European Plant Phenotyping Network

EPPN 2020

D1.1 3D CAD Designs and Simulated Images

Tony Pridmore (UNOTT), Sixtine Passot (UCL), Xavier Draye (UCL)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731013. This publication reflects only the view of the author, and the European Commission cannot be held responsible for any use which may be made of the information contained therein.

Document information

EU Project N°	731013	**Acronym**	EPPN2020
Full Title	European Plant Phenotyping Network 2020		
Project website	www.eppn2020.plant-phenotyping.eu		

Deliverable	**N°**	D1.1	**Title**	3D CAD designs and simulated images
Work Package	**N°**	JRA1	**Title**	Novel techniques and methods for environmental and plant measurements, including model-assisted phenotyping

Date of delivery	**Contractual**	29/07/2019 (Month 27)	**Actual**	05/05/2020 (Month 37)
Dissemination level	X	**PU Public, fully open, e.g. web**		
		CO Confidential, restricted under conditions set out in Model Grant Agreement		
		CI Classified, information as referred to in Commission Decision 2001/844/EC.		

Authors (Partner)	UNOTT, UCL			
Responsible Author	**Name**	Tony Pridmore	**Email**	tony.pridmore@nottingham.ac.uk

Version log			
Issue Date	**Revision N°**	**Author**	**Change**
08/04/2020	1	Tony Pridmore	Reviewed by WP leader
29/04/2020	2	Cloé Paul-Victor	Reviewed by Project Manager
05/05/2020	3	François Tardieu	Reviewed by Project Coordinator
05/05/2020	4	Xavier Draye	Reviewed by WP leader

Executive Summary

The primary goal of this deliverable is to secure the requirements enabling experiments spanning EPPN2020 phenotyping installations or post-hoc experiments combining datasets from several platforms. The ability to perform such experiments is vital if the maximum possible scientific benefit is to be obtained from the pan-European network of facilities which EPPN2020 seeks to create. Sound methods for the integration of data from multiple sites would support the production of richer data sets and exponentially increase the range of experiments accessible to the community.

This document first examines the need for calibration of phenotyping platforms and catalogues the range of instruments and calibration procedures currently used. EPPN2020 has decided to recognise as Level 1 calibration the (independent) adoption by each platform of procedures that ensure the correspondence between values measured across installations viz. the same leaf should be attributed the same size in different installations. It then details a Level 2 calibration procedure which allows assessing the quality of calibration procedures between installations. This new method is trialled in spring 2020 and should lead to the production of a generic calibration tool for current and future imaging installations and pipelines.

Phenotyping installations vary significantly in their structure and operation, and in the abilities of their operators to access and control their individual components. Some have been created in-house, resulting in deep understanding of their operation and complete access to all their parameters and intermediate results. Others, usually those acquired from commercial concerns, are black boxes: their operators can only provide input plants and receive output traits. Before attempting to design a Level 2 calibration process it was therefore important to assess the methods used in EPPN2020 installations. A survey was conducted (Section 2) which showed that Level 1 calibration is in place in a majority of installations. Attention then turned to possible Level 2 approaches, specifically for image-based phenotyping methods.

Previous attempts to calibrate phenotyping installations beyond Level 1 relied upon the use of common physical objects. However, it was impractical to image the same set of real or artificial plants in different phenotyping platforms.

In the approach adopted here, phenotyping installations apply their distinct analysis pipelines to shared images, rather than shared physical objects. Tools for the recovery of camera position and parameters are first made available and used to characterise individual imaging environments. This allows a set of real and artificial images, derived from real and virtual plants, to be made available. These images reflect the imaging geometry present in each installation. Therefore, differences between the results of their analyses allows improvements to be made.

Authors/Teams involved: Tony Pridmore (UNOTT), Sixtine Passot (UCL), Xavier Draye (UCL).

Table of contents

1. INTRODUCTION

The primary goal of this deliverable is to secure the requirements enabling experiments spanning EPPN2020 phenotyping installations or post-hoc experiments combining datasets from several platforms. The ability to perform such experiments is vital if the maximum possible scientific benefit is to be obtained from the pan-European network of facilities which EPPN2020 seeks to create. Sound methods for the integration of data from multiple sites would support the production of richer data sets and exponentially increase the range of experiments accessible to the network.

Central to the integration of data from multiple platforms is the availability of consistent calibration procedures. Reliable methods are required of ensuring that measurements from different installations are compatible; that e.g. leaf area measurements taken in France can be safely used alongside leaf area measurements obtained from plants of the same genotypes grown in the UK. These calibration methods, complemented with suitable environmental characterization (JRA1.1), statistical methods (JRA2) and information management systems, will leverage the network capacity to perform cross-platform experiments as conducted in JRA1.4.

