KU LEUVEN

FACULTY OF SCIENCE

Comparison of statistical methods and designs for a high throughput phenotyping experiment

Alexandre BOHYN

Supervisor: Prof. P. Goos
KULeuven

Mentor: Pr. X. Draye
UCLouvain

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Statistics

Academic year 2018-2019

Preface

The basis for this thesis stemmed from my desire to join knowledge from my previous master in bio-engineering to my newly acquired statistical analysis skills. Instead of choosing a pre-made thesis topic, I decided to reach out to different professors both in my current and former university, to create something unique. My goal was to find an interesting topic that would spark my interest and motivate me to work hard. Students in the master of statistics are often given datasets to analyse in the framework of their thesis. However, I wanted to give more than just a context to my dataset. For this reason, I decided to set up my own experiment, harvest the data and analyse them with a goal in mind. It took more time than expected and I had a few bumps along the road, but it was worth it. It allowed me to see the bigger pictures of scientific research. I realized that data analysis and statistical thinking are everywhere, from the pre-conception of an experiment to the detailed analysis of its results.

In truth, I could not have achieved this thesis without the essential help of some key people. First of all, I want to sincerely thank Professor Peter Goos from KULeuven. He agreed to be my promoter when I approached him with a personal thesis topic, and ever since he has been guiding me throughout this project, giving key insights and great support. He helped me conceiving the experiment, and advised me wisely on data analysis and interpretation. Then, I would like to thank Professor Xavier Draye from UCLouvain. Not only did he agree to let me use the phenotyping platform in the UCLouvain greenhouses, he especially helped me setting up and carrying out the experiment correctly. Finally, I would like to thank Emilie Millet and Professor Fred van Euwijk from the University of Wageningen. They were at the base of this project and were the one that helped me creating and defining the goals of this master thesis. Emilie and Xavier were also my main advisors for the phenotyping experiment. Without their support and the collaboration of the INRA in Montpellier, the data used in this thesis simply would not be here today. Finally I could not have achieved this work without the strong and constant support of my family. Thank you all for your unwavering support.

Summary

The rising threat posed by climate change and overpopulation has put food security as one of the major concern for the next decades. Plant breeding has been seen has one of the solution to take care of this global issue. In the recent years, large amount of progress has been made in genetic editing and sequencing techniques, to improve plants yield and make them more resistant. However, to fully take advantage of those innovations, similar progress needs to be made in plant phenotyping. Indeed, there is currently a bottleneck created by the lack of efficient high-throughput phenotyping platforms, to link the genetic data to useful traits in plants. While these platforms are slowly emerging, the necessary tools to analyse their data are not ready yet.

Over the years, a lot of complex experimental designs for field trials have been developed to better account for spatial variability. However, in practice, experimenters are often not using those designs because of their complexity and their difficulty of interpretation. In parallel, different kind of spatial models have been developed, also to account for variability in the data. Two categories of models are standing out: models based on the modelling of the spatial covariance structure, and other based on the modelling of spatial variation using polynomial splines. In the latter category, the SpATS model was recently created and showed promising results in the analysis of both simulated data and real trials.

Because of the need to constantly characterize plant growth, phenotyping platforms often involve moving plants. Since most of these platforms are relatively new, the impact of movement has not yet been well studied. In this thesis, we aimed at characterizing the effect of movement on plant growth, in the phenotyping platform located in the UCLouvain greenhouses. Another goal was to analyse the efficiency of two different spatial models, to account for the spatial variation on the platform, and also to estimate the genotypic effects. For this purpose, we created a custom experimental design, fitted to the platform setup and conducted an experiment including 30 different genotypes of maize. We split the seeds between a moving tank and a still tank. After the experiment, we harvested the plants and used their dry and fresh weights as variable for our spatial models. We tested the SpATS model against a standard spatial model, using auto-regressive processes ($AR(1)$) and linear variance structures (LV) to model the spatial covariance structure. We compared the models on their estimation of the genotypic effects, on the fitted values they provided and on the way they modelled spatial trends.

