European Plant Phenotyping Network

EPPN 2020

D2.1 Quality control of experimental design for phenotyping platforms

Fred van Eeuwijk, Emilie Millet & Darren Murray

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°**	D2.1	**Title**	Quality control of experimental design for phenotyping platform
Work Package	**N°**	WP2	**Title**	Design and analysis of phenotyping experiments across multiple platforms, scales of plant organisation, traits and management conditions

Date of delivery	**Contractual**	Month 27	**Actual**	Month 30
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)	Fred van Eeuwijk, Emilie Millet & Darren Murray			
Responsible Author	**Name**	Fred van Eeuwijk	**Email**	fred.vaneeuwijk@wur.nl

Version log			
Issue Date	**Revision N°**	**Author**	**Change**
15/07/2019	0	Emilie Millet	first version
06/09/2019	1	Darren Murray	Correction from co-authors
22/09/2019	2	Fred van Eeuwijk	first review by WP leader
07/10/2019	3	Project Coordinator	Final version

Executive Summary

This document presents a strategy developed in WP2 of the EPPN2020 project for quality control of experimental designs for in-doors phenotyping platforms. The main objective was to help platform users finding an adequate experimental design taking into account restrictions and limitations imposed by the specific conditions encountered in each individual phenotyping installation. A number of installations part of EPPN2020 were visited to make an inventory of their specifics and a survey was conducted to extract statistically intelligible descriptions of the installations. The survey revealed that platform managers are already aware of existing spatial trends in their installation and that they already try to account for these trends in experimental designs. Descriptions formulated on the basis of visits and surveys served to propose experimental designs following the rules for experimental design (replication, randomisation and restriction). Further improvements in experimental design are possible when spatial trends are mapped more precisely. A model for spatial adjustment of treatment estimates in phenotyping platforms was developed for experiments performed in time.

Table of contents

1. INTRODUCTION

1.1. Aim of the EPPN2020 project

The EPPN2020 project aims at providing public and private European plant scientists access to a wide range of state-of-the-art plant phenotyping facilities, techniques and methods. It will help the plant community in progressing towards excellence across the whole phenotyping pipeline that includes sensor and imaging techniques, data analysis adjusting treatment contrasts for environmental conditions and placing interpretation in a biological context, data organization and storage, and analysis of series of experiments as well as meta-analyses of experiments.

EPPN2020 coordinates its activities with the future infrastructure EMPHASIS, listed in the ESFRI roadmap, and with national programs. EPPN2020 involves:

- access to 31 key installations in 15 infrastructures,
- a Work Package on sensors (WP1),
- a Work Package on data analysis (WP2),
- a Work Package about data management (WP3),
- networking activities for establishing cooperation and increasing integration between facilities both within and outside EPPN2020.

1.2. Scope and aim of the document

New phenotyping platforms require a reconsideration of classical experimental design and analysis techniques currently used for field experiments. Although this is not widely recognised, spatial and temporal heterogeneities in controlled conditions are as large as in the field, if not larger, so it is essential that users choose appropriate experimental designs, models and analysis methods. WP2 addressees the lack of statistical design guidelines and analysis tools for data from phenotyping platforms. It has developed protocols for choosing experimental designs tools and statistical analysis of phenotyping experiments across platforms and scales of plant organization.

This document describes guidelines and diagnostics for choosing experimental designs for phenotyping platform. We arrived at these guidelines and diagnostics by (1) helping partners to describe their installation in a statistically intelligible way, (2) giving advice and feedback on experimental design applying statistical rules for choosing experimental design and (3) using a flexible model for estimating spatial trends whose intensity and structure will determine which experimental designs are suitable for next experiments.

2. INVESTIGATING TRENDS AND DISTURBANCES IN PLATFORMS

Phenotyping facilities display spatial heterogeneity. For example, the spatial variability of incident light can go up to 100% between pots within a glasshouse (Cabrera-Bosquet *et al.*, 2016). Taking into account these spatial trends and correcting for them in an appropriate way is a prerequisite for precise and unbiased estimation of genetic and treatment effects. In the same way as in field trials, a platform experiment needs to follow standard principles for experimental design and statistical modelling.

