with the others, to create a wide and nearly continuous variation in root mass from 6 to 63 kg per row. This overlap in distribution is displayed in Figure 4.7 below, and is significant because it reduces gaps in model training data.

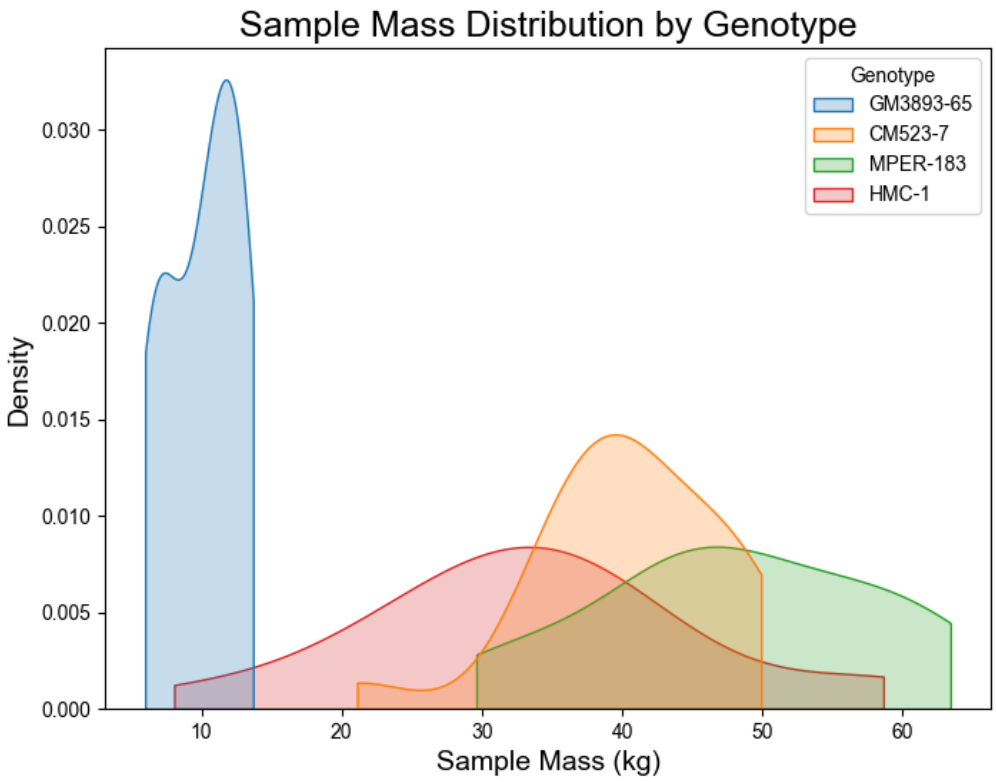

Figure 4.7 The distribution of root mass by genotype. This shown as a kernal density plot, which smooths the distribution between bins, but simplifies the display making interpretation easier.

Our model prediction results demonstrate a significant relationship between the reflected GPR signal and the bulked root mass of cassava plants. Figure 4.8 below illustrates the relationship between each feature and the root biomass. The scatter matrix shows the relationship between each component of the model in piecewise fashion, each component occupying both a column and a row, the intersection displaying the scatterplot relationship between. The lower triangle groups the genotypes together and fits a linear regression line to display the correlation coefficient of each component. Note the high correlation between the GPR features, demonstrating the need for penalized regression to protect against bias and overfitting. The upper triangle shows the data by

genotype, illustrating how genotypic variation serves to cover a wider spectrum than any genotype alone. The bottom row and the last column show feature relationships to the observed root mass, while the top row and first column show the predicted root mass.

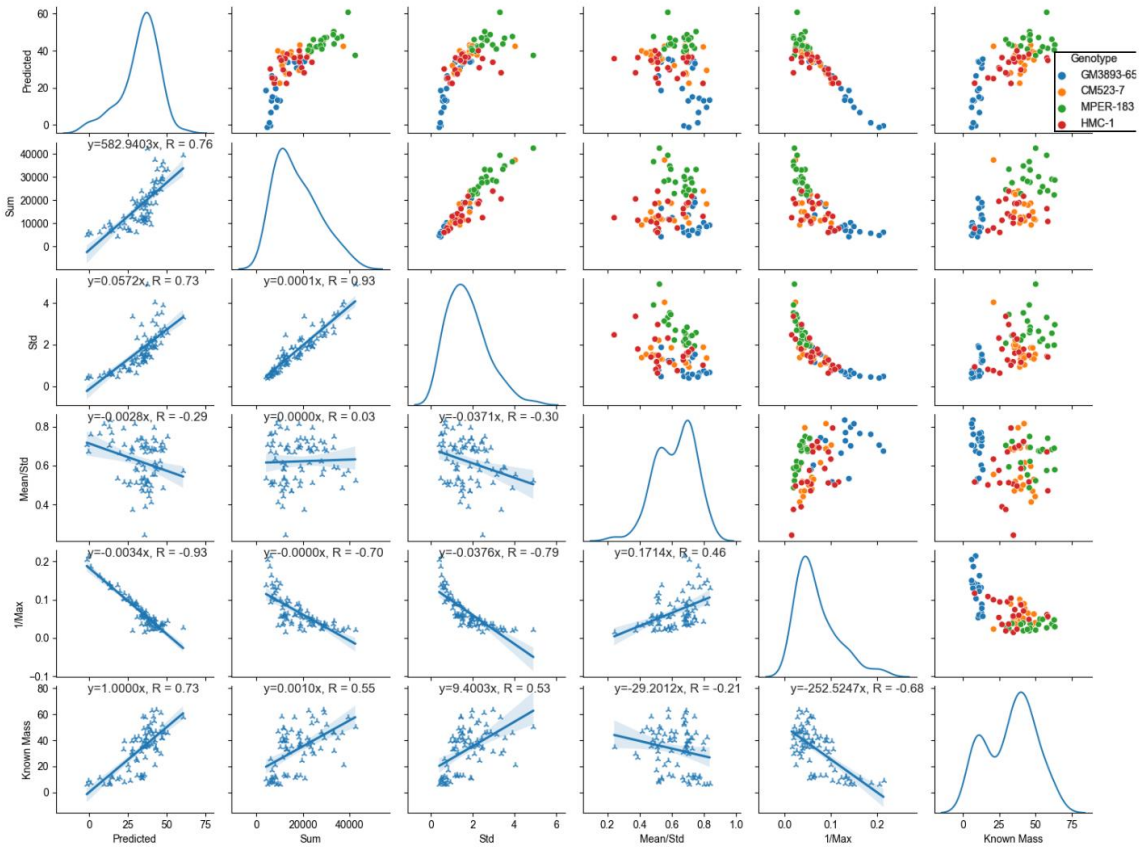

Figure 4.8 A scatter matrix showing the relationship between each feature and individual relationships to root mass. Correlation values are calculated by simple linear regression between the values shown. Shaded areas represent the 95% CI of the mean calculated by 10,000 iterations of bootstrap resampling.

Model correlation to observed root mass was significant at the $p < 0.005$ level within all genotypes and between all genotypes. CM523-7 showed the strongest correlation, likely driven by a single low weight sample acting as a leverage point, and also the shallowest roots. MPER-183 had the weakest correlation, possibly because of high phenotypic variation in the roots. Correlation between genotypes reduced significantly compared to

within genotypes when calculated from a single window depth, but remained high when the model was built using GPR features for each genotype were derived at representative window depths. The regression coefficient between predicted and observed values compares favorably to a 1:1 line, indicating little tendency to favor either under or over-estimation.

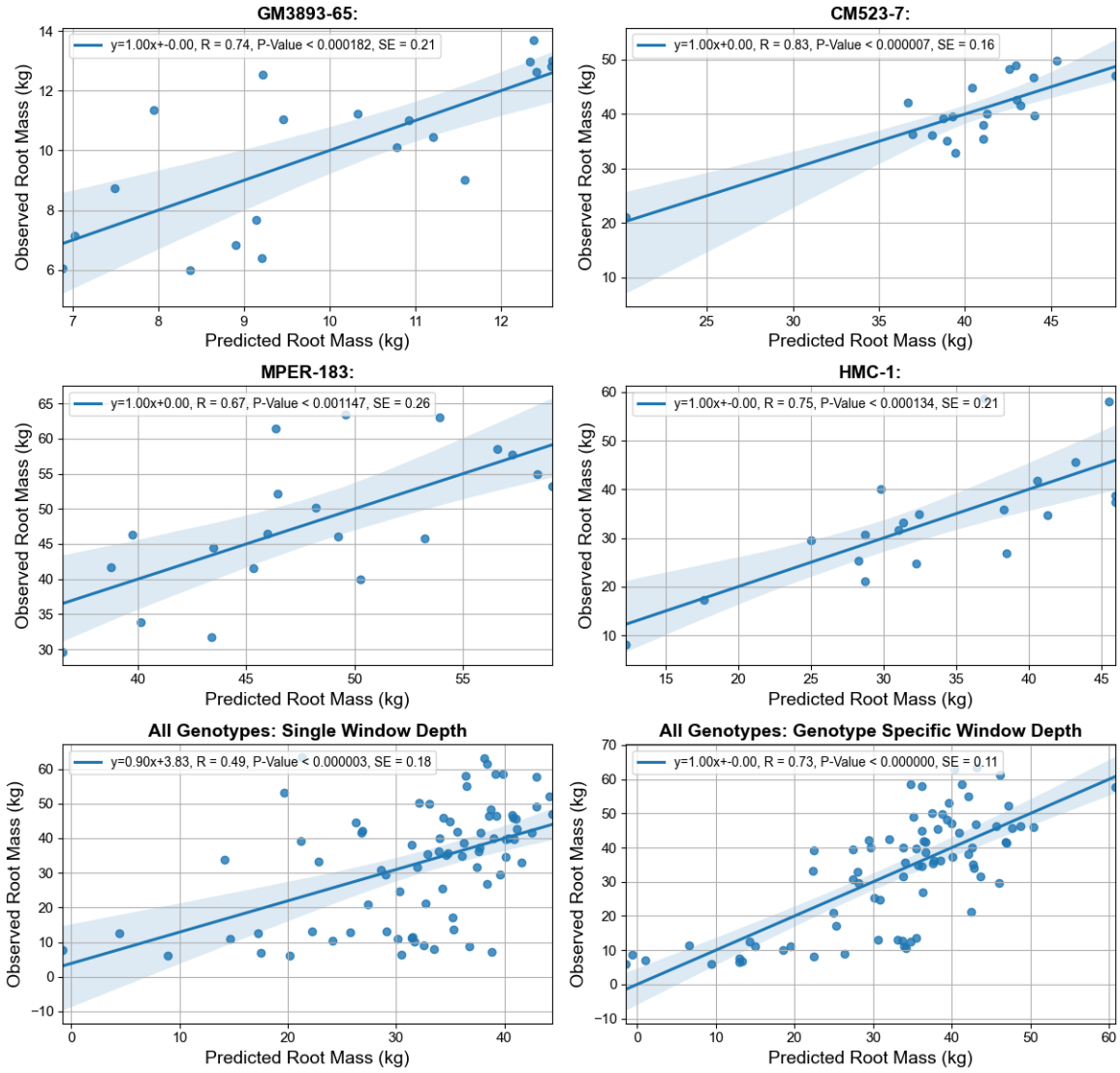

Figure 4.9 Regression plots between the observed root mass and the model predicted root mass, demonstrating the accuracy of the model. The top 4 plots show the model at the genotype level, while the bottom 2 plots show the model at the field level. Model at the bottom left used features extracted at the field level, with no window depth adjustment between genotypesShaded areas indicate the 95% C.I. of the mean based on 10,000 iterations of bootrapping. SE = Standard Error of regression.