Calibration methods depend on the instrument or protocol that are used to generate the data. Many instruments such as scales, thermometers or spectrometers (among others) generate raw signals that are easily converted to absolute measurements using (internal) calibration functions. Provided these calibration functions are re-adjusted regularly (in house or factory calibration, typically against standard samples), data generated in different installations by such instruments can be combined without any further correction.

Beyond these simple cases, however, plant phenotyping involves instruments whose output cannot be converted into absolute measurements. This happens especially when 2D cameras are used to image 3D objects in order to determine size measurements (length, area, volume). In this case, the raw size measurements, in pixels of projected objects, depend on known camera parameters but also on the unknown extent to which each object develops in the third dimension. For example, a perfectly flat object positioned orthogonal to the line of sight of the camera can be measured exactly if the camera parameters are known, while a 3D object will present occlusion and parallax effects that cannot be corrected for without knowing the exact 3D structure of the object. A crude but practical way to cope with this constraint is to perform manual measurements and collect image-derived data on a set of ‘calibration plants’ to estimate statistical regression functions that allow to predict absolute measurements from image-derived data collected on ‘experimental plants’. When using manual measurements, sufficient plants should be examined as to allow a reliable estimation of the regression parameters. This method will also yield an estimate the accuracy of the predicted data.

These methods, which include the recourse to standard samples or to manual measurements, are recognised by EPPN2020 as ‘Level 1 calibration’. Their adoption allows users to be confident that the data generated by a given installation is reliable and can be combined with that produced by other Level 1 calibrated facilities. The first part of this deliverable (see 2. below) reports on the monitoring of the adoption of Level 1 calibration for instruments and installations within EPPN2020.

Level 1 calibration is applied within individual phenotyping installations and allows transfer of phenotyping data (plant traits) between installations. However, when applied to image-derived size measurements, it has several inherent limits.

- It considers the performance of an installation as a whole. No insight is provided into the properties of distinct elements of the phenotyping pipeline used. Typically, the

effects of segmentation (lighting issue) and of projections (camera – object geometry) on size measurements are both combined.
- It does not support the export of intermediate data from within a phenotyping pipeline, only its final product. Specifically, it does not help images acquired in one installation to be analysed with another image analysis pipeline.
- In a network management context, the capacity to compare the performance of different installations (benchmark) is rather limited because different sets of plants (e.g. size or age, genotypes) are used by installations.
- It has to be applied individually to every variable exported from the pipeline. This is contradictory to a common argument for the use of image-based methods, viz. that once captured, image data can be revisited to extract further traits, or provide higher quality estimates of previously studied values when improved image analysis methods become available.

To help solving these limitations, we propose a practical and effective Level 2 calibration mechanism, based on the recovery of installation parameters (see 3. below) and the recourse to common calibration material. This mechanism will contribute to open data strategies by assessing variation across installations at the component level, and by supporting the transfer of raw data between installations – we focus here on image data – and the re-analysis of raw data as new methods emerge.

2. ADOPTION OF CALIBRATION METHODS IN EPPN2020

2.1. Overview of Calibration Approaches in Phenotyping Installations

In order to get an overview of the different situations encountered in phenotyping installations and the calibration solution put already in place, we conveyed a first survey in 2018. The questions of this survey were oriented towards the comparison of measured plant traits with actual plant traits. This survey was filled in by 13 installations and represented a wide panel of possible situations encountered in the network. This survey was rich in teachings. Calibration procedures seemed to be much more diverse than what we initially thought. We initially focused on traits extracted from images but many other situations exist in EPPN2020 installations, and the calibration procedures put in place are not always straightforward.

Following this analysis, two situations can be distinguished. In the first case, a measured trait can be calibrated against measurements obtained with a reference method considered as more reliable, for example calibrating the length extracted from picture against that measured with a ruler. In the second case, the measuring instrument used is already the most reliable we can think about, for example a scale to measure biomass. In that case, the only way to validate the measurement is to calibrate the instrument itself.

2.2. Extended Calibration Principles

To formalize this idea, we produced a list of phenotyping instruments that are used in phenotyping installations.

Imaging	*Other devices that need calibration*	*Devices that do not need calibration*
Camera (2D sensor)	Chlorophylmeter	Ruler

Scanner (line sensor)	IRGA	Hg or alcohol thermometer
3D scanner (Phenospex-like)	Porometer	Caliper
CT scanner	Scale	
RGB camera	Microwaves	
Fluorescence	Pressure bomb	
IR/thermal camera	Tensiometer	
NIR	Displacement sensor	
Multispectral	Thermal sensor	
Hyperspectral		
Fisheye camera		
MRI		

In this table, imaging devices have a special place, as their calibration depends on the type of object they should image. Imaging device may be used to retrieve a shape information (eg: leaf length) or an information based on pixel value (eg: photosystem efficiency based on chlorophyll fluorescence). Possible calibration procedures thus vary with the type of instrument and the type of output. We conceptualized these reflections in the form of a decision tree (**Erreur ! Source du renvoi introuvable.**).