In the EPPN2020 project, for platforms to be allowed to provide Trans-national access (TNA), the platform owners are expected to use state-of-the-art experimental designs and statistical analyses. A first step towards this expectation was a clear description of the installation in a statistically intelligible way and a motivation for current design and analysis choices. This description included the identification of the major sources of variation. Compared to field experiments, there are additional constraints and limitations to take care of in platform experiments. These can come from the platform physical structure (e.g. pillars in greenhouse or rainout shelter), the layout and available area (e.g. limited number of plants in growth chamber), or the technology/management (e.g. specific foliar treatment, robots).

We conducted a survey and visited a number of facilities in 2017/2018. Based on the outcome survey and visits we concluded that EPPN2020 platform managers have a good knowledge about the possible sources of error variation on their installations: they are able to describe the main trends likely to affect the plants (Fig.1) and a majority of them have already quantified the temperature, water and/or light variability in their installations (Fig.2). Therefore, they are in a good position to choose a suitable design, provided that basic rules for choosing experimental design are followed.

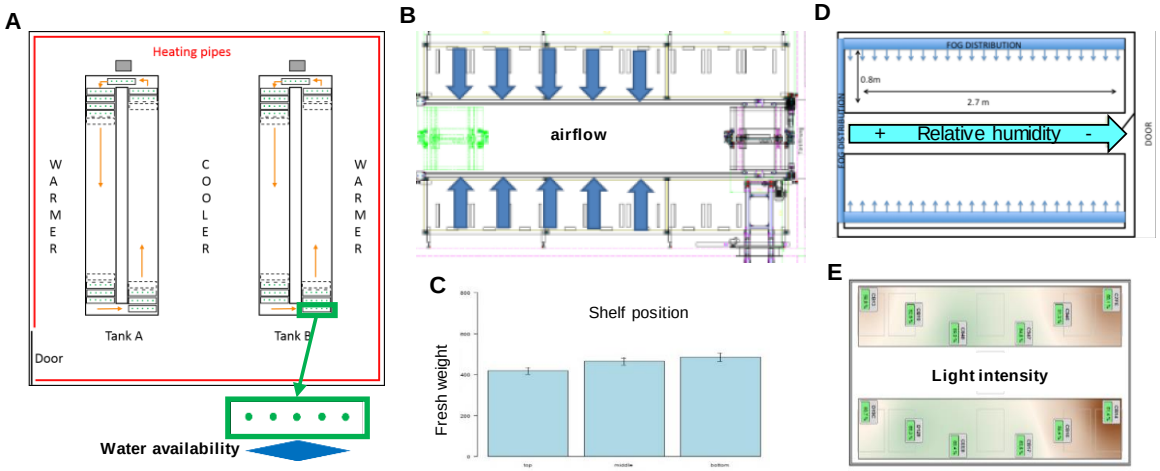

Figure 1. Examples of environmental variability in three EPPN2020 platforms. A, The heating pipes in the greenhouse create a temperature gradient. The water sprays create a water gradient within the tank (platform RootPhAir, UCLouvain, Belgium, courtesy of Sixtine Passot and Xavier Draye). B, Environmental variation is caused by the air ventilation on the left and right side behind the shelves. C, The light variability at the different layer positions of the shelves cause significant differences in the growth of the plants (B, C: platform Growthscreen, FZJ, Germany, courtesy of Fabio Fiorani). D, Scheme of the distribution of the devices for the control of relative humidity. Based on observations, there is a probable gradient of the relative humidity. E, Light map representative of light variations within a growth chamber equipped with sodium lights (D,E: Phenotic platform, INRA Angers, France, courtesy of Tristan Boureau and Etienne Belin).