4.4. Discussion

Reflected GPR energy contains information about cassava bulked root mass, making the estimation of that mass possible. Our results show correlation strength up to 83% within a single genotype, and 73% between 4 genotypes. Importantly, this demonstrates that the generality of the model extends to multiple genotypes. However, it remains to be seen if the model can be generalized between locations, and especially divergent soil types and soil moisture. This must be a focus of future work – if the predictive model cannot be made robust against location or soil conditions, a calibration would be required for nearly every application. The model may even need to be retrained for different soil moisture conditions within the same field. We therefore recommend further work exploring this aspect of agricultural GPR; it is entirely possible that a correction factor may be found to account for varying soil moisture or even varying soil textures. One such solution could be the inclusion of no-root control plots in every acquisition. These control plots could be used to standardize the radar data to the soil conditions present at that time and location. Another possibility lies in adding a soil moisture parameter to the model, which would require the measurement of soil moisture at each application, though it is possible the radar system itself could be used to estimate that parameter [179, 180].

This report represents the first non-destructive use of GPR for cassava root estimation as previous authors removed plant stands to aid data capture [8]. However, like the preceding reports, this method requires a supervised method for determining the correct root depth to optimize the model. In this case, we require the correlation of features from

several depths to observed root mass. Even so, the model built from a single window depth (Figure 4.9, bottom left) resulted in a small standard error of 0.18. Reliance on supervised root depth detection is a weakness when considering GPR as a predictive tool; however, the application is still in its infancy – future work will undoubtedly present a method to predict the correct depth without a priori knowledge of root mass or depth. Indeed, we hypothesize that a machine learning model could be trained to recognize the proper depth using the derived GPR features described above, and perhaps other features. Such a tool would allow for the use of the ideal depth on each sample, leading to even more robust models.

The smallest usable sample in this study was a plot row because the individual plants were not reliably discernable within the GPR data, therefore, each sample represents up to 9 plants within a single plot row. Because the repetitions were planted across the soil texture gradient, models for each genotype include samples from the range of soil type. Previous publications suggest model specificity to soil textures and moisture content, which if true, could explain some of the error in our models [8, 178, 181]. However, it's also possible the standardization of the channels in processing, and the transformation of the extracted features for model building help to counteract that effect. Further work is planned to explore whether this model will continue to perform well when considering more varied soil conditions. As it is now, GPR models for root mass estimation should not be considered as generalized to multiple locations.

We have suggested here two directions for future work, the automatic detection of root depth without prior knowledge of the roots, and the generalization of predictive models

across locations and soil conditions. While these directions focus on processing and modeling, further research is also required in agricultural GPR hardware. Optimization of antenna design, frequency, and arrangement could improve the suitability to widely varying field conditions while simultaneously reducing error, thus increasing the usefulness of GPR as a predictive tool.

4.5. Conclusion

The methodology and results presented here support and build upon previous publications, especially those of Delgado et al. and Teare et al. [8, 178]. They represent an important milestone in the development of a non-destructive agricultural GPR instrument and data analysis method for cassava. The use of GPR for estimating the bulked root mass of agricultural crops is still in its infancy and significant investment and research will be needed before breeders will be able to use it reliably. However, interest among researchers and breeders is increasing, and progress is continuing. With up to 69% explained variability (r^2=0.689, r=0.83) and standard error as low as 0.06, this study is beginning to approach a level of accuracy which will allow the high-throughput nature to overcome the error when used in multi-location heritability studies [182]. The continuing improvement of agricultural GPR will enable cassava breeders to evaluate breeding trials for early bulking, among other traits, and contribute to food security throughout the tropics.

5. CROSS LOCATIONAL VALIDATION OF THE RELATIVE ENERGY DESNITY MODEL FOR ESTIMATION OF BULKED CASSAVA ROOTS WITH GPR

5.1. Introduction

Cassava (*Manihot esculenta* Cranz) is a tropical shrub grown mainly for its large starchy roots. Often considered a food security crop, cassava accounts for a large portion of daily caloric intake for over 800 million people globally [9]. Recent years have seen the adoption of the crop into the industrial supply chain as a major source of plant starch in manufacturing [183]. Besides being tolerant of poor growing conditions, cassava has a long harvest window between 6 and 24 months after planting, however, maximum root mass generally occurs at 12 months [7]. This long maturity period is a significant challenge for producers and breeders, exacerbated by the low reproduction rate of the vegetatively propagated crop [6]. Early maturity is therefore a major goal for cassava breeders, but is hampered by the inability to estimate root mass without destructive harvest – something untenable in trials with as few as 3 plants per genotype [170]. Therefore, the need exists for a method to non-destructively estimate cassava root mass, and ground penetrating radar (GPR) has been proposed as such a method [8].

GPR has been used for some time in the detection and mapping of coarse tree roots [152]. More recently it has been used to estimate the root mass of agricultural crops, including sugar beet, wheat, peanut, energy cane, and cassava [8, 155, 161, 162, 178]. These studies showed that GPR data taken in an agricultural field contains information

about the plant root mass, and demonstrate the potential to predict root mass from GPR data. However, the authors also acknowledge that further development is needed to achieve the high precision required of phenotyping tools used by plant breeders and researchers. Each of the studies was somewhat limited in size, with few, if any, replications in different locations. It has been noted that GPR prediction models are likely tuned to local soil conditions such as texture and moisture, and as such, may not generalize well [8, 181, 184].

GPR works by emitting an electromagnetic (EM) pulse into the ground. The energy is either transmitted, scattered, absorbed, or reflected by the subsoil features, including root mass. Reflection is caused by changes in dielectric, a property which describes a materials ability to be polarized at the molecular level. Therefore, in a uniform material, such as air, there is no significant interaction between the EM energy and the material, and no reflections would be recorded by a GPR. Agricultural soils, however, are far from uniform, and EM scattering and reflection are caused by variations in texture, structure, aggregates, soil water, roots, and more [155, 178]. Additionally, the amount of interaction between the EM energy and subsoil features will be affected by the GPR frequency, with higher frequencies (smaller wavelengths) interacting with smaller features [20]. These characteristics of GPR data in agricultural soils result in a large amount of noise, that is, energy reflections not caused by root mass. Thus, predictive models which do not account for variability in soil conditions may not generalize well across locations.

In this paper, we build a root mass prediction model for cassava using GPR data collected at two distinct locations. We then test the performance of the models across the locations.

5.2. Methods

5.2.1. Location

Cassava fields were planted and maintained at the Center for International Tropical Agriculture (CIAT) in Cali, Colombia, and at the International Institute of Tropical Agriculture (IITA) in Ibadan, Nigeria. The soil at CIAT is loamy, bordering on clay loam, while the soil at IITA is sandy loam, with frequent ironstone concretions and some gravel. Both soils are tropical derived savannah.

At CIAT four cassava varieties were planted in four replications. The varieties were GM3893-65, a CIAT breeding line, CM523-7, MPER-183, and HMC-1, commercial varieties. Each plot was made of 5 rows, with 9 plants per row. Plant spacing within row was 1 m, and between row spacing was 2 m. Planting was done at the end of December 2017 using fresh stems. The field was kept weed free, irrigated, and fertilized according to standard practice at CIAT until GPR scanning in January 2019.

At IITA, three cassava varieties were planted in 3 replications with 7 planting dates. The varieties were TMEB419, TMEB693, and IBA070539, all sourced from IITA germplasm. The plots consisted of 6 plants per row, and 5 rows per plot. Plant spacing within rows was 0.8 m, and between row spacing was 1 m. A staggered planting design was utilized to increase root mass variation at harvest. Planting 1, consisting of 3 replications for each of the 3 varieties, began in June 2018, the subsequent 7 plantings

occurred every 3 weeks, with 3 replications for each of the 3 varieties. The field was kept weed free and irrigated, but no fertilizer was applied, according to standard practice at IITA, until harvest in April 2019.

5.2.2. GPR Sensor

The GPR sensor used an array of loaded vee dipole prototype antennas, developed by IDS Georadar (Pizza, Italy). The antennas have been previously described [163, 165]. The array was air-launched on a two wheeled cart which also carried the electronics and data collection portion of the GPR system. The GPR system is wide-band, centered at 1.8 GHz, with 4 transmit, and 4 receive antennas in the array, each spaced 4 cm from its neighbor. In both locations, data channels paired each transmitter with the immediately following receiver, so that there were four equally spaced channels centered 8 cm apart (see Figure 4.2 above). The system was configured to collect a data trace every 1 cm of travel, as measured by an encoder wheel. The time window was 18 ns, and 512 samples were collected per trace, or one sample per 0.035 ns. Because the plants were too tall to move the cart directly over, the cart moved along side each row, and the antenna was angled towards the bases of the plants, or approximately 30° up from nadir (Figure 5.1).

Figure 5.1 The GPR cart traveled on 2 wheels to accommodate the narrow row spacing, but still required some branches be pulled back to allow passage. The electronics were mounted on top, the tablet computer at the rear, and the antenna slung beneath at an angle towards the plant row.

5.2.3. Data Collection

At both locations, the GPR cart was assembled before transport to the field, allowing time for the GPR system to come to equilibrium with outdoor temperature and humidity. Several preliminary scans were performed to "warm up" the antenna before scanning the fields. Data collection began outside the plot row, and digital markers (fiducials) were placed in the data to mark the start and end of each plot. The fields were scanned only once because of time constraints. At IITA, only the inner 3 plot rows were scanned, but at CIAT all rows were scanned. Scanning is performed by walking the GPR cart along

the plant row while the data collection program is running. Plants were harvested immediately after GPR scanning was completed. Harvest was performed by experienced field technicians according to the standard practice at each research center, and fresh root mass was weighed in the field by a portable scale.

5.2.4. Data Processing and Analysis

Data processing was performed in Python 3.8 using GPR Studio (Crop Phenomics LLC, College Station, TX (cropphenomics.com)), a software package designed especially for the quantitative analysis of agricultural GPR data. GPR transects were split into plot rows using the digital markers placed during collection, then plot rows were assembled into plot replications and genotypes using the database management in GPR Studio. Analysis was also performed in Python using the Scikit-learn library. The workflow is summarized in Figure 5.2 below.

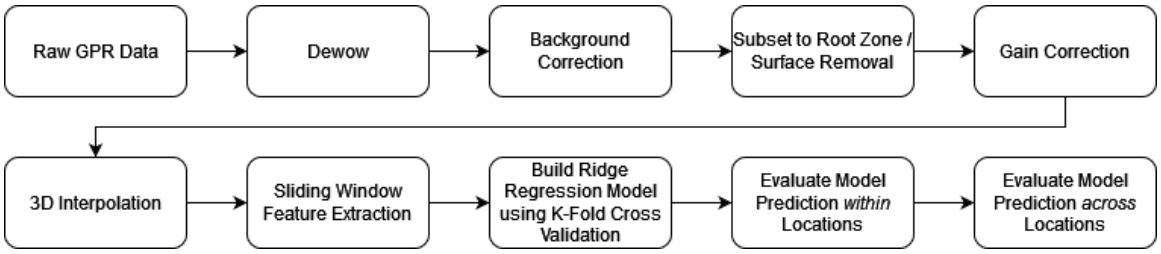

Figure 5.2 Data processing and analysis workflow done in Python.

In both data sets, channels 3 and 4 exhibited significant and apparently random noise and were therefore unfit for use in data analysis. The data were first passed through a dewow filter which removes the effect of very low frequency noise and also centers the data traces at 0. The data offset can vary by channel and is an artifact of the GPR control

system. Background correction was applied as a mean subtraction filter to remove horizontal banding noise inherent to this prototype hardware. The data were then subset to the approximate root zone to remove the surface reflection and ground clutter. For the CIAT data, an automatic surface detection algorithm was used to find and remove the ground surface, but the IITA data exhibited greater variation in the position of the surface reflection within each radargram, requiring manual detection of the soil surface. Gain correction was applied to adjust signal amplitude for attenuation and avoid a depth bias in the model. Finally, the channels were interpolated into a 3D block. The results of each step can be seen in Figure 5.3 below.