The experimental results displayed a strong difference between the two tanks and between genotypes and a lot of spatial variability. Both models proved the effect of movement to be highly significant on plant growth. They were also able to capture the difference in genotypes correctly. The estimates of the genotypic effects were correlated to more than 99% between the two models. The spatial variations were well accounted for in both models, as they gave very similar fitted values. However the spatial trends were more smooth in the still tank than in the moving tank. The main difference between the models was the parametrization. SpATS uses the concept of effective dimensions to assess the relative contribution of each component to the modelled spatial surface. This allows an easy interpretation of the directions and intensity of the spatial trends.

These findings proved that genotypes differences can be correctly estimated on the UCLouvain phenotyping platform, as the two models showed satisfying results. The difference between the tank also had a significant influence on the plant weights. The moving tank was a better environment for plant growth. However, even in greenhouses, growing conditions remain highly variable and the efficiency of the models cannot be fully assessed with a single trial. Overall the SpATS model was still better in terms of interpretation and adaptability to highly heterogeneous environment. It showed promising results and a great potential for spatial data analysis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the recent years, food security has grown to become a worldwide concern due to the challenges imposed on farmers by climate change. In this respect, genetic improvement to increase crop resistance to abiotic stresses is seen as a solution to generate novel traits in plants and help fight those threats (Tester & Langridge 2010). In the last decades, genetic editing and sequencing techniques have greatly improved, leading to a high-throughput of genetic data and associated data analysis techniques (Schiml & Puchta 2016). However, to capitalize on those discoveries, similar advancements need to be made regarding plant phenotyping. Even though many breakthrough regarding sensors and imaging techniques are made, plant phenotyping still remains a challenge (Furbank & Tester 2011). Due to the extreme sensitivity of plants to their growing environment, a large amount of the research is dedicated to study plants' variations that are unrelated to their genetic traits. Researchers are setting up high-throughput platforms, where thousands of plants are sequentially analysed, to be able to quantify the influence of environmental conditions on plant growth (Tardieu *et al.* 2017).

In parallel, plant breeding trials often involve a large number of genotypes and large areas where spatial variation is likely to be an obstacle to reliable genetic prediction. To account for complex spatial variations, researchers often use spatial analysis methods that model the correlation between neighbouring plots (Velazco *et al.* 2017). There is a large array of different methods such as nearest neighbours analysis (Wilkinson *et al.* 1983), spatial covariance structures (Gilmour *et al.* 1997; Piepho & Williams 2010), polynomials models (Federer 1998) or smoothing splines (Durban *et al.* 2001). However, for these approaches to be efficient, they require a complementing experimental design. Over the years, several types of complex experimental designs have been created and tested in field trials (Cullis, Smith, *et al.* 2006; Patterson & Hunter 1983; Yates 1939). However, even though those designs allow good correction for field trends, they are rarely used in practice because of their complexity and the specific cases to which they apply. This is even more relevant for phenotyping platforms, where the experimental set-up is rarely like the one of a field.

With this background in mind, the motivation behind this thesis was to design and conduct an experiment in a phenotyping platform and then analyse the results using different spatial models. We created an optimal design, fitted for the platform, and applied in experiment with different genotypes of maize. We then harvested the plants

and analysed the weights using different spatial models, to compare their efficiency on this given platform. The fact that we over-viewed each step of the experiment from the seeds germination to the analysis of the results took a large amount of time but allowed more control over the different phases of the process. On this platform, plants are constantly moving to be characterized, which renders the spatial trend quantification complicated. Even though this feature is often available in platforms, it is poorly evaluated. Therefore, to make this thesis relevant we wanted to evaluate the impact of constant movements on the plants growth.