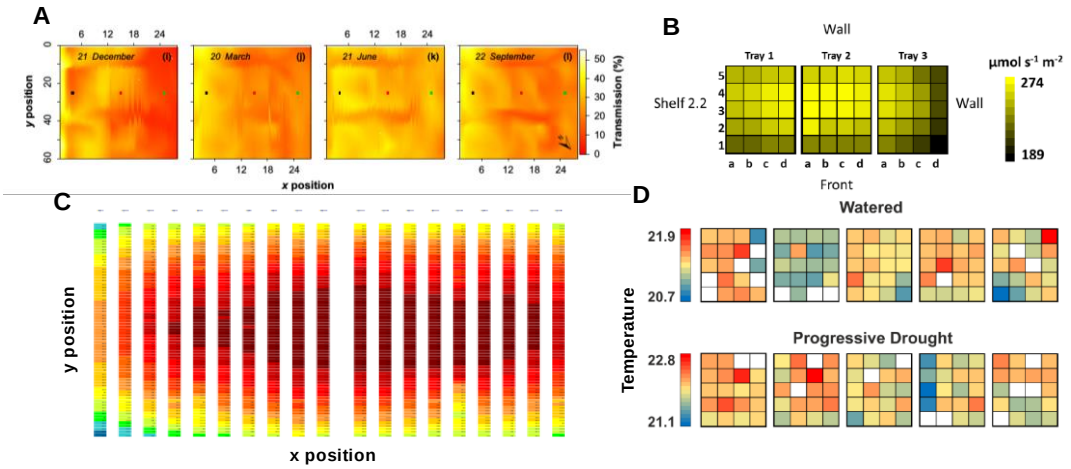

Figure 2. Examples of quantified environmental variability in four EPPN2020 platforms. A, Variability of transmitted light in the Phenoarch platform, at four different dates (INRA Montpellier, France, (Cabrera-Bosquet et al., 2016)). B, Light variability of three trays at one shelf level at the NaPPI platform (Helsinki, Finland, courtesy of Mirko Pavicic and Kristiina Himanen). C, distribution of the photoactive radiation when the lights are on in March at the 4PMI platform (INRA Dijon, France, courtesy Christophe Salon and Julien Martinet). D, Temperature variability at another NaPPI platform in four trays in two water scenarios.

There are three basic principles of experimental design:

- replication, that allows quantifying the experimental variation between experimental units and increasing the precision of estimated effects,
- randomization, to avoid confounding of treatment differences and (unknown) other differences between (groups of) units,
- restriction of randomization, or blocking, which is a local control to reduce the experimental error by grouping experimental units into blocks.

The platform user need to provide proper consideration to these principles before carrying out the experiment. A non-optimal experimental design could lead to imprecision, with large variability in estimates, bias (systematically wrong estimates of treatments), extra costs and wrong conclusions. It is important to note that, usually, even the use of an advanced statistical model cannot overcome the basic design flaws, and neither can the availability of a high-dimensional response.

There are major constraints and limitations in the platforms that were visited. One first outcome of the survey was that platform users often apply plausible blocking strategies in their experimental designs when the direction of blocking direction is evident. Figure 3 shows an example of a randomized complete block design (RCBD). RCBD is the simplest design including blocking: the assignment and randomization of treatments is not performed across the whole of all experimental units in the experiment, but treatments are assigned to blocks, at block level blocks are randomized, while within blocks randomization of the treatments assigned to that block takes place.

When blocks have not been defined a priori, they can be imposed a posterori. Such a posteriori blocking (post blocking) is less efficient than a priori blocking, but may still remove unwanted noise and allow some spatial adjustment. An example is given in Figure 4.

53	22	29	41	13	26	40	55	42	46	55	37	48	9
31	2	33	32	49	24	34	28	30	22	27	35	33	56
38	8	12	3	43	17	27	39	6	45	47	51	44	53
11	9	14	37	50	35	36	54	15	34	23	40	29	11
46	25	21	48	51	7	20	1	47	13	28	41	25	42
5	44	18	52	45	19	4	16	56	6	21	39	32	4
19	51	27	15	42	1	29	26	10	49	24	30	12	36
37	9	13	47	24	34	44	33	23	20	16	38	54	7
25	4	56	40	50	31	6	48	18	38	1	2	50	3
2	52	11	55	53	45	43	30	8	17	5	31	18	52
36	21	41	20	23	35	3	54	5	12	26	43	19	15
32	28	16	39	49	46	14	7	10	22	10	8	14	17