GPR feature extraction was done by a sliding window, based on the description in Chapter 4 above. The window thickness was set at 5 data rows, representing 0.175 ns of signal capture, and passed from the top to the bottom of the subset. At each depth multiple features were calculate based on the signal amplitude values contained within the window. The features extracted at each window depth were correlated to fresh root mass using ridge regression and evaluated by k-fold cross validation. Data processing and feature parameters were tuned based on the cross validation results. Cross validation was also used to select the appropriate penalty parameter, lambda, for the ridge regression. Lambda values ranging from 1e-3 to 1e2 on a log scale were considered, 0 inclusive, and were scored and selected by the median absolute error of the model.

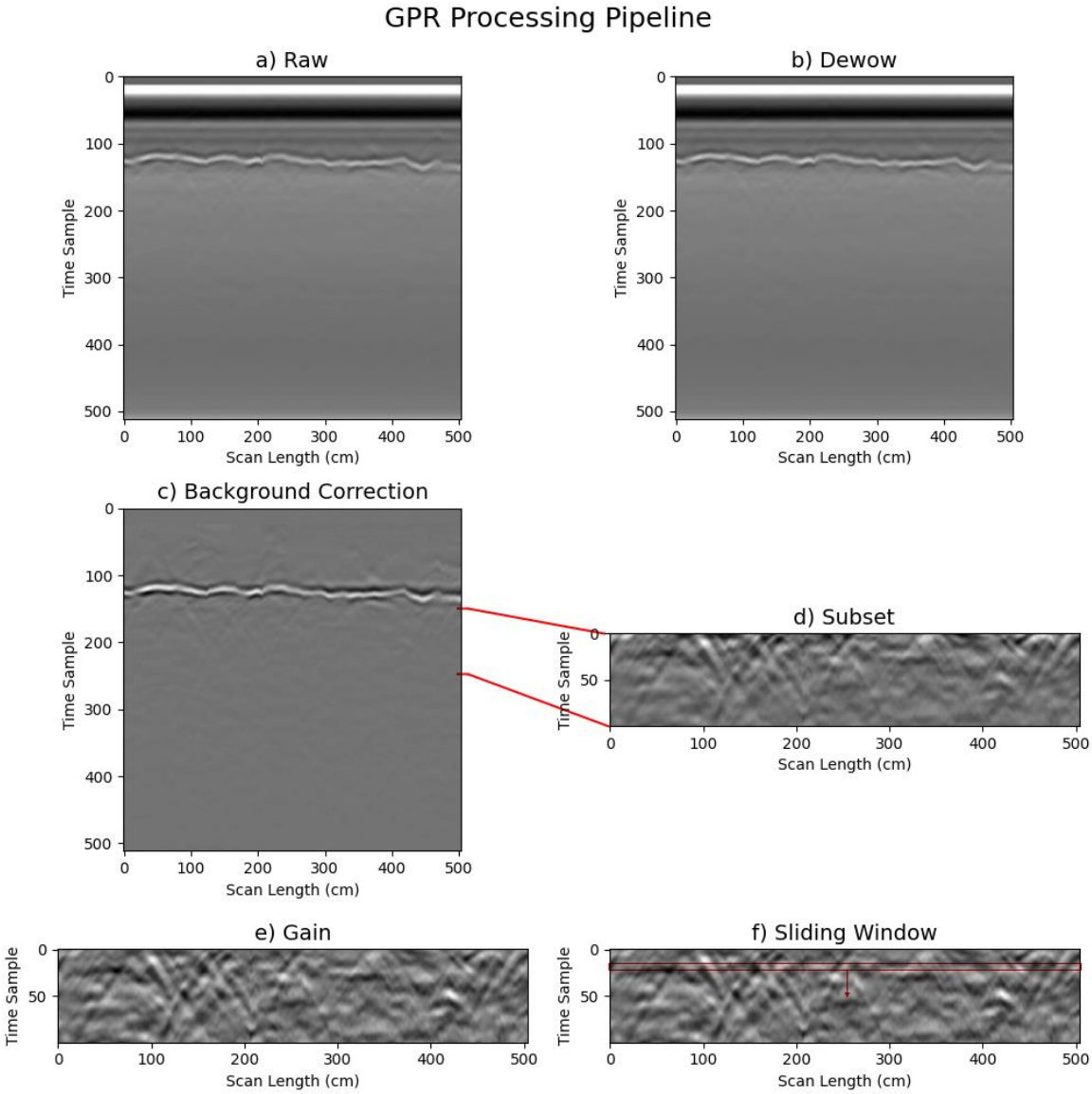

Figure 5.3 The visualized results of each processing step in the workflow. Notice especially the effect of the background correction in c) as compared to b), and the effect of gain correction in e) as compared to d). The lines connecting c) and d) indicate the area from which d) is taken. The red box in f) shows the 5 line thick sliding window in its downward movement.

Feature tuning at both locations resulted in four features used as predictors in the final model, summarized in Table 5.1 below. The first feature is the sum of amplitudes divided by the length of the scan. The division by scan length is important to avoid introducing error or potential bias caused by variance in the scan length, which can be

caused by both error in plant spacing and errors in placing the digital markers. This feature has previously been considered a measure of total reflected energy, but because these data were not forced into the positive domain and are relatively centered about zero amplitude, this feature trends towards the mean, in which case, it represents the soil matrix. The second feature is the standard deviation of signal amplitude, which can be considered a measure of heterogeneity in the subsoil such as is caused by roots. The third feature is the 75^{th} percentile, the value at which 75% of amplitudes lie below, and 25% above. Similarly, the fourth feature is the 25^{th} percentile. These values are robust against outliers compared to the maximum and minimum, and, because the amplitudes are centered about zero (including both positive and negative values), they both relate to amplitude values caused by reflections significantly brighter than the soil matrix, such as caused by root mass. Extracted features were conditioned for model building using Scikit-learn RobustScaler function, which is similar to standardization except that it uses the inner-quartile range rather than the mean, making it robust to outliers. Even so, some extreme outliers were removed from both datasets during quality control.

Table 5.1 Summary of the GPR features extracted. Features are relative to the amplitudes contained within the sliding window.

Feature 1	Feature 2	Feature 3	Feature 4
Sum of window amplitudes, divided by scan length.	Standard deviation of signal amplitudes.	Amplitude value at the 75^{th} percentile.	Amplitude value at the 25^{th} percentile.

The most appropriate depth estimate for root location was selected by correlation strength between model predictions and observed values, accounting for appropriateness and filtering for significance level $p < 0.01$. The data were analyzed at the genotype level to account for variation in rooting depth by genotype.

Independent models were built and tested for each location using 10 by 3 k-fold cross validation. Then, a model was trained using the features from Location 1 and used to predict mass at Location 2 based on the features from Location 2. The model was tested in both directions, that is, training on IITA to predict CIAT, and training on CIAT to predict IITA. Training and validation were performed using the same k-fold cross validation procedure.

5.3. Results

Analysis of the model performance by linear regression between predicted and observed root mass showed significant prediction strength which varied by model, shown below in Figure 5.4 ($p < 0.001$). The CIAT model performed best with 91% correlation ($R = 0.91$), while the IITA model predictions correlated at 59% ($R = 0.59$). When the models were applied across locations, performance decreased markedly while maintaining significance. The CIAT model predicted IITA root mass at $R = 0.56$, while IITA predicted CIAT root mass at $R=0.45$. The median absolute error calculated during k-fold cross validation is summarized in Table 5.1.

Table 5.2 A table summary of performance metrics for each model. The mean absolute error was calculated from 30 iterations of k-fold cross validation, while Pearson's was measured from the final model predictions.

	CIAT Model	IITA Model	CIAT predicts IITA	IITA predicts CIAT
Mean Median Absolute Error (MAE)	11.89	1.58	11.88	1.58
Standard Deviation of MAE	4.23	0.44	4.36	0.45
Lambda	1.0	1.0	0.1	0.1
Pearson’s R	0.91	0.59	0.56	0.45
Mean Observed Mass	32.65 kg	3.34 kg	3.34	32.65
Mean Predicted Mass	32.65 kg	3.34 kg	30.24	3.64

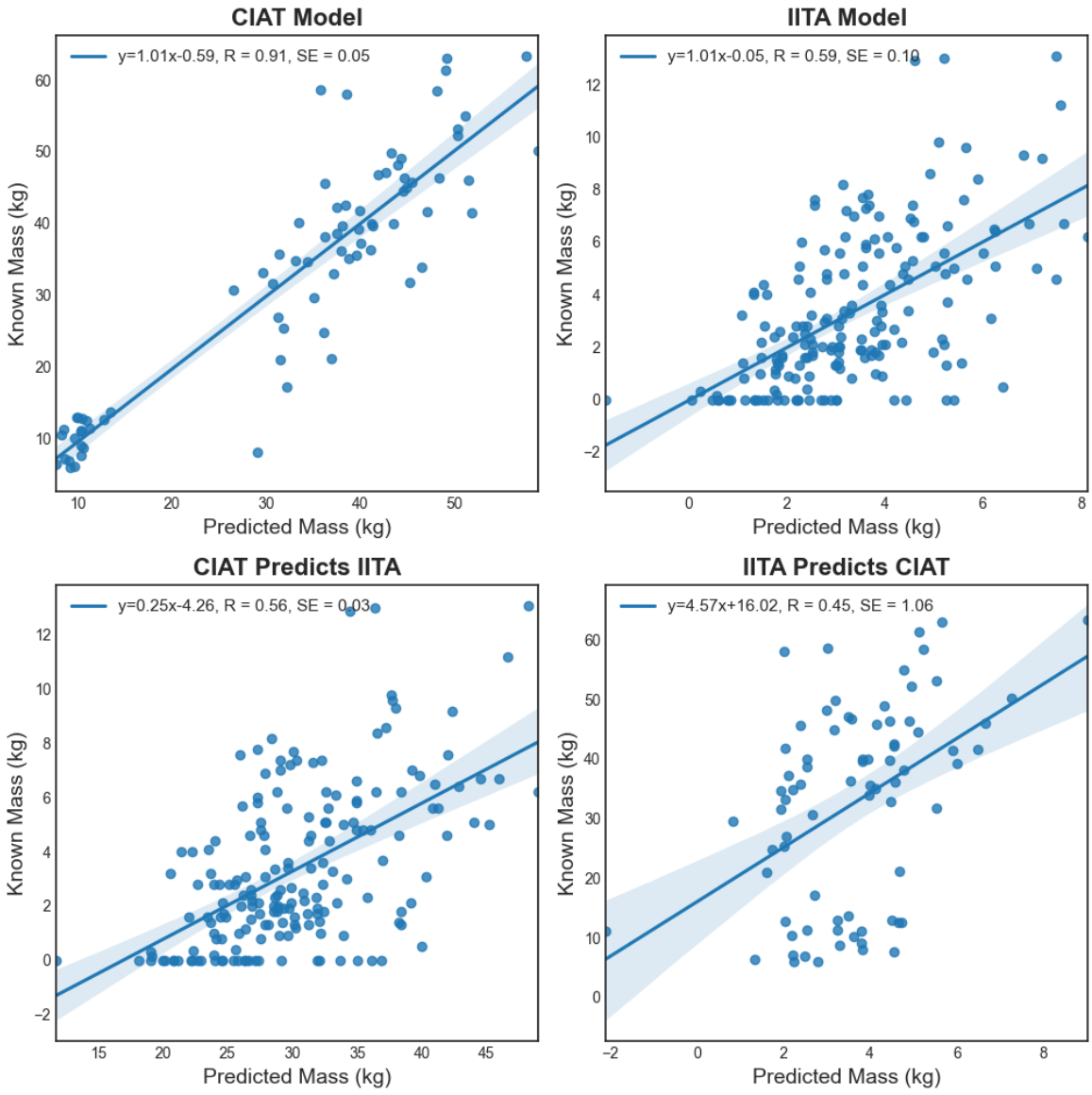

Figure 5.4 The regression results of the predicted root mass versus the observed root mass for each model. The top two represent models built and validated at a single location, while the bottome two represent models trained at one location and tested at the other.