To summarize, this master thesis can be divided in two sections. The first one is about creating an experiment adapted to the platform's characteristics to evaluate the impact of movements on plant growth, as well as the effect of each genotype. The second part is about fitting different spatial models to see which one is better able to capture, and explain, the spatial trends in the data. We do not aim at creating new spatial models or revolutionizing experimental design, but rather to apply those concepts on a practical case, to see how well they allow the analysis and interpretation of phenotypic data. The following sections are a literature review of the topic; then a development of the material and methods used (that is the experimental procedure and the spatial models); a presentation and discussion of the results; and finally a conclusion to discuss if and how the goals, presented in this introduction, were met in the thesis.

Chapter 2

Literature review

2.1 Plant phenotyping

The terms phenotype and phenotyping are often interpreted in diverse ways between authors and between studies. In order to avoid any confusion, it is important to define these concepts clearly. Plant phenotyping is defined as the identification of effects on the phenotype (i.e., the plant appearance and performance) as a result of genotype differences (i.e., differences in the genetic code) and the environmental conditions to which a plant has been exposed (Fiorani & Schurr 2013; Houle *et al.* 2010). In this thesis, we refer to phenotyping more precisely as the set of methodologies and protocols used to measure plant growth, architecture, and composition with a certain accuracy and precision at different scales of organization (Fiorani & Schurr 2013).
Plant phenotyping is an important tool to address and understand plant environment interaction and its translation into application in crop management practices, effects of biostimulants, microbial communities, etc... (Pieruschka, Schurr, *et al.* 2019). In our current society, food security is a rising issue and genetic crop improvement is seen as a solution to deal with this problem. While genetic editing techniques and genome mapping technologies are blooming, they depend on a similar improvement in phenotyping, since they are key to analyse plant responses to environmental characterization. In recent years, high-throughput and high-resolution phenotyping tools have made impressive progress and can now help relieving the current phenotyping bottleneck (Fiorani & Schurr 2013; Furbank & Tester 2011; Tardieu *et al.* 2017). Different phenotyping platforms are emerging around the world. They range from high-precision platforms for cell and organ characterization (Vargas *et al.* 2006) to multi-environment networks of fields, exploiting remote sensing (Virlet *et al.* 2017). A typical phenotyping experiment generates a large amount of raw data that provides a condensed set of multi-dimensional information (2D usually, but 3D scanning platforms are developping (Mooney *et al.* 2012)). Futhertmore, a lot of tools are available for data analysis in phenotyping platforms (Lobet, Draye & Périlleux 2013). This makes the choice complicated for an external user, especially since most of these softwares are designed for a single specific purpose. Another challenge in phenotyping, is root system architecture (RSA) characterisation because of the inherent complexity of the system. Different techniques have been developed to best characterize the RSA in a cost-efficient way (Lobet & Draye 2013; Pound *et al.* 2013). However, at all scales, phenotyping facilities display spatial heterogeneity that needs to be separated from the genetic signal. For

example, the spatial variability of incident light raises up to 30% between pots within a glasshouse or a growth chamber (Cabrera-Bosquet *et al.* 2016). There is also evidence of microclimate variation in greenhouses experiments (Brien *et al.* 2013). Therefore, correcting for spatial trends and using appropriate experimental designs is crucial for a precise estimation of genetic effects. Hence, the existing design and modelling theory for field experiments needs to be adapted for the new emerging phenotyping platforms.