It turns out from Figure 1 that the situation of morphological traits is the more complex. Morphological traits include length, area, volume and angles. In some installations, plant part mass is estimated from their computed biovolume and may also be considered as a morphological output. Morphological traits can be measured directly with ruler, tape measurer or caliper that do not need calibration and can serve as reference for other measurement methods. In most cases, however, morphological traits are retrieved from images and their raw value is expressed in pixels, or voxels.

When objects are in 2D, typically roots growing on a rhizotron glass or individual leaves put on a flatbed scanner, a simple scale factor is sufficient to convert pixels into length or area. Some 3D objects can also be considered as near 2D from some specific viewing angles. It is the case of Arabidopsis rosette in top view or young maize leaves correctly positioned in a side view. In that case, a scale factor estimated with a reference object may be sufficient to compute lengths and angles.

Some platforms are equipped with devices that produce 3D images of 3D objects. These devices may be CT-Scanner, MRI or a combination of 2D cameras with a specially designed system to combine image and reproduce a 3D scene. For these cases, scale factors may be sufficient to measure length, in the same way as for scanner and 2D objects. Physical validation of estimated length with manual measurements would also make sense.

The most challenging situation involves installations that use 2D imaging devices to estimate morphological traits from 3D objects. For example, leaf length and leaf angle of maize plants can be estimated from a 2D image if the camera is placed wisely. However, leaf width and therefore leaf area cannot be computed easily in that configuration. For other plant species like wheat or tomato, even measuring leaf length may be tricky. In those cases, a global estimate of aerial biomass is generally performed from the total area of green pixels, calibrated against actual measurements on sample plants. Finer estimation can be done using images taken from different viewpoints. This situation is developed in the next part.