The GPR features extracted showed varying degrees or relationship to the observed root mass, and the strength of relationship for each variable differed between locations. The locational variance of features and their relationship to observed mass meant that finding a feature set which worked for both locations precluded optimizing the model for either location. The relationship between each feature is illustrated below in Figures 5.5 and

5.6. The relationship of each feature to the observed and predicted root mass is also shown. These scatter matrixes arrange several scatter plots, the variables of each plot being labeled on the outer axes. Each variable is represented on both the X and Y axis, and the relationship between any two features can be seen at the intersection of their row and column. This arrangement is convenient in showing the complexity of multiple relationships in multilinear regression but duplicates the results in an upper and lower triangle, albeit rotated by 90 degrees. We have taken advantage of this duplication to show different aspects of the data.

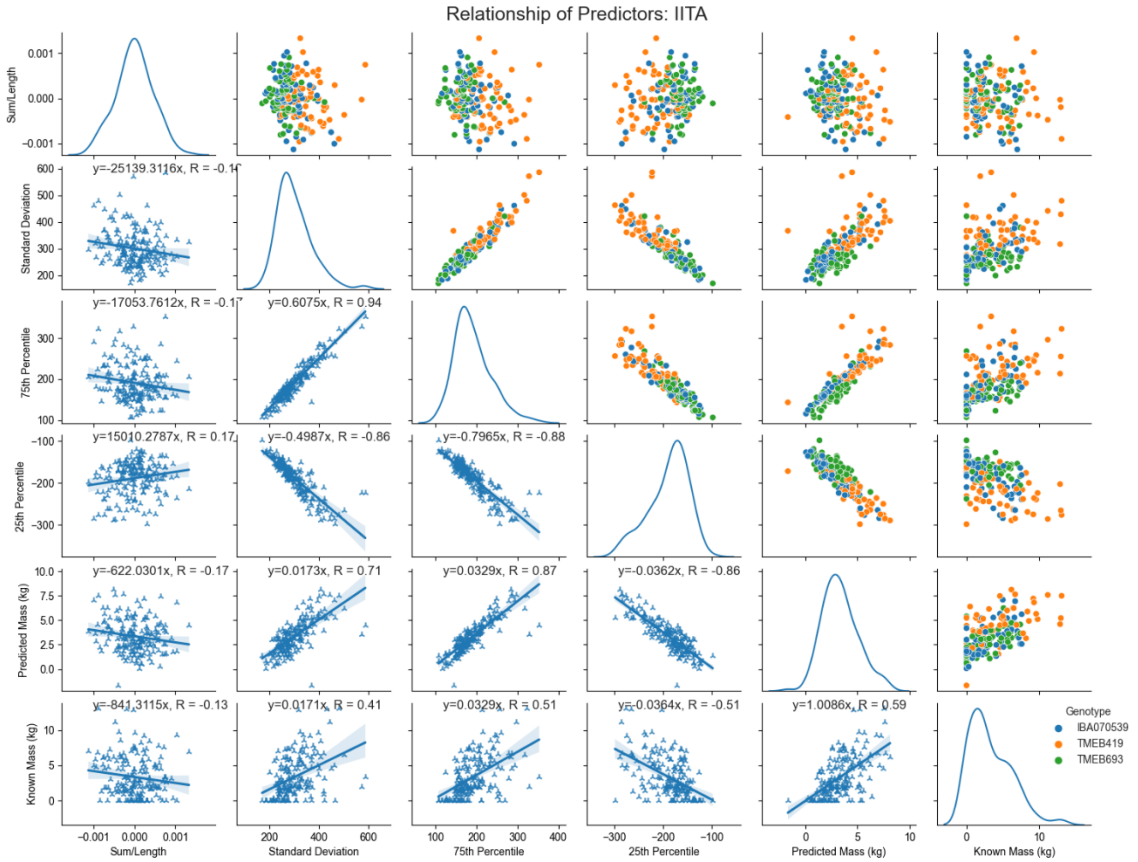

Figure 5.5 A scatter matrix showing the relationship between each feature and individual relationships to root mass for the model built from IITA data. Correlation values are calculated by simple linear regression between the values shown. Shaded areas represent the 95% CI of the mean calculated by 10,000 iterations of bootstrap resampling. The feature values displayed here are not regularized.

The lower triangle shows the entire data set without distinction by genotype or replication, allowing the regression between the entire feature set. This triangle also shows the relationship of each feature to the observed mass in the bottom row, and the predicted mass in the second to bottom row. Note that none of the features show a strong relationship to observed root mass at both locations, illustrating the need for multilinear regression. The strength of relationships between features explains the need for penalized regression. Cross validation was used to estimate the optimal penalty value (lambda). The penalty term shrinks the regression coefficients to reduce bias and avoid

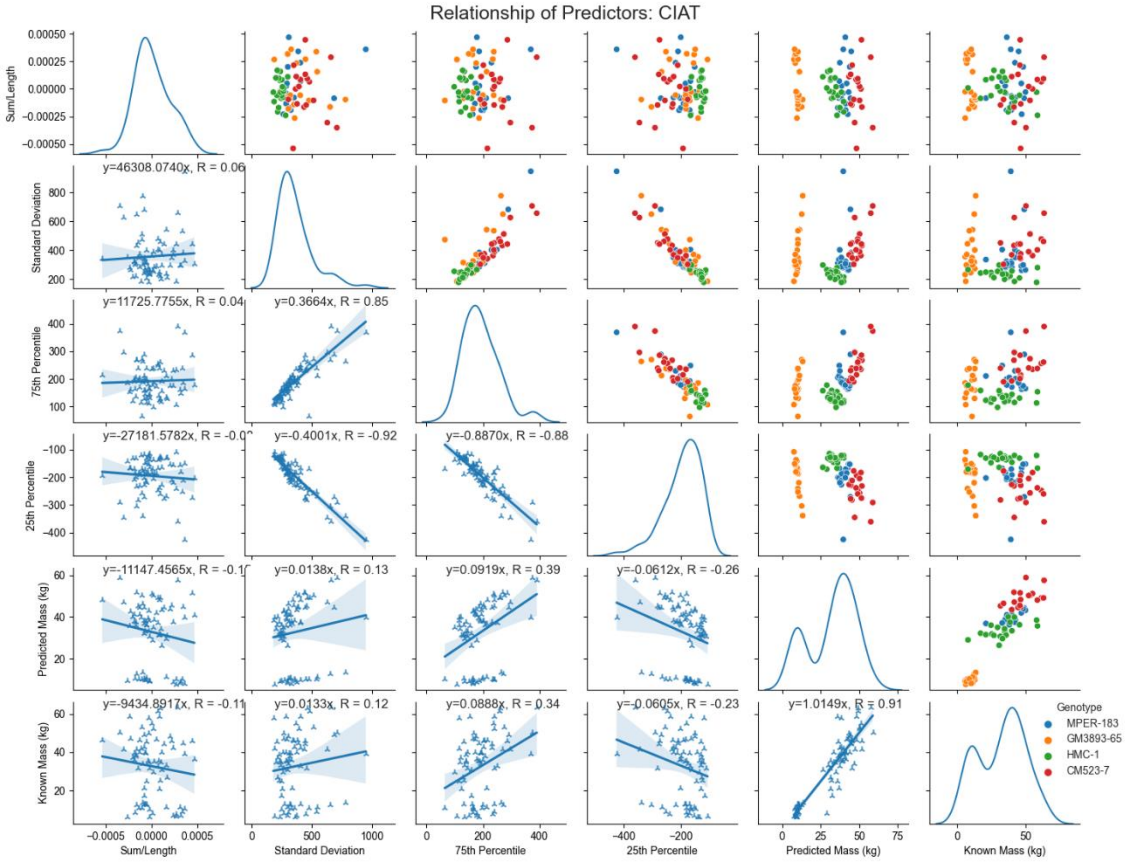

Figure 5.6 A scatter matrix showing the relationship between each feature and individual relationships to root mass for the model built from CIAT data. Correlation values are calculated by simple linear regression between the values shown. Shaded areas represent the 95% CI of the mean calculated by 10,000 iterations of bootstrap resampling. The feature values displayed here are not regularized.

overfitting, shown in Table 5.1. The upper triangle shows the same data as the lower triangle, but colorized based on genotype, which illustrates how the genotypic variance extends the range within each feature, including the root mass. It is interesting to note that Feature 1, the sum of amplitudes divided by scan length, is not strongly correlated to root mass in either model, however, models tested without it did not produce significant predictions at both locations. The diagonal shows the smoothed distribution of each feature as a density plot. The values of the features were not regularized for these figures so that the differences in amplitude would be apparent and demonstrate the need for regularization. Large differences in amplitude between predictors in penalized regression reduces the effectiveness of the penalty parameter because it will not affect each predictor equally, requiring regularization to prevent bias.

Figure 5.8 shows the distribution of each feature by location on a common axis so the features can be compared across locations. Again, this feature shows un-regularized features. Comparing the features across location is informative in diagnosing the cross-location performance of the models. It can be seen that the features have very similar distribution centers, as well as similar range, at both locations. In contrast, the distribution of observed mass is well separated by location, with very little overlap. The CIAT data tends to be more widely distributed, no doubt a function of the greater range in observed mass values.