2.2 Experimental design in field trials

Experimental field trials in agriculture have always been affected by soil heterogeneity. As Van Es (2002) explains, soil is a continuum with variability on multiple scales. The heterogeneity is as much affected by microscopic interactions as by field-sized effects. Therefore, agricultural trials have always heavily relied on randomisation, blocking and replication to account for spatial variability and remove bias from the estimation of the treatment effects (Atkinson & Bailey 2001). For randomisation to be truly effective, stationarity of the mean and spatial independence assumptions need to be verified. Several studies have proven that it is rare that both these assumptions hold in field trials (Davidoff *et al.* 1986; Iqbal *et al.* 2005; Nielsen *et al.* 1973). Moreover, Van Es (1993) showed that even randomized designs can still be problematic for experiments with large numbers of treatments and low numbers of replications in the presence of spatial autocorrelation. A new class of design has been proposed involving the use of replicated plots for a percentage of the test lines: the “p-rep” designs (Cullis, Smith, *et al.* 2006; Velazco *et al.* 2017). Local field trends can influence groups of treatments in specific blocks. As a solution, several authors (Fagroud & Van Meirvenne 2002; Watson 2000) have suggested considering the spatial trends and autocorrelation structures when creating the design, by using prior soil information, but taking into consideration spatial variability in the design of a trial not only require previous information on the plot but is often costly and cumbersome. Furthermore, in practice, most experimenters have neither the capacity to implement advanced designs (in terms of computation power and statistical training), nor the capacity to analyse them. Finally, Van Es *et al.* (2007) showed that completely randomized (43 % in greenhouse trials) and random block designs (70 % in field trials) are still widely used. Considering this global issue, finding and using an appropriate design is complex task.

2.3 Spatial modelling for field trials

In order to increase the precision of the estimation of genetic effects, experimental designs need to be complemented with appropriate models of analysis. Mixed model analyses using the autoregressive ($AR(p)$) functions (Cullis & Gleeson 1991) have become a standard strategy in field trials. However, Piepho, Möhring, *et al.* (2015) recently discussed several issues with this model and have therefore proposed the use of the linear variance (LV) model (Williams & Luckett 1988) instead. More specifically, Piepho & Williams (2010) have proposed a revised version of this model, augmenting

it into two dimensions ($LV \times LV$). The main novelty resides in the addition of spatial components to a classic rows-columns model. Recently, Rodriguez-Alvarez *et al.* (2018) introduced a novel spatial model that adjusts for both global and local trends simultaneously: the SpATS model (Spatial Analysis of field Trials with Splines). The new spatial method makes use of penalized splines (Eilers & Marx 1996) to estimate a bivariate smooth function over the rows and columns of a plot. Using the work of Lee Hwang (2010), Lee & Durbán (2011), and Lee, Durbán & Eilers (2013) the spatial variability is characterized using tensor products of two-dimensional P-splines (Dierckx 1995) and decomposed in a PS-ANOVA system. By exploiting the similarities between P-splines and mixed models (Currie & Durban 2002; Durban *et al.* 2001; Wand 2003), the P-splines are expressed as a mixed model, which allows the use of classical mixed-model software but also the use of additional random and fixed effects to the model to better capture the variation along the 2-dimensional field. It has already been tested on simulated data (Rodriguez-Alvarez *et al.* 2018) and previous field trials data (Lado *et al.* 2013) and showed promising results.

As Wilkinson *et al.* (1983) and Gilmour *et al.* (1997) highlight, in field trials data modelling, three main sources of spatial variations need to be accounted for:

Stationary[1] **variations:** Large scale trends across the field (e.g. fertility trend, depth of soil, moisture)

Non-stationary variations: Also natural variations but localized on part of the field (e.g. patch of soil moisture)

Extraneous variations: Variations unrelated to a natural process, often due to the way the field is prepared (e.g. tillage, sowing practices, etc. . .)

A part of these variations can be attributed to systematic effects, e.g. sowing or planting, another part to random effects such as fertility trends. While systematic effects can easily be modelled using factors and row-columns attributes, it is not case the case for random spatial variation. They are harder to model because there are no covariates to relate it to. Since the spatial variation has both random and systematic components, it makes sense to use the mixed model framework.

There are two main approaches to model spatial trends: one based on spatial variance-covariance structures; and the other based on smoothing techniques. The SpATS model uses a smooth bivariate surface to model both the global and local trends, while accounting for the extraneous variations by using extra random and/or fixed coefficients. Models using spatial covariance structure (such as the $LV \times LV$) model the global trends using functions of the spatial coordinates (both linear trends and smoothing splines), while the local trends are estimated with the use of spatially dependant error term (thus the reason why these models use spatial covariance structures) and the extraneous trends are managed similarly to the SpATS model. In this thesis, the data extracted from the phenotyping platform are modelled using these 2 different kinds of models.