Figure 3. Randomized Complete Block Design (RCBD) at the Phenodyn platform (INRA Montpellier, France) with 168 plants (experimental units) grouped into 3 complete but irregular blocks (white, grey and black), each of which contains 56 genotypes. *Courtesy: Claude Welcker and Boris Parent.*

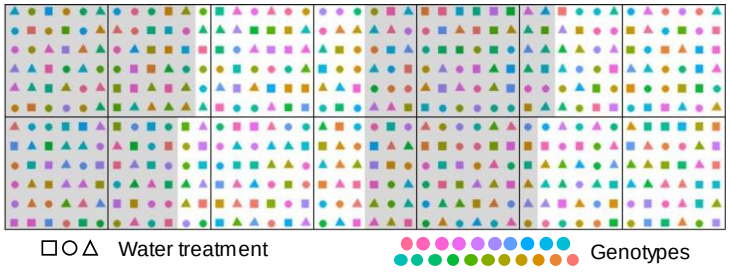

Figure 4. Completely Randomized Design (CRD), in the Phenopsis platform (INRA Montpellier, France). The experiment investigated the effect of three water treatments (shapes) applied to 21 Arabidopsis genotypes (colours) replicated 8 times per water treatment: a total of 504 experimental units to which the treatments are randomly assigned. The grey and white areas are incomplete blocks defined for post-blocking. *Courtesy: Denis Vile.*

In the majority of phenotyping installations, control and/or correction of micro-climatic conditions is essential. To this end, we can use two-way blocking strategies, like the Row-Column Design (RCD), where the blocks are best chosen following prior knowledge of the structure and magnitude of existing noise variation. The RCD approach consists in viewing the phenotyping experiment as a rectangular grid on a set of row and column coordinates ($r \times c$). Row and column blocks can be defined as incomplete blocks in two directions. To ensure that treatments will be as evenly spread as possible over columns and/or rows, it is possible, and sometimes desirable, to use a resolvable row-column design (Piepho *et al.*, 2015). In this case, complete blocks are first defined and then, within complete blocks, incomplete blocks are defined in the two directions ($r \times c$).

In platforms the measurement unit is either a small plot or an individual plant, which increase the variability between experimental units in comparison to what we are used to in field experiments. Moreover, typically there is a limited number of experimental units available in a limited space. Consequently, experimental designs may vary from one experiment to the next, even on the same installation (Fig.5).

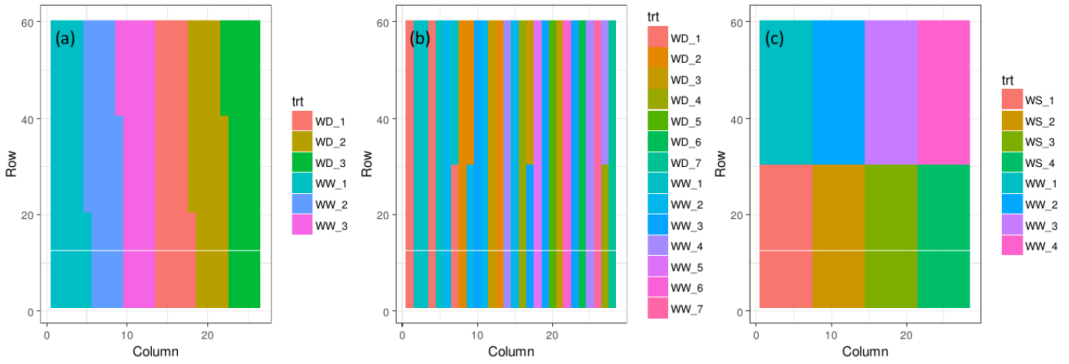

Figure 5. Three resolvable row-col designs at the Phenoarch platform (INRA Montpellier, France). Each coloured rectangle indicates a resolvable block (replicate); the blocking direction changes with the experiment and the plant material tested. *Courtesy Llorenç Cabrera-Bosquet and Claude Welcker*