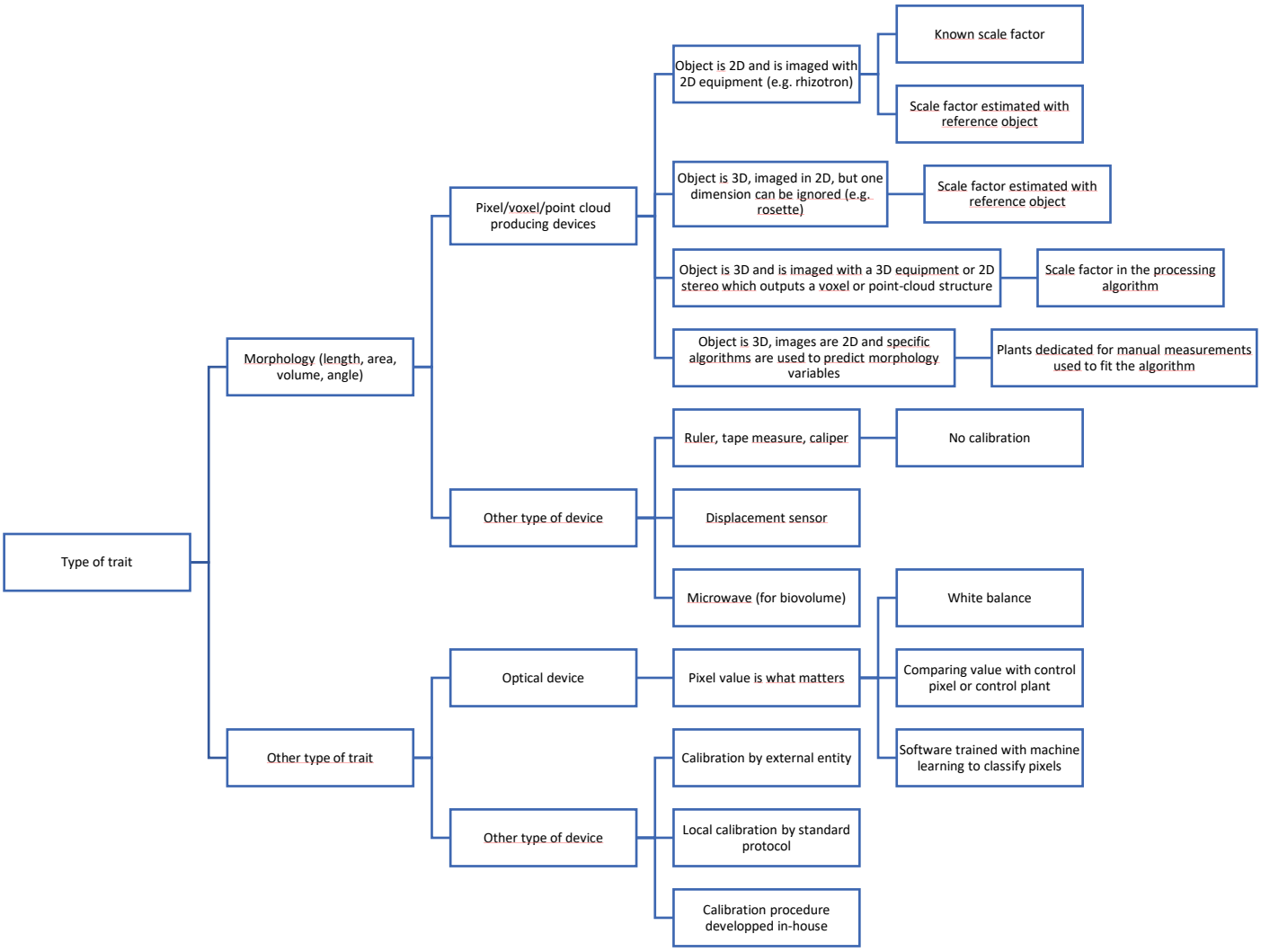


Figure 1 Types of calibration depending on device and measured objects

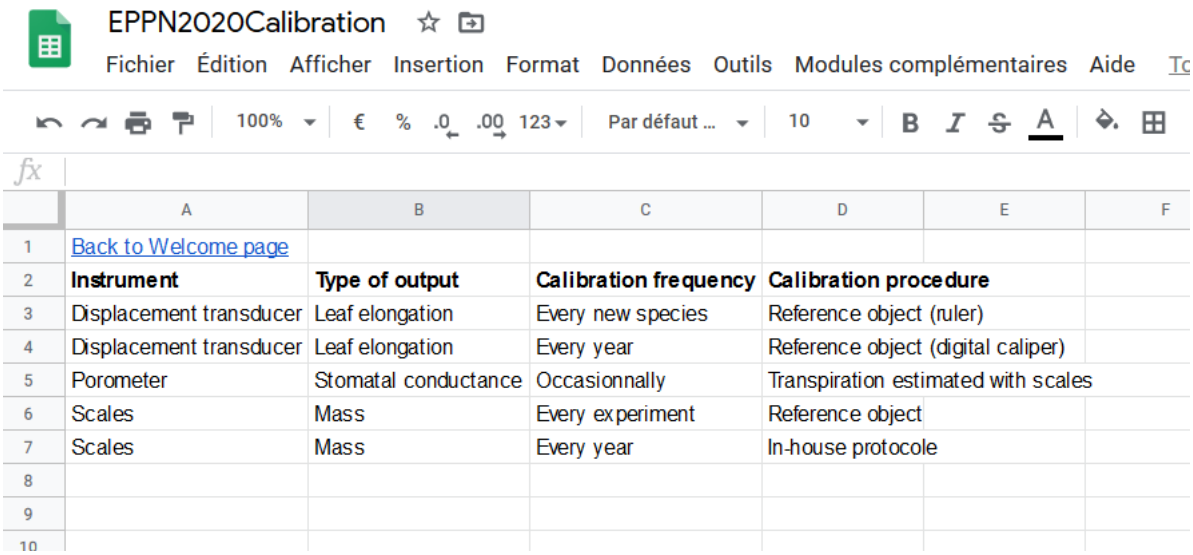
Automated size measurements from images almost invariably start with the segmentation of the plant or plant organs from the background. Usually, the imaging configuration (e.g. color of background) is designed to maximise the segmentation performance. However, several aspects of an imaging cabinet influence that performance. The positioning of the light sources relative to the leaves often create specular light reflection (hot spots) towards the camera, yielding to low performance segmentation. The imaging cabinet specifications and the image processing pipelines should therefore be considered in the calibration procedure.

Finally, some optical instruments are used to retrieve pixel value. This concerns for example the measurement of reflectance or chlorophyll fluorescence. As long as the area of interest can be identified, calibration is only needed for pixel signal intensity and wavelength accuracy and is typically made on the instrument itself. The situation becomes complicated again if an organ pixel values have to be weighted by the organ area, for example.

2.3. Comprehensive status of calibration in EPPN2020 installations

Keeping these ideas in mind, we performed a second survey over all installations in February 2020, in the form of a google sheet. This survey was more generic than the first one, to ensure

that every installation can fit into it. We favored universality upon specificity. We asked installations to list their plant measuring instruments and to specify, for each of them, the type of data that is obtained, the kind of calibration technique which is used and when this calibration is performed (Figure 1). In this survey, the same instrument may appear several times if different calibration methods are used for it.