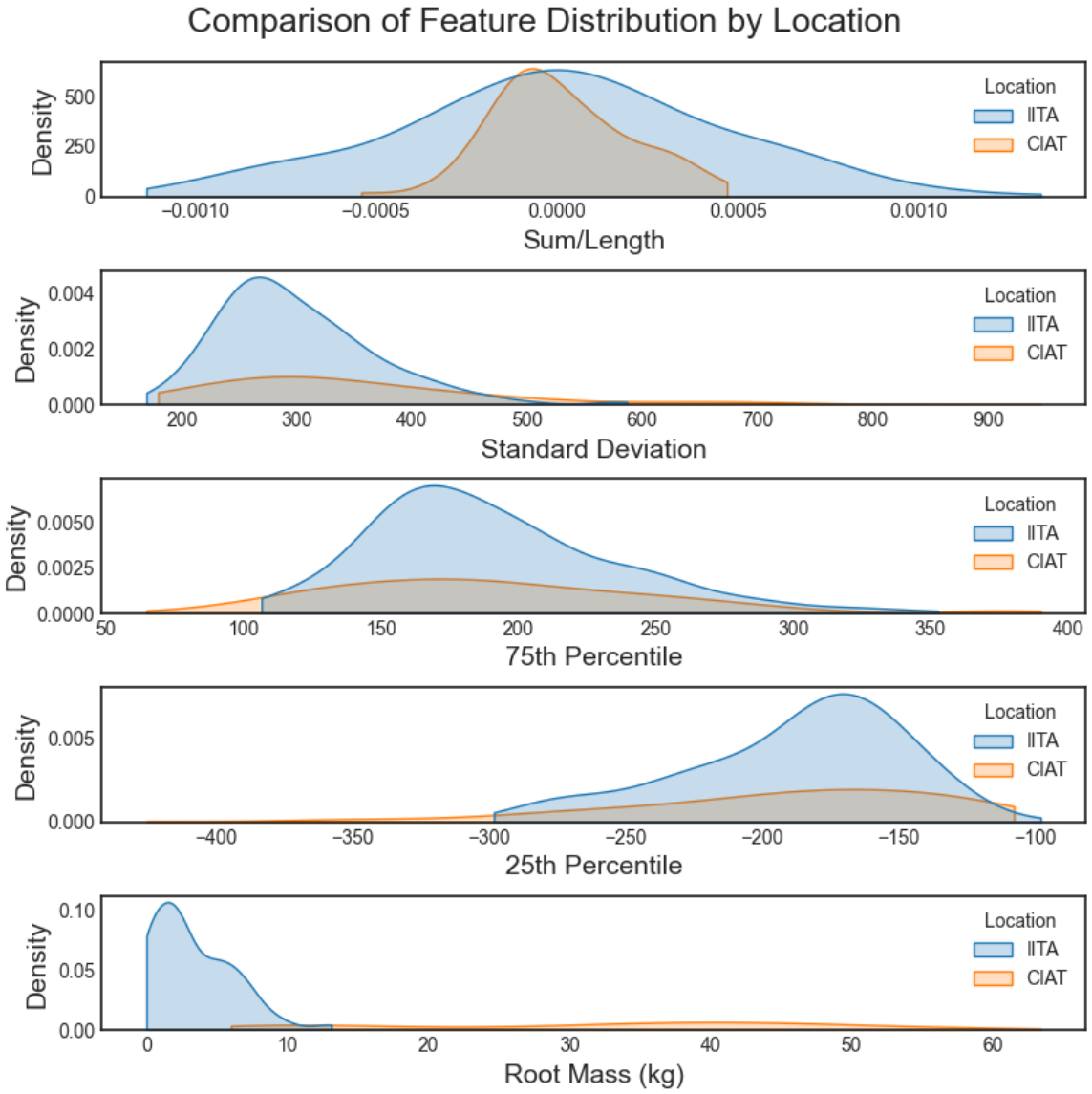

Figure 5.7 Distribution of the GPR features by location, including the observed root mass. Distribution is displayed by kernel density estimation, such that the area under each curve sums to 1. It can be easily seen that feature distributions are heavily overlapped while the mass distribution is almost entirely separate.

5.4. Discussion

The results presented here demonstrate correlation between GPR data and bulked root mass in cassava, a finding which agrees with previous findings. To date, this appears to be the first study to examine this relationship at multiple locations, and to assess the

ability of the predictive model to generalize across locations. Within locations, the CIAT model performed best, with 82% explained variation. The IITA model did not perform as well, with only 35% explained variation between predicted mass and observed mass. The distribution of root mass at each location becomes significant in this evaluation, because even though IITA had a much larger sample population, the sample unit consisted of fewer plants with lower yield, causing a marked mismatch in ground truth values between locations. Indeed, as shown in Figure 5.7 above, there is barely any overlap at all in root mass values, with IITA having a total range of 0 – 13.1 kg, and CIAT having a total range of 6.0 – 63.41 kg, with a distance between the means of 29.3 kg, a value twice the range of mass at IITA. The larger range of sample values at CIAT likely contributed to the improved model performance within the location.

This method is exceptionally dependent on successfully removing the surface reflection from the radargram before feature extraction – the surface reflection tends to be an order of magnitude brighter than the subsurface reflections. As noted above, the surface reflection at IITA displayed greater variance in relative position within the radargram, which resulted in lower quality estimations and removal of the surface reflection. This lower data quality may have contributed to the reduced model performance at that location. We should also note here that feature tuning resulted in models with improved performance for a location, but severely reduced performance at the other, indicating the feature set was not generalized. These differences are likely the result of differing soil properties, as noted in the introduction.

The cross locational performance of the models was greatly reduced from the within location performance, though the CIAT trained model still outperformed the IITA trained model when predicting the opposing location. The difference in root mass again becomes significant here, as each model is essentially attempting to extrapolate outside the training data set. As seen in Table 5.1, the extrapolation is a significant problem, such that the mean of predicted mass is closer to the mean of the training data than the test data, i.e. the CIAT model predicts a mean closer to the observed root mass at CIAT than at IITA when attempting to predict the root mass at IITA. The problem is also apparent when examining the slope of the regression lines in Figure 5.6 – the CIAT model tends to overestimate the root mass at IITA by a factor of 4, while the IITA model tends to underestimate the root mass at CIAT by the same factor. This may be an indication of overfitting, however, examination of the feature distribution at each location, as seen in Figure 5.7, demonstrates a more fundamental issue. While the distributions vary slightly, the central tendencies are very similar, despite root mass distribution barely overlapping. The resulting implication is that model predictions are location specific, and the model should be calibrated at each location where it will be applied.

The most likely explanation for this locational specificity is variation in soil conditions, such as texture, structure, and moisture content as noted in the introduction. Another possibility is the difference in sample unit of one plot row – at CIAT a single row consisted of 9 plants, while a single row at IITA was limited to 4 plants, which also helps to explain the difference in observed root mass. However, it is as yet impossible to

rule out the possible effect of genotypic variance in cassava root structure, as different genotypes were used at each location, and the effect of root orientation on GPR mass estimation has not been extensively studied [151]. Genotypic differences may exhibit as differences in root number, length, diameter, growing angle, and depth. The possibility remains that other methods, such as those reported by Delgado et al. and Dobreva et al., may be more robust to locational variance [8, 155]. It is also possible that these methods could yet be modified to better generalize. One possible solution is to normalize either the radar data or the extracted features to the location. This could be accomplished through the inclusion of soil blanks in the radar data, from which could be derived normalization characteristics for that soil environment.

5.5. Conclusion

The work presented here is significant in two important ways. First, it supports and expands on the growing body of evidence that GPR data contain information about bulked root mass, suggesting that the question is no longer whether root mass can be estimated by GPR, but rather, what is the best way to extract that information. Secondly, it begins to answer the question as to whether predictive models can be generalized across locations and soil environments. As the study is not comprehensive to all reported methods and models, it's not yet possible to conclusively answer that question. Indeed, as work in this area continues, old conclusions will need to be updated for new methods. Already previously held assumptions about GPR work in agricultural settings have been upset when Teare et al. demonstrated improved GPR performance under wet soil

conditions [178]. What has been demonstrated clearly is the possibility of building a predictive model which is valid at multiple locations, even if it requires local calibration. With up to 82% explained variation, the model performance based on the CIAT data represents one of the most precise prediction models reported. Undoubtedly, data collection methods, processing methods, and model prediction will continue to improve.

6. CONCLUSIONS

6.1. Current Status of GPR in Cassava

The application of GPR for the estimation of root mass in agricultural fields is yet in its infancy. To date, only 5 studies have been published, though several more are expected soon, including the contents of this work. While most of these publications have operated on the null hypothesis that GPR data cannot be related to root mass, the conversation is beginning to move towards how to optimize processing and models. While the body of work supporting the ability to estimate root mass is not overwhelming, the results have been consistent. What's more, at least 3 different models have been demonstrated with success [8, 155, 178].

The body of work presented here has demonstrated that agricultural GPR data contains information about bulked root mass, though at varying degrees of quality. It is clear that GPR data contain a lot of information not related to root mass – noise. The ability of a processing pipeline to reduce the noise and isolate the root information is a major driver of the prediction quality. Chapter 3 showed the effect of soil water, that increased soil water can improve the quality of GPR data, but also that extracted features will vary with water content. In Chapter 5, the cross locational validation of prediction models was possibly hampered by differences in soil water content. While water content was not measured at either CIAT or IITA, the average reflection brightness was greater at IITA than at CIAT, a difference which, based on Chapter 3 results, could be caused by a drier soil at IITA. This is in line with GPR theory as increased water content causes increased

signal attenuation, which should result in lower total energy return to the receiving antenna.

Chapter 4 confirmed the model presented in Chapter 3, called the Relative Energy Density Model, and expanded on the predictive features that could be extracted using the sliding window method. Chapter 5 confirmed that features based on the range and mean of the windowed amplitudes are descriptive of root mass. The difference in features used in Chapters 3, 4, and 5 suggests that any number of correlative features may exist, and these are just a few. While a model can be made which works at multiple locations is possible, the current methods require location specific calibration, an arduous requirement in practical applications of the technology.

6.2. Future Directions for Further Research

While several important questions have been answered here, many more remain. Perhaps foremost of the needed improvements is the need to detect root depth, or rather, the ability to identify the subset of the radargram with information about the roots, without prior knowledge of the root location or mass. We have suggested that machine learning is one possible solution. Using supervised datasets, a model could be trained to recognize the relationships of GPR features which are peculiar to root zones.

Another need is a way to account for differences in soil conditions, especially soil moisture. There are certainly many ways this could be done, though one proposed method is the use of blank control plots which can be used to standardize data to soil conditions. Another possibility exists that a model could be derived which is robust

against differences in soil conditions, and therefore does not require adjustments to account for variation across applications.

Finally, the effect of root orientation requires further study. Though Delgado et al. have reported some findings on this topic already, they did not provide a solution to account for the effect on mass estimation caused by root orientation outside of suggesting the most effective antenna polarization. Currently, we have no suggestions on how to correct for estimation error caused by root orientation. Indeed, the error has not even been quantified, though it is almost certainly significant.

The range of prediction accuracies achieved both in this work and by others indicates that no method has succeeded in accounting for all the sources of noise. As we've concluded here, likely sources include soil conditions, root orientation, and improper selection of the root zone within the radargram. Other sources may include variations in the position and orientation of the radar antenna, changes in the ground surface roughness or slope, and electromagnetic interference from outside sources. These problems remain to be solved before agricultural GPR can become a useful phenotyping and breeding tool.

REFERENCES

1. Lipper L, Thornton P, Campbell BM, et al (2014) Climate-smart agriculture for food security. Nat Clim Chang 4:1068–1072
2. FAO © (2015) FAOSTAT - Country/Territorial Notes: Yield of Commodity. 2016:Group-Production, Domain-Crops, Area-Africa
3. Ceballos H, Ramirez J, Bellotti AC, Jarvis A, Alvarez E (2011) Adaptation of Cassava to Changing Climates. In: Crop Adapt. to Clim. Chang. Wiley-Blackwell, Oxford, UK, pp 411–425
4. Bassey EE, Harry GI (2013) Screening cassava (Manihot esculenta Crantz) genotypes for tuber bulking, early maturity and optimum harvesting time in Uyo, South Eastern Nigeria. Peak J Agric Sci 1:83–88
5. Suja G, John KS, Sreekumar J, Srinivas T (2010) Short-duration cassava genotypes for crop diversification in the humid tropics: growth dynamics, biomass, yield and quality. J Sci Food Agric 90:188–198
6. Ceballos H, Iglesias CA, Pérez JC, Dixon AGO (2004) Cassava breeding: opportunities and challenges. Plant Mol Biol 56:503–516
7. El-Sharkawy MA (2004) Cassava biology and physiology. Plant Mol Biol 56:481–501
8. Delgado A, Hays DB, Bruton RK, Ceballos H, Novo A, Boi E, Selvaraj MG (2017) Ground penetrating radar: a case study for estimating root bulking rate in cassava (Manihot esculenta Crantz). Plant Methods 13:65
9. Nassar N, Ortiz R (2010) Breeding Cassava to Feed the Poor. Sci Am 302:78–84