In plant genetics, platform users try to maximize the number of genotypes they test. In this case, to be able to estimate the error variance and correct for the global and local trends, one

strategy is to partially replicate only a small number of genotypes of interest: the p-rep design (partially replicated design, Fig.6A). A p-rep design can be generated using any block design for the replicated entries, usually about 25-30% of them, and then augmenting it with the unreplicated entries by allocating them to the free plots in completely randomized order (Cullis *et al.*, 2006). Another strategy is the Augmented Row-Col design (Piepho & Williams, 2016) (Fig.6B).). In this design a large number of unreplicated varieties are arranged in a row-column design along with some check varieties that are highly replicated.

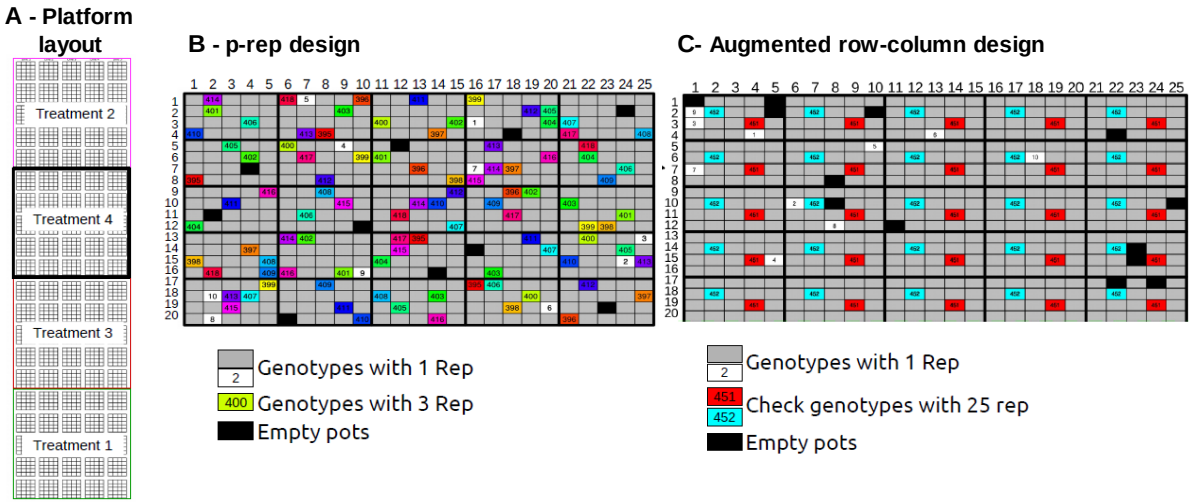

Figure 6. P-rep and augmented row-column design at the small plant platform in Aberystwyth. A, The full platform and experiment layout, split into four different treatments (nitrogen × plant density). Each treatment has been defined as blocks and within each block either a p-rep (B) or augmented RCD (C) has been applied. B, p-rep design with replicated genotypes in colour and unreplicated genotypes in grey. C, Augmented RCD with two highly replicated check genotypes (cyan and red) and unreplicated genotypes (grey). *Courtesy John Doonan and Gina Garzon.*

In addition, having one or more highly replicated check genotype(s) could help better quantify the impact of the spatial variability on the plants without dedicating a whole experiment to a uniformity trial (trial with only one treatment).

Another complication in relation to choosing a suitable design for phenotyping experiments is that at certain types of installations plants change position during the course of the experiment. Ideally, at every step of the experiment a suitable randomization should be chosen. However, in practice the position of the plants after the initial round of observations is determined by the mechanical restrictions of the installation that allocates the plants to a position on the platform. When plants are not allocated following a randomization scheme dictated by a statistical design, the subsequent statistical analysis is not obvious.

It is still statistically conceivable if a proper randomization is used at every time points. In that case, the RCD (augmented or not) is an attractive solution.