[1] A stationary process has the property that the mean, (co)variance and autocorrelation structure do not depend on spatial location (Risser 2016).

2.4 Thesis objectives

This master thesis falls within the scope of the second research activity of the European project EPPN^{2020} [2]. It is a research infrastructure project funded by Horizon 2020, that will provide access to 31 key plant phenotyping installations. It defines three research activities: (1) novel technologies and methods for environmental and plant measurements, (2) innovative design and analysis of phenotyping experiments across multiple platforms and (3) a European plant phenotyping information system. The project revolves around data acquisition, data analysis and data networking, so that every platform uses common, standardized practices and analysis protocols, that have been tested for robustness and quality.

The main goal was to assess the utility of statistical designs and mixed models to identify and correct for spatial trends (heterogeneity) in an aeroponic root installation at UCLouvain (Louvain-la-neuve). The idea was to set up an experiment in this installation using different genotypes (plant varieties) and a custom experimental design to account for possible complex environmental variations. The design was created using JMP®(Version 14.3, SAS Institute Inc., Cary, NC, 1989-2019.), taking into account the specificities of the platform and the number of genotypes used. This approach allows the design to fit the experiment properly and avoids having to use a pre-made design, that may not be optimal for the experiment. After data collection and image analysis, two different models will be used to model the spatial variability and to assess the quality of spatial prediction. The first one is a two-dimensional version of the linear variance model, revised by Piepho & Williams (2010). The second one is the SpATS model, recently created by Rodriguez-Alvarez *et al.* (2018). The two models will be compared in term of their ability to estimate genotypic effects and to quantify spatial variability. These comparisons will be made using classical indicators (RMSE, ...) and other indicators, specific to spatial models for field trials (Oakey *et al.* 2006).

The experiment took place in February 2019 in the UCLouvain greenhouses. The installation consists of two aeroponic tanks of 495 plants located in a 64 m^2 greenhouse. Plants are held on strips, 5 plants per strip, 99 strips per tank. The specificity of the platform is that the plant rotate constantly so that their root system can be photographed every two hours. The experiment lasted 3 weeks, after which the plants became too large for the platform. The experiment included two tanks. In the first one, plants constantly moved to be pictured every 2 hours (usual set-up on this platform). In the second one, plants moved twice or three times a day to be pictured. This allowed comparing the effect of moving versus non-moving plants, which is a feature often available in the phenotyping platforms but poorly evaluated so far.

Since the UCLouvain platform focuses on the analysis of the root system, the main variable of interest in the experiment is the overall growth of the root system of each plant. Scientists of the UCLouvain platform have developed pipelines[3] that allow easy processing of the images captured in the platform to extract quantitative root architec-

[2] European Plant Phenotyping Network 2020 `https://eppn2020.plant-phenotyping.eu/`

[3] Here, pipelines are defined as computer programs designed to analyse raw data from phenotyping platforms.

ture information for the spatial models (Lobet & Draye 2013; Lobet, Pagès, *et al.* 2011). However in this thesis the main variable of focus will be the weights of the plants. Since they are a direct indicator of the overall growth.

This thesis was divided in four main parts: creating an appropriate experimental design for a phenotyping experiment and conducting the experiment; analysing data from a high-throughput platform; comparing the efficiency of various spatial models to correct for heterogeneous spatial trends and estimate genotypic features; and developing the appropriate R scripts.

Chapter 3

Material and methods

3.1 Optimal experimental designs

In the context of design of experiments (DOE), experimenters are always looking for the ideal design. However, researchers often use pre-made designs that fit many experiments instead of creating an optimal one for each problem at hand. Since experiments exist in all sizes and forms, not all pre-made designs are ideal. Therefore, creating an optimal design is a great choice to ensure that the design fits the experiment and not the other way around. In this section, we explain how and why we created a custom, optimal, design to fit our experiment.