EPPN2020Calibration

Fichier Édition Afficher Insertion Format Données Outils Modules complémentaires Aide

	A	B	C	D	E	F
1	Back to Welcome page					
2	**Instrument**	**Type of output**	**Calibration frequency**	**Calibration procedure**		
3	Displacement transducer	Leaf elongation	Every new species	Reference object (ruler)		
4	Displacement transducer	Leaf elongation	Every year	Reference object (digital caliper)		
5	Porometer	Stomatal conductance	Occasionnally	Transpiration estimated with scales		
6	Scales	Mass	Every experiment	Reference object		
7	Scales	Mass	Every year	In-house protocole		
8						
9						
10						

Figure 1. Screenshot of the 2020 survey for one platform. Single devices use one line per calibration method or frequency

As of April 2020, 63% of the installations have completed the survey (22/35)

2.4. Examples of calibration procedures

In this section, we present several examples of original calibration methods performed in phenotyping installation (as described by partners). They were detailed in the first survey or during direct discussions with installations.

2.4.1. Stomatal conductance measurement

Transpiration measured with scales is classically used as a reference technique. It can be compared with stomatal conductance measured at leaf level. For our installation, the comparison of whole plant transpiration with stomatal conductance measurements revealed very consistent ranking among genotypes.

2.4.2. Chlorophyll fluorescence parameters

We measure chlorophyll fluorescence parameters (CFP), depending of position and radiation under sensors. We employ a reference CF card to evaluate homogeneity of active lights. We check the homogeneity of illumination across the imaged zone. If the illumination is heterogeneous, we change LED panel incidence.

Every experiment includes internal control plants. They serve for thresholding calibration in each experiment. Based on these thresholds, we extract the diseased area on each leaf. The analysis methods used were published in Plant Methods (Rousseau et al. 2013, Rousseau et al. 2015).

2.4.3. GFP fluorescence

It is possible to calibrate reported traits against actual plant parameters only when these traits are already known. We do it in the starting phase as we optimize a particular screen. For

example, we want to measure the GFP fluorescence of a GFP-marked pathogenic bacterium as a measure for bacterial growth in Arabidopsis seedlings. In this particular experiment, we use 5 Arabidopsis lines with known phenotypes (1 control, 2 more resistant than control, 2 more susceptible than control) and compare bacterial growth curves with fluorescence measurements. Afterwards, we will only measure fluorescence and include the control genotype in all analyses.

The platform measures fluorescent signals, e.g. growth of GFP-marked pathogens in plants or fluorescently stained pathogens or pathogen defence markers. Measurements are always at the microscopic level and performed in multiwall plate formats. The platform output includes a ranked list of fluorescence intensities comparing individual wells in a series of multi-well plates. The analysis software measures the fluorescence intensities as 'grey values' and gives the RFU (relative fluorescence units) for each image. We compare these versus our control. The average RFU from control samples is deducted from experimental samples as part of normalization.

2.4.4. Metabolomics

In our platform, we are extracting and quantifying metabolite concentrations and enzyme activities using targeted assays. Our calibration protocols have been developed and validated in-house. We are working in batches, corresponding to a microplate. Every batch includes several replicates of a biological standard, which enable the calculation of levels or activities as well as the calculation of a CV for every analyte. The CV is used to validate/invalidate the measurements. Every assay is optimised by checking recovery (spiking), extract dilution effect and technical error. When technical error gets unsatisfactory, pipetting parameters are modified (e.g., aspiration height, aspiration speed, etc.).

The major source of error is matrix heterogeneity (e.g., quality of sample grinding). We constantly try to improve the quality of grinding. Biological standards are used to calibrate and validate the methods in a given matrix (value, recovery, dilution, technical error). Aliquot weighing and pipetting are further sources of error, which we also try to improve regularly. Elsewhere, components of the installation are regularly checked (5-6 times a year) and a thorough verification is performed every 2 years (maintenance operations).

2.4.5. Roots in 2D on a contrasted background

In this installation, roots grow in 2D on a uniform background, which colour has been chosen to offer a very good contrast. A semi-automated software was developed specifically to retrieve root axes with minimal human intervention. The scale factor converting pixels to distance is given as an input once for the every batch of pictures, as the camera keeps the same fixed position during the whole experiment. In this situation, measuring length was not a problem, the challenge was to accurately retrieve the position of root axes (RootNav, Pound et al, 2013).

3. TOWARDS LEVEL 2 CALIBRATION

3.1. Objects vs Images

Previous attempts to calibrate phenotyping installations beyond Level 1 have sought to provide platform operators with access to common objects. However, it is impractical to image the same set of real plants in different phenotyping platforms.