3. AN AUTOMATED ANALYSIS OF INDIVIDUAL AND SERIES OF PLATFORM EXPERIMENTS AND DIAGNOSTICS FOR IMPROVING FUTURE PHENOTYPING EXPERIMENTS

The statistical analysis of an experiment should separate genetic and treatment information from noise variation. From a statistical modelling perspective, the total phenotypic variation will be partitioned into: (i) extraneous variation that comes from the design, such as replicate or row and column effect, (ii) global trend variation that is an additive function of spatial coordinates, and (iii) the local trend variation with spatially dependent noise. Popular mixed models to separate spatial trends from treatment and genetic effects rely on the use of autoregressive correlation functions defined on rows and columns (AR1×AR1) to model the local trends (Gilmour *et al.*, 1997) . These models are difficult to fit and the selection of a best model is complicated therefore preventing an automated phenotypic analysis of series of trials. An attractive alternative is the use of 2-dimensional P-spline surfaces, the SpATS model (Spatial Analysis of Trials using Splines, (Rodríguez-Álvarez *et al.*, 2018). This model corrects for spatial trends, row and column effects and has the advantage of avoiding the model selection step. It also provides the user with a graphical output that is easy to interpret (Figure 7). The SpATS model was applied to several datasets generated within the consortium. The SpATS analyses clearly showed the existence of spatial trends on platforms. Experimental designs need to take such trends into account.

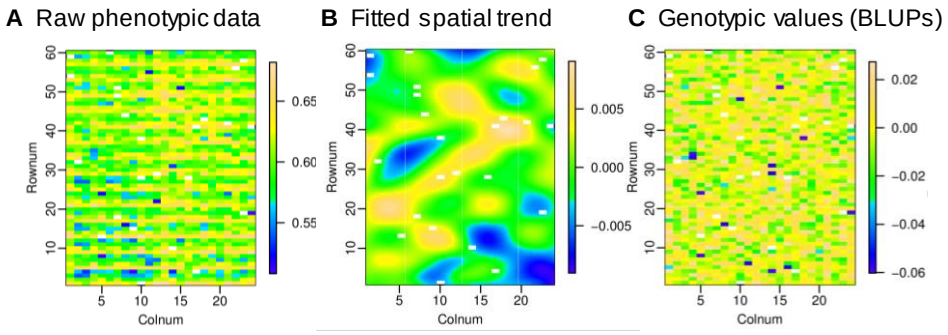

Figure 7. Graphical output of the SpATS model fitted on data coming from an experiment at the Phenovator platform (WUR, Netherlands). Raw plot data (A) are adjusted for experimental design factors and spatial variation (B). The results are adjusted genotypic means (best linear unbiased estimators, BLUES) or predictions (best linear unbiased predictors, BLUPS) (C). *Courtesy: Mark Aarts and René Boesten.*

Platform measurements are often taken over time and then it is important to acknowledge that spatial trends and patterns will change over time. We investigated the changes in spatial trends in time series data from platforms by fitting SpATS models to individual time points and then comparing the fitted spatial trends over time as well as modelling the genetic signals as a function of time (Figure 8). An important conclusion was that methods to correct for spatial trends in field trials are also effective for the correction of spatial trends in platform data. Genetic predictions become more accurate and knowledge of fitted spatial trends helps to choose better experimental designs for future phenotyping experiments.