3.1.1 Orthogonal designs

Orthogonal designs are interesting because they guarantee that each main effect and interaction can be estimated independently. Meaning that the effect of one factor or interaction can be estimated separately from the effect of other factors and that the addition or removal of terms in the model does not affect the estimates. In a regression or ANOVA-type model, the best linear unbiased estimator (BLUE) of the regression coefficients is obtained by using the ordinary least squares (OLS) method, because it minimizes the variance of the estimators. These variances are often represented in variance-covariance (VCOV) matrix, where the diagonal is the variance of the estimators and the non-diagonal elements are the pairwise covariances between estimators. The inverse of this matrix is the information matrix, because it summarizes the available information on the models coefficients (a low variance means a lot of information, and inversely). As detailed in Goos & Jones (2011), when the information matrix is diagonal, then the design is said to be orthogonal. Many standard designs are orthogonal but they often do not fit the experiment at hand.

3.1.2 Optimality criteria

To compare different designs, the two main criteria are the D-optimality and the I-optimality. The first one aims at minimizing the variance of the factors effects estimates and is more useful for significance testing. D-optimal designs are therefore more ap-

propriate for screening experiments[1]. The second one aims at minimizing the average relative prediction variance over the experimental region. I-optimal designs are focused on prediction and thus are more suited to response surface experiments. There also exists a G-optimality criterion that is similar to the I-optimality criterion but minimizes the maximum prediction variance. Recent work (Rodríguez *et al.* 2010) has shown that I-optimal designs are often better choices than the G-optimal ones. Since the experiment discussed in this thesis is a screening experiment, only the D-optimality is detailed here. More information about I-optimal and G-optimal design is available in Atkinson (2014) and Goos & Jones (2011).

The elements of the information matrix are inversely proportional to the variances and covariances of the models parameters. Therefore, the design with factor settings that maximize the determinant of information matrix, will maximize the available information about the models parameters. This design is called the "D-optimal design", where the "D" stands for determinant and the value of the determinant itself is called the "D-optimality criterion".

For any model with two-levels factors and two-factor interaction effects, orthogonal designs will always be D-optimal. However if the number of runs is not a multiple of 4 then there are no orthogonal designs available for two-level factors. This condition offers little flexibility for experimenters and is not always feasible. In contrast, the optimal experimental design approach allows for any number of runs. However, in non-orthogonal designs the variance of the estimates is inflated and the estimates are correlated. Nevertheless this inflation is usually small and the correlation is too small to cause any concerns. Therefore there exist non-orthogonal designs that still maximize the information of the model being estimated. The D-optimal designs may not be unique. For a specified number of runs, several designs might have the maximal value for the determinant of the information matrix.

3.1.3 Generating optimal designs

Several algorithms exist to generate optimal designs, but the most common one is the coordinate exchange algorithm, created by Meyer & Nachtsheim (1995). It has the advantage to run in polynomial time, which means that the time needed to find an optimal design does not explode when the size of the design increases. Another similar algorithm is the point-exchange algorithm, created by Fedorov (1972) and modified several times to speed it up (Atkinson & Donev 1989; Johnson & Nachtsheim 1983). The main drawback of this algorithm is that it needs a list of possible design points as input, which can be quite tedious to do for large designs. In recent years, other types of algorithm such as genetic algorithms (Heredia-Langner, Carlyle, *et al.* 2003; Heredia-Langner, Montgomery, *et al.* 2004), simulating annealing algorithms (Bohachevsky *et al.* 1986;

[1] Screening experiments are experiments designed to evaluate the significance of factors and factor-interactions in a model. The factor are usually two-level factor because they are either present in the model or not. They are opposed to response surface (RS) experiments that are designed to find the optimal settings for the factors. In RS experiments, the significant factors of the model are already determined.

Meyer & Nachtsheim 1988) and tabu search algorithms (Jung & Yum 1996) have been used in experimental designs. While these algorithms maintain a level of performance comparable to more traditional design construction techniques, they are not as popular because they are either far more complex, only feasible in some specific cases or better for some specific models and do not lead to designs that make a significant difference in practice.