The first EPPN project addressed this problem using more robust, artificial plant-like objects. A collection of commercially available artificial plants of different sizes and configuration was gathered. These were very similar in structure and appearance to real plants. In some ways they were too similar; while stronger than natural plant material, the thin plastic these were

constructed from was also flexible. The plastic plants were often distorted in transit and could not be relied upon to be the same shape when they arrived at each installation.

To combat this, a set of more abstract but recognisably plant-like objects were constructed from metal bar and plates. Three such plants were made available, all similar in appearance (painted green) but of different sizes. These are robust enough to travel between sites, but their abstract design raised questions as to their validity: it was not clear that the assumptions made by image analysis methods designed to operate on real plants were satisfied by images of the metal objects.

More importantly, however, these objects did not allow details of the viewing geometry to be recovered. Although the installations all viewed the same objects, they did so from different viewpoints. There was no way to ensure that viewing positions were identical, or to know what the viewing parameters were. This made it impossible to determine if differences in trait measurements were due to variations in processing pipelines, or simply in camera placement.

The work outlined in the initial proposal relied upon the use of emerging 3D printing technology to allow model plants to be generated locally, avoiding the need for expensive and potentially damaging transport of calibration objects between sites. CAD (Computer-aided Design) files describing the calibration objects were to be prepared and distributed across EPPN2020. The intention was to embed standard calibration targets (traditionally chequerboards, see figure 3) within the objects, and require platform operators to position the object so that those targets were visible. This has the dual effect of constraining the viewpoints from which images were acquired and allowing recovery of the actual viewpoints. Given an image of a standard calibration target, computational methods and software tools exist which can recover both the intrinsic (e.g. focal length of the lens used) and extrinsic (position and orientation of the camera in 3D space) to a high degree of accuracy.

Figure 3. A standard chequerboard calibration target

Following discussion within the consortium, however, this approach was abandoned. Comparatively few partners were confident that they could obtain access to 3D printing facilities of sufficient size to be relevant for plant calibration .

3.2. A Novel Approach to Platform Calibration

The approach to platform calibration that has emerged only requires simple objects and digital images to be shared between sites. The physical objects are standard chequerboard targets, which can be manufactured centrally and easily distributed to participating installations. There is no need to pass objects between installations. The complete approach is described below and has been deployed in the network early 2020, giving a priority to installations participating to the cross-platform experiment (JRA1.4).
The principle consists in providing installations with a collection of images of two sets of plants (the same for all installations). The first is a *training set*, delivered with ground truth data (corresponding to manual measurements). Partners compute image-derived data using their pipeline and analyse it with ground truth data to estimate a suitable regression function. The

second is a *test set*, from which partners compute image-derived data and apply their regression function to estimate (predict) absolute size variables. To separate the effect of camera parameters from all other effects (e.g. thresholding), the images sent to partners are captured using the same camera settings as the one used in their installation.

The different methods that are deployed are explained below. All of them require a characterisation of imaging environments, as proposed in Level 2 calibration.

3.2.1. Characterisation of imaging environments

The following procedure is used to characterise imaging environments:

- Each participating installation to provide UNOTT with 2-3 images showing the environment in which plant images are acquired. This gives an initial impression of view distances and angles, and any physical constraints on plant placement.
- UNOTT designs a calibration target, or small set of targets, based on a standard chequerboard design and distributes it to each participating installation. Partners acquire images of the calibration target in their imaging environment and return them to UNOTT
- UNOTT applies standard camera calibration methods to recover the intrinsic and extrinsic parameters of the installations' image acquisition methods.

This process allows images to be generated that are in line with the imaging geometries used (see below). It also provides an overview of the range of variation in camera parameters across EPPN2020 facilities. This is a broad indicator of the scale of the calibration problem: the more tightly clustered the imaging geometries reported by this exercise the more likely it is that variation in measurements will be small.

3.2.2. Artificial Images from Real Plants

Given (i) details (intrinsic and extrinsic camera parameters) of an installation's imaging system and (ii) access to 3D models of the training and test plants, artificial images of these plants as they would be captured in the context of the installation can be generated. The procedure deployed in EPPN2020 is the following:

UNOTT acquires 3D models of the two sets of plants using X-ray CT at the Hounsfield Facility. Separation of plant material from air in the resulting images is straightforward. The resulting plant model (the set of voxels considered to represent the plant) will be approximated by a triangular mesh surface. This approach provides very accurate 3D ground truth which has been used to assess both 3D measurements acquired from multi-view reconstruction methods (Gibbs et el 2019) and to generate artificial images for use in evaluating image analysis methods (Pound et al). Surface mesh models can be projected onto an imaging plane to produce a photorealistic artificial image from any chosen viewpoint (Figure 4).