10. Okechukwu RU, Dixon AGOO (2009) Performance of Improved Cassava Genotypes for Early Bulking, Disease Resistance, and Culinary Qualities in an Inland Valley Ecosystem. Agron J 101:1258–1265
11. Ceballos H, Hershey C, Becerra-López-Lavalle LA (2012) New Approaches to Cassava Breeding. In: Plant Breed. Rev. John Wiley & Sons, Inc., Hoboken, NJ, USA, pp 427–504
12. Zaheer K, Akhtar MH (2016) Potato Production, Usage, and Nutrition—A Review. Crit Rev Food Sci Nutr 56:711–721
13. De Haan S, Rodriguez F (2016) Potato Origin and Production. In: Adv. Potato Chem. Technol. Second Ed. Elsevier Inc., pp 1–32
14. Furrer AN, Chegeni M, Ferruzzi MG (2018) Impact of potato processing on nutrients, phytochemicals, and human health. Crit Rev Food Sci Nutr 58:146–168
15. USDA (2019) United States Department of Agriculture National Agricultural Statistics Service Crop Production Historical Track Records.
16. Camire ME, Kubow S, Donnelly DJ (2009) Potatoes and human health. Crit Rev Food Sci Nutr 49:823–840
17. Vreugdenhil D, Bradshaw J, Gebhardt C, Govers F, MacKerron DKL, Taylor MA, Ross HA (2007) Potato Biology and Biotechnology. Potato Biol Biotechnol Adv Perspect. https://doi.org/10.1016/B978-0-444-51018-1.X5040-4
18. Jol HM (2009) Ground Penetrating Radar Theory and Applications., First. Elsevier, The Boulevard, Landford Lane, Kidlington, Oxford OX5 1GB, UK
19. Baker GS, Jordan TE, Pardy J (2007) An introduction to ground penetrating radar

(GPR). In: Spec. Pap. 432 Stratigr. Anal. Using GPR. Geological Society of America, pp 1–18

20. Utsi EC (2017) Ground Penetrating Radar : Theory and Practice. Elsevier Science, Oxford, UK

21. Everett ME (2013) Near-Surface Applied Geophysics. Near-Surface Appl Geophys. https://doi.org/10.1017/CBO9781139088435

22. Annan AP (2002) GPR - History, Trends, and Future Developments. Subsurf Sens Technol Appl 3:253–270

23. Hubbard S, Chen J, Williams K, Peterson J, Rubin Y (2005) Environmental and agricultural applications of GPR. In: Proc. 3rd Int. Work. Adv. Gr. Penetrating Radar. pp 45–49

24. Zhu S, Huang C, Su Y, Sato M (2014) 3D Ground Penetrating Radar to Detect Tree Roots and Estimate Root Biomass in the Field. Remote Sens 6:5754–5773

25. Lal R (2004) Soil carbon sequestration to mitigate climate change. Geoderma 123:1–22

26. Lynch JP (2007) Roots of the Second Green Revolution. Aust J Bot 55:493

27. Atkinson JA, Pound MP, Bennett MJ, Wells DM (2019) Uncovering the hidden half of plants using new advances in root phenotyping. Curr Opin Biotechnol 55:1–8

28. Kuijken RCP, Van Eeuwijk FA, Marcelis LFM, Bouwmeester HJ (2015) Root phenotyping: From component trait in the lab to breeding. J Exp Bot 66:5389–5401

29. Paez-Garcia A, Motes C, Scheible W-R, Chen R, Blancaflor E, Monteros M (2015) Root Traits and Phenotyping Strategies for Plant Improvement. Plants 4:334–355

30. Ehosioke S, Nguyen F, Rao S, Kremer T, Placencia-Gomez E, Huisman JA, Kemna A, Javaux M, Garré S (2020) Sensing the electrical properties of roots: A review. Vadose Zo J 19:e20082

31. Hund A, Trachsel S, Stamp P (2009) Growth of axile and lateral roots of maize: I development of a phenotying platform. Plant Soil 325:335–349

32. Bonser AM, Lynch J, Snapp S (1996) Effect of phosphorus deficiency on growth angle of basal roots in Phaseolus vulgaris. New Phytol 132:281–288

33. Falk KG, Jubery TZ, Mirnezami S V., Parmley KA, Sarkar S, Singh A, Ganapathysubramanian B, Singh AK (2020) Computer vision and machine learning enabled soybean root phenotyping pipeline. Plant Methods 16:5

34. French A, Ubeda-Tomás S, Holman TJ, Bennett MJ, Pridmore T (2009) High-Throughput Quantification of Root Growth Using a Novel Image-Analysis Tool. Plant Physiol 150:1784–1795

35. Bengough AG, Gordon DC, Al-Menaie H, Ellis RP, Allan D, Keith R, Thomas WTB, Forster BP (2004) Gel observation chamber for rapid screening of root traits in cereal seedlings. Plant Soil 262:63–70

36. Nagel KA, Lenz H, Kastenholz B, et al (2020) The platform GrowScreen-Agar enables identification of phenotypic diversity in root and shoot growth traits of agar grown plants. Plant Methods 16:89

37. Futsaether CM, Oxaal U (2002) A growth chamber for idealized studies of seedling root growth dynamics and structure. Plant Soil 246:221–230

38. Fang S, Yan X, Liao H (2009) 3D reconstruction and dynamic modeling of root architecture in situ and its application to crop phosphorus research. Plant J 60:1096–1108

39. Clark RT, MacCurdy RB, Jung JK, Shaff JE, McCouch SR, Aneshansley DJ, Kochian L V. (2011) Three-Dimensional Root Phenotyping with a Novel Imaging and Software Platform. Plant Physiol 156:455–465

40. Hargreaves CE, Gregory PJ, Bengough AG (2009) Measuring root traits in barley (Hordeum vulgare ssp. vulgare and ssp. spontaneum) seedlings using gel chambers, soil sacs and X-ray microtomography. Plant Soil 316:285–297

41. Gregory PJ, Bengough AG, Grinev D, Schmidt S, Thomas W (Bill) TB, Wojciechowski T, Young IM (2009) Root phenomics of crops: opportunities and challenges. Funct Plant Biol 36:922

42. Gruber BD, Giehl RFH, Friedel S, von Wirén N (2013) Plasticity of the Arabidopsis Root System under Nutrient Deficiencies. Plant Physiol 163:161–179

43. Passioura JB (2006) Viewpoint: The perils of pot experiments. Funct Plant Biol 33:1075–1079

44. Liao H, Rubio G, Yan X, Cao A, Brown KM, Lynch JP (2002) Effect of phosphorus availability on basal root shallowness in common bean. In: Interact. Root Environ. An Integr. Approach. Springer Netherlands, pp 69–79

45. Richard CAI, Hickey LT, Fletcher S, Jennings R, Chenu K, Christopher JT (2015)

High-throughput phenotyping of seminal root traits in wheat. Plant Methods 11:13

46. Manschadi AM, Hammer GL, Christopher JT, DeVoil P (2008) Genotypic variation in seedling root architectural traits and implications for drought adaptation in wheat (Triticum aestivum L.). Plant Soil 303:115–129

47. Maeght J-L, Rewald B, Pierret A (2013) How to study deep roots—and why it matters. Front Plant Sci 4:299

48. van Noordwijk M, Brouwer G, Meijboom F, do Rosário G. Oliveira M, Bengough AG (2001) Trench Profile Techniques and Core Break Methods. In: Root Methods. Springer Berlin Heidelberg, pp 211–233

49. do Rosário G. Oliveira M, van Noordwijk M, Gaze SR, Brouwer G, Bona S, Mosca G, Hairiah K (2000) Auger Sampling, Ingrowth Cores and Pinboard Methods. In: Root Methods. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 175–210

50. Vepraskas MJ, Hoyt GD (1988) Comparison of the Trench-Profile and Core Methods for Evaluating Root Distributions in Tillage Studies. Agron J 80:166–172

51. Nemoto H, Suga R, Ishihara M, Okutsu Y (1998) Deep Rooted Rice Varieties Detected through the Observation of Root Characteristics Using the Trench Method. Ikushugaku zasshi 48:321–324

52. de Azevedo MCB, Chopart JL, de Conti Medina C (2011) Sugarcane root length density and distribution from root intersection counting on a trench-profile. Sci Agric 68:94–101

53. Hairiah K, Van Noordwijk M, Setijono S (1991) Tolerance to acid soil conditions of the velvet beans Mucuna pruriens var. utilis and M. deeringiana - I. Root development. Plant Soil 134:95–105

54. Floris J, Van Noordwijk M (1984) Improved methods for the extraction of soil samples for root research. Plant Soil 77:369–372

55. Brito JM, Jankiewicz LS, Orduña VMF, Escobar FC (1986) The root system of the husk tomato (Physalis ixocarpa Brot.). Acta Agrobot 39:367–383

56. Aerts R, Berendse F, Klerk NM, Bakker C (1989) Root production and root turnover in two dominant species of wet heathlands. Oecologia 81:374–378

57. Burridge JD, Black CK, Nord EA, Postma JA, Sidhu JS, York LM, Lynch JP (2020) An Analysis of Soil Coring Strategies to Estimate Root Depth in Maize (Zea mays) and Common Bean (Phaseolus vulgaris). Plant Phenomics 2020:1–20

58. Ogilvie CM, Ashiq W, Vasava HB, Biswas A (2021) Quantifying Root-Soil Interactions in Cover Crop Systems: A Review. Agriculture 11:218

59. KONO Y, YAMAUCHI A, NONOYAMA T, TATSUMI J, KAWAMURA N (1987) A revised experimental system of root-soil interaction for laboratory work. Environ Control Biol 25:141–151

60. Delory BM, Weidlich EWA, van Duijnen R, Pagès L, Temperton VM (2018) Measuring plant root traits under controlled and field conditions: Step-by-step procedures. In: Methods Mol. Biol. Humana Press Inc., pp 3–22

61. Kano-Nakata M, Inukai Y, Wade LJ, Siopongco JDLC, Yamauchi A (2011) Root Development, Water Uptake, and Shoot Dry Matter Production under Water

Deficit Conditions in Two CSSLs of Rice: Functional Roles of Root Plasticity. Plant Prod Sci 14:307–317

62. Thangthong N, Jogloy S, Jongrungklang N, Kvien CK, Pensuk V, Kesmala T, Vorasoot N (2018) Root distribution patterns of peanut genotypes with different drought resistance levels under early-season drought stress. J Agron Crop Sci 204:111–122

63. Miyazaki A, Arita N (2020) Deep rooting development and growth in upland rice NERICA induced by subsurface irrigation. Plant Prod Sci 23:211–219

64. Smit AL, George E, Groenwold J (2000) Root Observations and Measurements at (Transparent) Interfaces with Soil. In: Root Methods. Springer Berlin Heidelberg, pp 235–271

65. Taylor HM, Upchurch DR, McMichael BL (1990) Applications and limitations of rhizotrons and minirhizotrons for root studies. Plant Soil 129:29–35

66. Eberbach PL, Hoffmann J, Moroni SJ, Wade LJ, Weston LA (2013) Rhizo-lysimetry: Facilities for the simultaneous study of root behaviour and resource use by agricultural crop and pasture systems. Plant Methods 9:3

67. Van de Geijn SC, Vos J, Groenwold J, Goudriaan J, Leffelaar PA (1994) The Wageningen Rhizolab — a facility to study soil-root-shoot-atmosphere interactions in crops. Plant Soil 161:275–287

68. Chen S, van der Graaff E, Ytting NK, Thorup-Kristensen K (2019) Evaluation of deep root phenotyping techniques in tube rhizotrons. Acta Agric Scand Sect B — Soil Plant Sci 69:62–74

69. Ytting NK, Kirkegaard JA, Thorup-Kristensen K (2019) Optimizing root measurements in rhizotrons. bioRxiv. https://doi.org/10.1101/587329

70. Martins SM, de Brito GG, Gonçalves W da C, Tripode BMD, Lartaud M, Duarte JB, Morello C de L, Giband M (2020) Phenoroots: An inexpensive non-invasive phenotyping system to assess the variability of the root system architecture. Sci Agric 77:2020

71. Huo C, Cheng W (2019) Improved root turnover assessment using field scanning rhizotrons with branch order analysis. Ecosphere 10:e02793

72. Mohamed A, Monnier Y, Mao Z, Lobet G, Maeght JL, Ramel M, Stokes A (2017) An evaluation of inexpensive methods for root image acquisition when using rhizotrons. Plant Methods 13:11

73. Boldt-Burisch K, Naeth MA (2017) Mycorrhization affects root distribution of Lotus corniculatus and Calamagrostis epigeios in a nutrient poor heterogeneous soil in a rhizotron experiment. Rhizosphere 4:36–47