The coordinate-exchange algorithm proceeds by iterating through the rows of the matrix of factors settings

$$\mathbf{D} = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{k1} \\ x_{12} & x_{22} & \dots & x_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1n} & x_{2n} & \dots & x_{kn} \end{bmatrix}, \tag{3.1.31}$$

called the design matrix, of an experiment with n runs and k factors. The rows of this matrix essentially represent the coordinates of the runs in the experimental space, where each factor of the experiment is a dimension. This algorithm is called the coordinate-exchange algorithm, because in each iteration of the algorithm, possible changes for every element of the design matrix are considered.
It can happen the D-optimality criterion is zero, in those cases, the design is called singular. To avoid singularity, the number of design points (different rows in the design matrix D) must be greater than or equal to the number of model parameters.

The algorithm starts by generating a random design. For all continuous factors, the algorithm generates random values on the interval $[-1, +1]$. For all factors that are categorical, the algorithm randomly chooses a value in a discrete set of levels. This random starting design is almost always non-singular. If the design happens to be singular, then another new random starting design is computed.

In the next step, the algorithm improves the design on an element-by-element basis. For each element of the starting design, x_{ij}, a change to either -1 or $+1$ (or to a value in the starting interval for non-quantitative values) is considered, and its impact on the D-optimality criterion is evaluated. The change that increases the value of the D-optimality criterion the most, is kept. After investigating changes in each element of the design, the process is repeated until no element changes within an entire iteration through the factor settings or until a prespecified maximum number of iterations is reached. The obtained design is the best among a set of neighbouring designs but it is often a locally optimal design that is different for each random starting design. To select the best among all locally optimal designs, the algorithm is repeated a large number of times. The globally optimal design is then selected among all the locally optimal ones, as the one that yields the highest D-optimality criterion.

3.1.4 Generating the design

For our experiment, a custom design was created for the phenotyping platform, where the goal was to quantify the genotype and tank effect. Four factors were considered:

Tank In which tank was the plant situated (moving or still).

Strip Which of the 99 strip was used (1 to 99).

Position What was the position on the strip (1 to 5).

Genotype Which one of the 30 genotypes was used (1 to 30).

To create the design, the design of experiment (DOE) tool of JMP was used. The four categorical factors were specified and *Tank* and *Strip* were set to "very hard to change" and "hard to change", respectively. Two whole plots of 99 sub-plots each were specified to match the tanks and the strips. With 99 strips of five positions inside two tanks, 990 experimental units were available.

Initially the design was supposed to take into account the 99 different strips individually but the program couldn't converge to an optimal design because of its complexity. Instead only 33 strips were considered and the design was replicated 3 times to match the number of strips available. Figures 3.1 and 3.2, display a schematic view of the design for the moving and still tank respectively.

The seeds were provided by researchers from the french national institute of agronomic research (INRA). A subset of 30 maize hybrids was chosen among an historical series[2] of maize hybrids that were characterised by their commercial success at the time of their release. Claude Wlecker (LEPSE, INRA) and Carine Palaffre (SMH Maïs, INRA) prepared and sent a total of 30 seeds per genotype plus an extra 150 seeds of genotype called "border genotype" hereafter, used to fill the gaps in the platform left by non germinated seeds and the 90 empty spots left. This genotype is not part of the historical series and was therefore not considered in the design of the experiment.

3.1.5 Design characteristics

In the planned design, each genotype was replicated 33 times on the platform with 15 or 18 replication in each tank. The genotypes were replicated 3, 6 or 9 times in each position of a strip, with usually at least 3 replication in each tank. Figure 3.3 shows a graphical representation of the planned repartition of genotypes among positions and tanks. It also depicts the actual repartition that was set up on the platform. No genotype stands out as poorly represented over the whole experiment. However, the border genotype (number 31) filled several empty spaces in the platform, more for the still tank than for the moving tank.

[2]Historical series correspond to varieties that have been cultivated and bred for some time, mainly due to their physiological specificities.