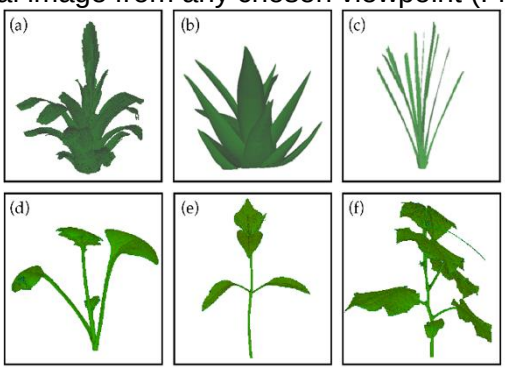


Figure 4. Artificial plant images (Gibbs et al 2019)

Using the parameters obtained in 3.2.1, to generate artificial images of the scanned plants for each installation. These images mimic the images that would be obtained if the same plants were imaged in the installation. The method makes it possible to obtain accurate ground truth size measurements for each plant.

This work relies on artificial images. While these exhibit a high degree of realism, any approach not using images of real plants is open to criticism that these are not truly representative. To address this, a second method employing real images will also be used.

3.2.3. Real images

UNOTT has developed a flexible imaging cell comprising a 6 degree of freedom robot arm holding a digital camera, and a computer-controlled turntable upon which the plant of interest is placed. The ability to synchronise control of turntable and arm means that the camera can be placed to acquire an image of the plant from any specified direction, and a good range of viewing distances (Figure 5).

Figure 5. The UNOTT imaging cell

In this second method, this facility will be used to capture images of the training and test real plants (the same as in 3.2.2) from camera locations specified by the parameters recovered from each installation (3.2.1). The calibration data acquired from installations during the setup phase is used to position the camera to mimic each installation's environment, and images are acquired.

The effect is that each installation is provided with real images of a set of standard plants acquired in a close approximation to their own facilities.

3.2.4. Artificial images from virtual plants

A parallel route is being used in which virtual plants are used to generate artificial images in the same way as**Erreur ! Source du renvoi introuvable.**. A collection of 4500 virtual maize plants representative of a broad scope of maize morphology has been assembled at INRA, and can be used to generate artificial images corresponding to each installation setup using standard algorithms. The realism of these images is usually good (Figure 4), although less than artificial images obtained from real plants. The main advantage of this procedure is that the number of plants that can be used to assemble training and test sets is considerably larger than what is achievable with real plants. In addition, because these plants are constructed based on predefined absolute size, their ground truth is known.

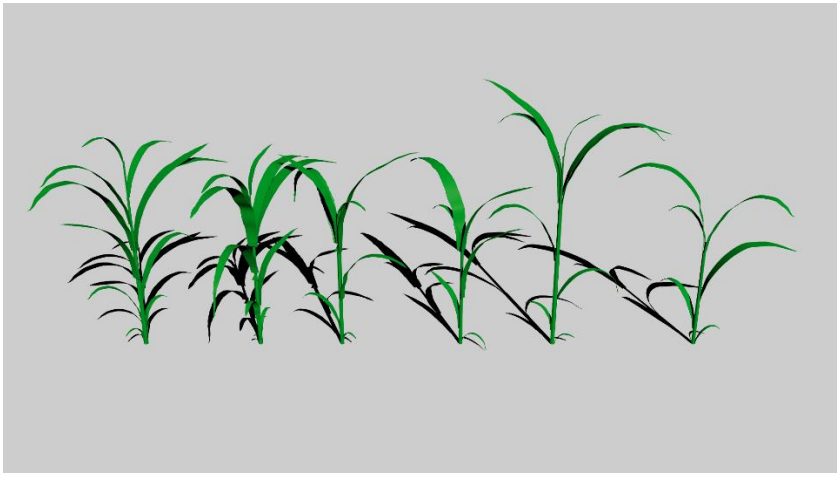


Figure 4. Illustration of virtual maize plants (C. Fournier, INRA).

3.3. Evaluation of installation performance

As explained above, partners now run the training set of images through their pipeline to compute image-derived data, which they analyse with the ground truth measurements provided with the training set to generate prediction functions. Afterwards, they run the test set images through their pipeline and use the newly adjusted prediction functions to estimate absolute size data which they send back to UNOTT.

The result of this sequence of operations are four-fold:

- a novel data set characterizing the imaging environments in use across the EPPN2020 community
- data assessing the variation in results obtained across platforms
- data assessing the expected variation in results (obtained by collating the trait estimates generated directly from the 3D models)
- comparisons between platform data and ground truth measures obtained from the 3D models, which will allow corrections to be made.

The strengths of this approach are that real plants are in some sense “shared”, a common calibration tool is used and that ground truth is available.