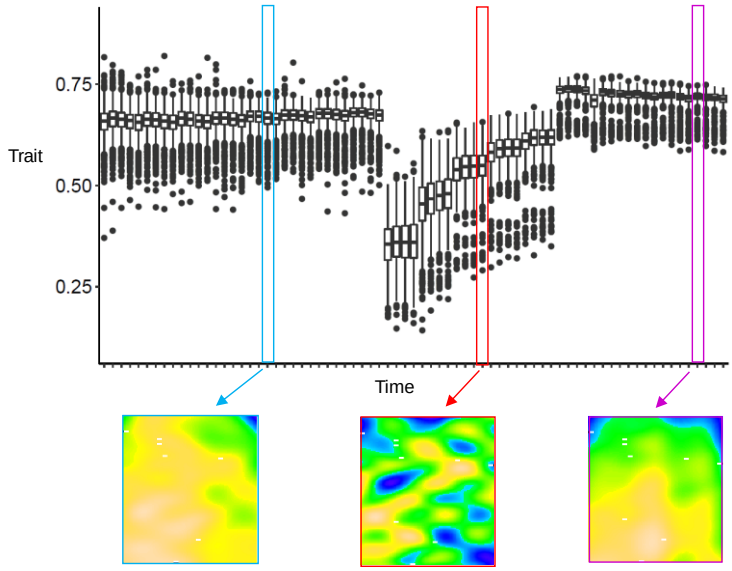

Figure 8. Raw plot data as boxplot per time point and spatial variation at three time points. The raw data are adjusted for experimental design factors and spatial variation per time point using the SpATS model as described in (van Eeuwijk *et al.*, 2019). Data coming from an experiment at the Phenovator platform (WUR, Netherlands). *Courtesy: Mark Aarts and René Boesten.*

4. CONCLUSION

Platform descriptions with sources of error variation were created by visits to installations, discussion with installation managers and a survey. Most installation managers know the sources of error variation on their installation. Some of them use information on direction and magnitude of error trends to improve their experimental designs. In WP1 of EPPN2020, installation managers are asked to quantify the environmental error variability by mapping environmental gradients on the coordinates of their platform. This mapping will help in defining complete and incomplete blocks in experimental designs. Similarly, use of the automated statistical analysis following the SpATS procedure produces spatially adjusted treatment means alongside with spatial trend diagnostics that are useful to improve experimental designs for future phenotyping experiments.

The inventory of installation layout and experimental designs performed in WP2 serves as input to a design generator tool that will become available to platform users within EPPN2020. An application was designed and prototyped to provide an API driven cloud-based app within a micro-service framework to generate statistical designs. An initial user interface was implemented to generate row-column and RCB designs and will be extended with the p-rep and augmented row-column designs. The app allows parameters relating to a platform to be provided to help in the construction of the design. Together with an R package built to run a spatial model over time, it will enable the data analysis in an automated, reproducible and traceable way.

Acknowledgement: This document was written by Emilie Millet (WU) and Fred van Eeuwijk (WP2 leader, WU), with contributions by Martin Boer (WR), Darren Murray (VSNi) and Maria Xose Alvarez-Rodriguez (BCAM).

References

Cabrera-Bosquet L, Fournier C, Brichet N, Welcker C, Suard B, Tardieu F. 2016. High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. *New Phytologist* **212**: 269-281.

Cullis BR, Smith AB, Coombes NE. 2006. On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological, and Environmental Statistics* **11**: 381-393.

Gilmour AR, Cullis BR, Verbyla AP, Verbyla AP. 1997. Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. *Journal of Agricultural, Biological, and Environmental Statistics* **2**(3).

Piepho H-P, Williams ER. 2016. Augmented Row–Column Designs for a Small Number of Checks. *Agronomy Journal* **108**(6).

Piepho H-P, Williams ER, Michel V. 2015. Beyond Latin Squares: A Brief Tour of Row-Column Designs. *Agronomy Journal* **107**(6).

Rodríguez-Álvarez MX, Boer MP, van Eeuwijk FA, Eilers PHC. 2018. Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spatial Statistics* **23**: 52-71.

van Eeuwijk FA, Bustos-Korts D, Millet EJ, Boer MP, Kruijer W, Thompson A, Malosetti M, Iwata H, Quiroz R, Kuppe C, et al. 2019. Modelling strategies for assessing and increasing the effectiveness of new phenotyping techniques in plant breeding. *Plant Sci* **282**: 23-39.

Glossary

CRD – Completely Randomized Design
EPPN2020 - European Plant Phenotyping Network – 2020
p-rep design – partially replicated design
RCBD – Randomize Complete Block Design
RCD - Row-Colum Design
SpATS - Spatial Analysis of Trials using Splines
TNA – Trans-National Access
WP – Work Package