74. Rasmussen CR, Thorup-Kristensen K, Dresbøll DB (2020) Uptake of subsoil water below 2 m fails to alleviate drought response in deep-rooted Chicory (Cichorium intybus L.). Plant Soil 446:275–290

75. Montagnoli A, Baronti S, Alberto D, Chiatante D, Scippa GS, Terzaghi M (2021) Pioneer and fibrous root seasonal dynamics of Vitis vinifera L. are affected by biochar application to a low fertility soil: A rhizobox approach. Sci Total Environ 751:141455

76. Haarhoff SJ, Lötze E, Swanepoel PA (2021) Rainfed maize root morphology in

response to plant population under no-tillage. Agron J 113:75–87

77. Taylor BN, Beidler K V., Cooper ER, Strand AE, Pritchard SG (2013) Sampling volume in root studies: the pitfalls of under-sampling exposed using accumulation curves. Ecol Lett 16:862–869

78. Bragg PL, Govi G, Cannell RQ (1983) A comparison of methods, including angled and vertical minirhizotrons, for studying root growth and distribution in a spring oat crop. Plant Soil 73:435–440

79. Bodner G, Nakhforoosh A, Arnold T, Leitner D (2018) Hyperspectral imaging: a novel approach for plant root phenotyping. Plant Methods 14:84

80. Rahman G, Sohag H, Chowdhury R, Wahid KA, Dinh A, Arcand M, Vail S (2020) SoilCam: A Fully Automated Minirhizotron using Multispectral Imaging for Root Activity Monitoring. Sensors 20:787

81. Defrenne CE, Childs J, Fernandez CW, Taggart M, Nettles WR, Allen MF, Hanson PJ, Iversen CM (2021) High-resolution minirhizotrons advance our understanding of root-fungal dynamics in an experimentally warmed peatland. Plants People Planet 3:640–652

82. Xu W, Yu G, Zare A, Zurweller B, Rowland DL, Reyes-Cabrera J, Fritschi FB, Matamala R, Juenger TE (2020) Overcoming small minirhizotron datasets using transfer learning. Comput Electron Agric 175:105466

83. Yu G, Zare A, Sheng H, Matamala R, Reyes-Cabrera J, Fritschi FB, Juenger TE (2020) Root identification in minirhizotron imagery with multiple instance learning. Mach Vis Appl 31:1–13

84. Wang T, Rostamza M, Song Z, Wang L, McNickle G, Iyer-Pascuzzi AS, Qiu Z, Jin J (2019) SegRoot: A high throughput segmentation method for root image analysis. Comput Electron Agric 162:845–854

85. Crocker TL, Hendrick RL, Ruess RW, Pregitzer KS, Burton AJ, Allen MF, Shan J, Morris LA (2003) Substituting root numbers for length: Improving the use of minirhizotrons to study fine root dynamics. Appl Soil Ecol 23:127–135

86. Lee CG, Suzuki S, Noguchi K, Inubushi K (2016) Estimation of fine root biomass using a minirhizotron technique among three vegetation types in a cool-temperate brackish marsh. Soil Sci Plant Nutr 62:465–470

87. Herrera JM, Büchi L, Rubio G, Torres-Guerrero C, Wendling M, Stamp P, Pellet D (2017) Root decomposition at high and low N supply throughout a crop rotation. Eur J Agron 84:105–112

88. Postic F, Beauchêne K, Gouache D, Doussan C (2019) Scanner-Based Minirhizotrons Help to Highlight Relations between Deep Roots and Yield in Various Wheat Cultivars under Combined Water and Nitrogen Deficit Conditions. Agronomy 9:297

89. McBRATNEY AB, WEBSTER R (1983) Optimal interpolation and isarithmic mapping of soil properties. J Soil Sci 34:137–162

90. Fiorini A, Boselli R, Amaducci S, Tabaglio V (2018) Effects of no-till on root architecture and root-soil interactions in a three-year crop rotation. Eur J Agron 99:156–166

91. Baker CJ, Saxton KE, Ritchie WR, Chamen WCT, Reicosky DC, Ribeiro MFS,

Justice SE, Hobbs PR (2006) No-tillage seeding in conservation agriculture: Second edition. No-Tillage Seeding Conserv Agric Second Ed. https://doi.org/10.1079/9781845931162.0000

92. Muhandiram NPK, Humphreys MW, Fychan R, Davies JW, Sanderson R, Marley CL (2020) Do agricultural grasses bred for improved root systems provide resilience to machinery-derived soil compaction? Food Energy Secur 9:e227

93. Boehm W (1974) Mini-Rhizotrons for root observations under field conditions. Zeitschrift fuer Acker- und Pflanzenbau 140:282–287

94. Bennie ATP, Taylor HM, Georgen PG (1987) An assessment of the core-break method for estimating rooting density of different crops in the field. Soil Tillage Res 9:347–353

95. Kücke M, Schmid H, Spiess A (1995) A comparison of four methods for measuring roots of field crops in three contrasting soils. Plant Soil 172:63–71

96. Wasson AP, Chiu GS, Zwart AB, Binns TR (2017) Differentiating Wheat Genotypes by Bayesian Hierarchical Nonlinear Mixed Modeling of Wheat Root Density. Front Plant Sci 8:282

97. Li X, Ingvordsen CH, Weiss M, Rebetzke GJ, Condon AG, James RA, Richards RA (2019) Deeper roots associated with cooler canopies, higher normalized difference vegetation index, and greater yield in three wheat populations grown on stored soil water. J Exp Bot 70:4963–4974

98. Bai C, Ge Y, Ashton RW, et al (2019) The relationships between seedling root screens, root growth in the field and grain yield for wheat. Plant Soil 440:311–326

99. Steingrobe B, Schmid H, Claassen N (2000) The use of the ingrowth core method for measuring root production of arable crops — influence of soil conditions inside the ingrowth core on root growth. J Plant Nutr Soil Sci 163:617–622

100. Chen S, Lin S, Reinsch T, Loges R, Hasler M, Taube F (2016) Comparison of ingrowth core and sequential soil core methods for estimating belowground net primary production in grass-clover swards. Grass Forage Sci 71:515–528

101. Reinsch T, Struck IJA, Loges R, Kluß C, Taube F (2021) Soil carbon dynamics of no-till silage maize in ley systems. Soil Tillage Res 209:104957

102. Lei C, Abraha M, Chen J, Su Y-J (2021) Long-term variability of root production in bioenergy crops from ingrowth core measurements. J Plant Ecol 14:757–770

103. Trachsel S, Kaeppler SM, Brown KM, Lynch JP (2011) Shovelomics: High throughput phenotyping of maize (Zea mays L.) root architecture in the field. Plant Soil 341:75–87

104. Bucksch A, Burridge J, York LM, Das A, Nord E, Weitz JS, Lynch JP (2014) Image-Based High-Throughput Field Phenotyping of Crop Roots. PLANT Physiol 166:470–486

105. Colombi T, Kirchgessner N, Le Marié CA, York LM, Lynch JP, Hund A (2015) Next generation shovelomics: set up a tent and REST. Plant Soil 388:1–20

106. Abiven S, Hund A, Martinsen V, Cornelissen G (2015) Biochar amendment increases maize root surface areas and branching: a shovelomics study in Zambia. Plant Soil 395:45–55

107. Arifuzzaman M, Oladzadabbasabadi A, McClean P, Rahman M (2019)

Shovelomics for phenotyping root architectural traits of rapeseed/canola (Brassica napus L.) and genome-wide association mapping. Mol Genet Genomics 294:985–1000

108. Kengkanna J, Jakaew P, Amawan S, Busener N, Bucksch A, Saengwilai P (2019) Phenotypic variation of cassava root traits and their responses to drought. Appl Plant Sci 7:e01238

109. Le Marié CA, York LM, Strigens A, Malosetti M, Camp KH, Giuliani S, Lynch JP, Hund A (2019) Shovelomics root traits assessed on the EURoot maize panel are highly heritable across environments but show low genotype-by-nitrogen interaction. Euphytica 215:1–22

110. Chloupek O (1972) The relationship between electric capacitance and some other parameters of plant roots. Biol Plant 14:227–230

111. Chloupek O (1977) Evaluation of the size of a plant's root system using its electrical capacitance. Plant Soil 48:525–532

112. Kendall WA, Pederson GA, Hill RR (1982) Root size estimates of red clover and alfalfa based on electrical capacitance and root diameter measurements. Grass Forage Sci 37:253–256

113. Cseresnyés I, Vozáry E, Rajkai K (2020) Does electrical capacitance represent roots in the soil? Acta Physiol Plant 42:71

114. Dalton FN (1995) In-situ root extent measurements by electrical capacitance methods. Plant Soil 173:157–165

115. Beem J, Smith ME, Zobel RW (1998) Estimating Root Mass in Maize Using a

Portable Capacitance Meter. Agron J 90:566–570

116. Psarras G, Merwin IA (2000) Water stress affects rhizosphere respiration rates and root morphology of young “Mutsu” apple trees on M.9 and MM.111 rootstocks. J Am Soc Hortic Sci 125:588–595

117. Ozier-Lafontaine H, Bajazet T (2005) Analysis of Root Growth by Impedance Spectroscopy (EIS). Plant Soil 277:299–313

118. Ellis T, Murray W, Kavalieris L (2013) Electrical capacitance of bean (Vicia faba) root systems was related to tissue density-a test for the Dalton Model. Plant Soil 366:575–584

119. McBride R, Candido M, Ferguson J (2008) Estimating root mass in maize genotypes using the electrical capacitance method. Arch Agron Soil Sci 54:215–226

120. Chloupek O, Ehrenbergerová J, Opitz Von Boberfeld W, Říha P (2003) Selection of white clover (Trifolium repens L.) using root traits related to dinitrogen fixation. F Crop Res 80:57–62

121. Dietrich RC, Bengough AG, Jones HG, White PJ (2013) Can root electrical capacitance be used to predict root mass in soil? Ann Bot 112:457–464

122. Aulen M, Shipley B (2012) Non-destructive estimation of root mass using electrical capacitance on ten herbaceous species. Plant Soil 355:41–49

123. Rajkai K, Végh KR, Nacsa T (2004) Electrical Capacitance as the Indicator of Root Size and Activity. Agrokémia és Talajt 51:89–98

124. Cseresnyés I, Rajkai K, Szitár K, Radimszky L, Ónodi G, Kröel-Dulay G (2020)

Root capacitance measurements allow non-intrusive in-situ monitoring of the seasonal dynamics and drought response of root activity in two grassland species. Plant Soil 449:423–437

125. Chloupek O, Forster BP, Thomas WTB (2006) The effect of semi-dwarf genes on root system size in field-grown barley. Theor Appl Genet 112:779–786

126. Wu W, Duncan RW, Ma B-L (2017) Quantification of canola root morphological traits under heat and drought stresses with electrical measurements. Plant Soil 415:229–244

127. Cseresnyés I, Szitár K, Rajkai K, Füzy A, Mikó P, Kovács R, Takács T (2018) Application of Electrical Capacitance Method for Prediction of Plant Root Mass and Activity in Field-Grown Crops. Front Plant Sci 9:93

128. Ellis TW, Murray W, Paul K, Kavalieris L, Brophy J, Williams C, Maass M (2013) Electrical capacitance as a rapid and non-invasive indicator of root length. Tree Physiol 33:3–17

129. Furman A, Arnon-Zur A, Assouline S (2015) Electrical Resistivity Tomography of the Root Zone. In: Anderson SH, Hopmans JW (eds) Soil–Water–Root Process. Adv. Tomogr. Imaging. Soil Science Society of America, Madison, WI, USA, pp 223–245