Other analyses of the generated datasets will also be made. On the one side, we will focus on the statistical properties of the data set generated, rather than detailed evaluation of individual trait estimates. Plots will be produced of the traits reported by the installations involved, providing a distribution against which individual responses can be evaluated. This mirrors an approach used by the WHO when assessing measurements made by diagnostic centres in similar circumstances. On the other side, we will also analyse the image analysis pipelines used by contributing partners. Examination of intermediate representations, e.g. segmented images, will reveal critical stages in trait recovery, and highlight common image analysis challenges. This mirrors the standard evaluation methods used within the computer vision community, where statistical analysis is commonly followed by detailed consideration of failure modes.

3.4. Resources needed

The planned calibration method and experiment require some hardware and software resources. Those needed at participating installations are low cost and easily accessible, while

the more costly equipment is only needed at UNOTT, where it is available. The tools and resources needed to support this experiment are:

1. At participating sites, a suitable camera calibration target. This is a planar object of known geometry which can be placed in the field of view of the camera. A black and white checkerboard, as shown in figure 1, is traditionally used. The recent EMPHASIS-PREP survey of phenotyping platforms showed that many already have access to such targets. Those that do not can create one by printing a paper target which is then mounted on a rigid board. Thick (1cm) MDF board is stable enough to produce a usable target, though some groups prefer aluminium sheet, if that is available. The target can be downloaded, free of charge, from one of the two sites:

   http://boofcv.org/notwiki/calibration/letter_chessboard.pdf
   http://boofcv.org/notwiki/calibration/A4_chessboard.pdf.

   Standard laser or ink-jet printers produce a sufficiently high-resolution target. Artists' spray mount glue, available in aerosol form from art and craft shops, is an excellent, and low-cost way to mount the target on the board.

2. At UNOTT, camera calibration will be performed using EmguCV software, a cross platform .NET wrapper for the OpenCV software library. OpenCV's calibration tool imposes only one constraint on the input images; that each corner of the checkerboard must be visible. Output from this is a camera model consisting of:
   a. the intrinsic (internal) parameters of the camera, namely; focal length, sensor format and principal point
   b. the extrinsic (external) parameters (a rotation and translation)
   c. a re-projection error which specifies the accuracy of the calibration.

3. At UNOTT, 3D surface mesh models of a representative set of real and virtual plants and visualisation software capable of producing sufficiently realistic images. Some models from real plants are already available, others will be obtained using the X-ray CT scanners in the Hounsfield Facility. In addition, a collection of 4500 virtual maize plants at different growth stages is available from INRAE. All necessary software to compute artificial images are developed locally and are available at UNOTT.

4. At UNOTT the imaging cell shown in figure 3 will be used to capture real images. This is based around a Universal Robots UR5 arm and an LT360EX – Linear X Systems, Portland, USA – computer-controlled turntable. The cell is itself fully calibrated, allowing the camera to be placed at any location relative to the target object (the plant).

3.5. An Online Calibration Service

If the experimental evaluation described here is successful, and the approach considered valuable to the community, an online calibration service is envisaged which is open to use by installations at any time. Users of this service will:

- Login and provide some installation details
- Download a calibration target
- Image that target in their facility and upload the resulting images
- Receive artificial and real test images appropriate to their situation
- Run their analysis pipelines on these images and upload the results in a standard form
- Receive feedback on their method's performance against the ground truth (where available) and the community distribution (where no ground truth is available)

The community data set (distributions of results obtained) would be visible to all service users.

While it is possible for such a system to generate artificial images directly from users' calibration data, it will not be possible for it to capture new images of the real plants involved 'to order'. Instead, sets of images, each tagged with calibration metadata, would be pre-prepared and stored. User's calibration data would then be used to select the most appropriate member(s) of that set. While there would be some loss of accuracy – dependent on the number of images captured in each set – this would provide a valuable calibration and monitoring service in line with those in place in medical disciplines.

4. CONCLUSION

This document analysed the current calibration practices for the main instruments used in phenotyping installations. Based on this analysis, it has proposed two levels of calibration. Level 1 indicates the adoption by an individual installation of methods ensuring the compatibility of measurements with those from other installations. Level 2 moves to the adoption by the community of a common set of resources (3D plant models, calibration targets and software) that enable the calibration of each installation against common samples and the identification of possible improvements inside each installation pipeline. The approach minimises the transfer of materials between sites and requires no physical objects to be transferred. This new calibration approach, if found to be effective, holds out the possibility of an ongoing, online calibration service.

This work is in finalisation at partners sites. The Coronavirus situation at the time of delivery of the document is delaying some operations. The recourse to virtual plants, which was proposed early in the project, will reduce this impact.