130. Rao S, Meunier F, Ehosioke S, Lesparre N, Kemna A, Nguyen F, Garré S, Javaux M (2019) Impact of Maize Roots on Soil–Root Electrical Conductivity: A Simulation Study. Vadose Zo J 18:190037

131. Amato M, Basso B, Celano G, Bitella G, Morelli G, Rossi R (2008) In situ

detection of tree root distribution and biomass by multi-electrode resistivity imaging. Tree Physiol 28:1441–1448

132. Amato M, Bitella G, Rossi R, Gómez JA, Lovelli S, Gomes JJF (2009) Multi-electrode 3D resistivity imaging of alfalfa root zone. Eur J Agron 31:213–222

133. Corona-Lopez DDJ, Sommer S, Rolfe SA, Podd F, Grieve BD (2019) Electrical impedance tomography as a tool for phenotyping plant roots. Plant Methods 15:49

134. Rao S, Lesparre N, Flores-Orozco A, Wagner F, Kemna A, Javaux M (2020) Imaging plant responses to water deficit using electrical resistivity tomography. Plant Soil 454:261–281

135. Heeraman DA, Hopmans JW, Clausnitzer V (1997) Three dimensional imaging of plant roots in situ with X-ray computed tomography. Plant Soil 189:167–179

136. Perret JS, Al-Belushi ME, Deadman M (2007) Non-destructive visualization and quantification of roots using computed tomography. Soil Biol Biochem 39:391–399

137. Mooney SJ, Pridmore TP, Helliwell J, Bennett MJ (2012) Developing X-ray computed tomography to non-invasively image 3-D root systems architecture in soil. Plant Soil 352:1–22

138. Gao W, Schlüter S, Blaser SRGA, Shen J, Vetterlein D (2019) A shape-based method for automatic and rapid segmentation of roots in soil from X-ray computed tomography images: Rootine. Plant Soil 441:643–655

139. Pfeifer J, Kirchgessner N, Colombi T, Walter A (2015) Rapid phenotyping of crop root systems in undisturbed field soils using X-ray computed tomography.

Plant Methods 11:41

140. Mairhofer S, Zappala S, Tracy S, Sturrock C, Bennett MJ, Mooney SJ, Pridmore TP (2013) Recovering complete plant root system architectures from soil via X-ray μ-Computed Tomography. Plant Methods 9:8

141. Koebernick N, Daly KR, Keyes SD, et al (2017) High-resolution synchrotron imaging shows that root hairs influence rhizosphere soil structure formation. New Phytol 216:124–135

142. Helliwell JR, Sturrock CJ, Miller AJ, Whalley WR, Mooney SJ (2019) The role of plant species and soil condition in the structural development of the rhizosphere. Plant Cell Environ 42:1974–1986

143. Kirk GJD, Boghi A, Affholder M, Keyes SD, Heppell J, Roose T (2019) Soil carbon dioxide venting through rice roots. Plant Cell Environ 42:3197–3207

144. Pohost GM, Elgavish GA, Evanochko WT (1986) Nuclear magnetic resonance imaging: With or without nuclear? J Am Coll Cardiol 7:709–710

145. Asseng S, Aylmore LAG, MacFall JS, Hopmans JW, Gregory PJ (2000) Computer-Assisted Tomography and Magnetic Resonance Imaging. In: Root Methods. Springer Berlin Heidelberg, pp 343–363

146. Metzner R, Eggert A, van Dusschoten D, Pflugfelder D, Gerth S, Schurr U, Uhlmann N, Jahnke S (2015) Direct comparison of MRI and X-ray CT technologies for 3D imaging of root systems in soil: Potential and challenges for root trait quantification. Plant Methods 11:17

147. Rogers HH, Bottomley PA (1987) In Situ Nuclear Magnetic Resonance Imaging

of Roots: Influence of Soil Type, Ferromagnetic Particle Content, and Soil Water 1. Agron J 79:957–965

148. Bagnall GC, Koonjoo N, Altobelli SA, et al (2020) Low-field magnetic resonance imaging of roots in intact clayey and silty soils. Geoderma 370:114356

149. van Dusschoten D, Metzner R, Kochs J, Postma JA, Pflugfelder D, Bühler J, Schurr U, Jahnke S (2016) Quantitative 3D Analysis of Plant Roots Growing in Soil Using Magnetic Resonance Imaging. Plant Physiol 170:1176–1188

150. Perelman A, Lazarovitch N, Vanderborght J, Pohlmeier A (2020) Quantitative imaging of sodium concentrations in soil-root systems using magnetic resonance imaging (MRI). Plant Soil 454:171–185

151. Delgado A, Novo A, Hays DB (2019) Data Acquisition Methodologies Utilizing Ground Penetrating Radar for Cassava (Manihot esculenta Crantz) Root Architecture. Geosciences 9:171

152. Butnor JR, Doolittle JA, Kress L, Cohen S, Johnsen KH (2001) Use of ground-penetrating radar to study tree roots in the southeastern United States. Tree Physiol 21:1269–1278

153. M. Konstantinovic, S. Woeckel, P. Schulze Lammers, J. Sachs (2008) UWB Radar System for Yield Monitoring of Sugar Beet. Trans ASABE 51:753–761

154. Liu X, Dong X, Xue Q, Leskovar DI, Jifon J, Butnor JR, Marek T (2018) Ground penetrating radar (GPR) detects fine roots of agricultural crops in the field. Plant Soil 423:517–531

155. Dobreva ID, Ruiz-Guzman HA, Barrios-Perez I, Adams T, Teare BL, Payton P,

Everett ME, Burow MD, Hays DB (2021) Thresholding Analysis and Feature Extraction from 3D Ground Penetrating Radar Data for Noninvasive Assessment of Peanut Yield. Remote Sens 13:1896

156. Lynch JP (2013) Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. Ann Bot 112:347–357

157. Kamau J, Melis R, Laing M, et al (2011) Farmers ' participatory selection for early bulking cassava genotypes in semi-arid Eastern Kenya. J Agron Crop Sci. https://doi.org/10.5897/JPBCS.9000053

158. Durrence JS, Perry CD, Vellidis G, Thomas DL, Kvien CK (2015) Mapping Peanut Yield Variability with an Experimental Load Cell Yield Monitoring System. 1131–1140

159. Malay WJ, Gordon RJ, Madani A, Patterson GT, Esau C, Eaton LJ, LeBlanc P (2000) Root Crop Yield Monitor Evaluation for Carrots. In: 2000 ASAE Annu. Intenational Meet. Tech. Pap. Eng. Solut. a New Century. American Society of Agricultural Engineers, pp 3039–3050

160. Davenport JR, Redulla CA, Hattendorf MJ, Evans RG, Boydston RA (2002) Potato Yield Monitoring on Commercial Fields. Horttechnology 12:289–296

161. Miodrag Konstantinovic, Sebastian Woeckel, Peter Schulze Lammers, Juergen Sachs, Konstantinovic M, Woeckel S, Lammers PS, Sachs J (2007) Evaluation of a UWB radar system for yield mapping of sugar beet. Am Soc Agric Biol Eng Meet Present 71051:1–11

162. Liu X, Dong X, Leskovar DI, Xue Q, Marek T, Jifon J, Butnor JR (2018) Ground

penetrating radar (GPR) detects fine roots of agricultural crops in the field. Plant Soil 517–531

163. Nuzzo L, Alli G, Guidi R, Cortesi N, Sarri A, Manacorda G (2014) A new densely-sampled Ground Penetrating Radar array for landmine detection. In: Proc. 15th Int. Conf. Gr. Penetrating Radar. IEEE, pp 969–974

164. Montoya TP, Smith GS (1996) Vee dipoles with resistive loading for short-pulse ground-penetrating radar. Microw Opt Technol Lett 13:132–137

165. Kangwook Kim, Scott WR (2005) Design of a resistively loaded vee dipole for ultrawide-band ground-penetrating Radar applications. IEEE Trans Antennas Propag 53:2525–2532

166. Topp GC, Davis JL, Annan AP (1980) Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. Water Resour Res 16:574–582

167. Smit AL, Bengough AG, Engels C, van Noordwijk M, Pellerin S, van de Geijn SC (2000) Root Methods. https://doi.org/10.1007/978-3-662-04188-8

168. Li S, Cui Y, Zhou Y, Luo Z, Liu J, Zhao M (2017) The industrial applications of cassava: current status, opportunities and prospects. J Sci Food Agric 97:2282–2290

169. Wossen T, Alene A, Abdoulaye T, Feleke S, Rabbi IY, Manyong V (2019) Poverty Reduction Effects of Agricultural Technology Adoption: The Case of Improved Cassava Varieties in Nigeria. J Agric Econ 70:392–407

170. Ceballos H, Rojanaridpiched C, Phumichai C (2020) Excellence in Cassava

Breeding: Perspectives for the Future. Crop Breeding, Genet Genomics 1–31

171. Qiu R, Wei S, Zhang M, Li H, Sun H, Liu G, Li M (2018) Sensors for measuring plant phenotyping: A review. Int J Agric Biol Eng 11:1–17

172. Reynolds M, Chapman S, Crespo-Herrera L, et al (2020) Breeder friendly phenotyping. Plant Sci 295:110396

173. Chawade A, van Ham J, Blomquist H, Bagge O, Alexandersson E, Ortiz R (2019) High-Throughput Field-Phenotyping Tools for Plant Breeding and Precision Agriculture. Agronomy 9:258

174. Wei MCF, Maldaner LF, Ottoni PMN, Molin JP (2020) Carrot Yield Mapping: A Precision Agriculture Approach Based on Machine Learning. AI 2020, Vol 1, Pages 229-241 1:229–241

175. Al-Gaadi KA, Hassaballa AA, Tola E, Kayad AG, Madugundu R, Alblewi B, Assiri F (2016) Prediction of Potato Crop Yield Using Precision Agriculture Techniques. PLoS One 11:e0162219

176. Selvaraj MG, Valderrama M, Guzman D, Valencia M, Ruiz H, Acharjee A (2020) Machine learning for high-throughput field phenotyping and image processing provides insight into the association of above and below-ground traits in cassava (Manihot esculenta Crantz). Plant Methods 16:87

177. Annan AP (2003) Ground Penetrating Radar Principles, Procedures & Applications.

178. Teare BL, Ruiz H, Agbona A, et al (2021) Effect of Soil Water on GPR Estimation of Bulked Roots, Methods, and Suggestions. Plant Methods -

PREPRINT. https://doi.org/10.21203/rs.3.rs-907807/v1

179. Chanzy A, Tarussov A, Bonn F, Judge A (1996) Soil Water Content Determination Using a Digital Ground-Penetrating Radar. Soil Sci Soc Am J 60:1318–1326

180. Lambot S, Weihermüller L, Huisman JA, Vereecken H, Vanclooster M, Slob EC (2006) Analysis of air-launched ground-penetrating radar techniques to measure the soil surface water content. Water Resour Res 42:1–12

181. Miller TW, Borchers B, Hendrickx JMH, Hong S, Dekker LW, Ritsema CJ (2002) Effects of soil physical properties on GPR for landmine detection. Fifth Int. Symp. Technol. Mine Probl.

182. Lane HM, Murray SC (2021) High Throughput can produce better decisions than high accuracy when phenotyping plant populations. Crop Sci. https://doi.org/10.1002/csc2.20514

183. Mtunguja MK, Beckles DM, Laswai HS, Ndunguru JC, Sinha NJ (2019) Opportunities to commercialize cassava production for poverty alleviation and improved food security in Tanzania. African J Food, Agric Nutr Dev 19:13928–13946

184. Miller TW, Borchers B, Hendrickx JMH, Hong S-H, Lensen HA, Schwering PBW, Rhebergen JB (2002) Effect of soil moisture on land mine detection using ground penetrating radar. In: Broach JT, Harmon RS, Dobeck GJ (eds) Detect. Remediat. Technol. Mines Minelike Targets VII. pp 281–290