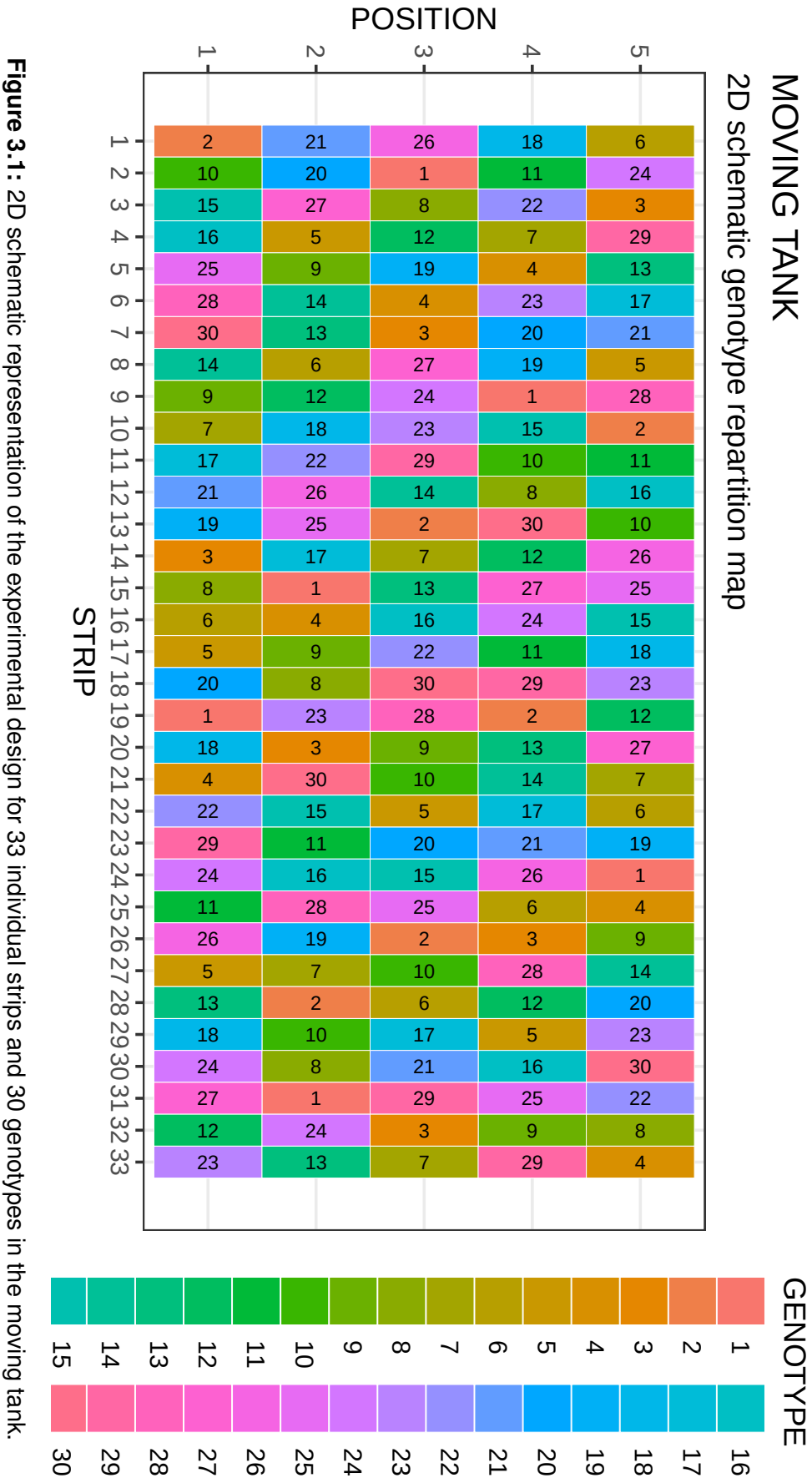


Figure 3.1: 2D schematic representation of the experimental design for 33 individual strips and 30 genotypes in the moving tank.

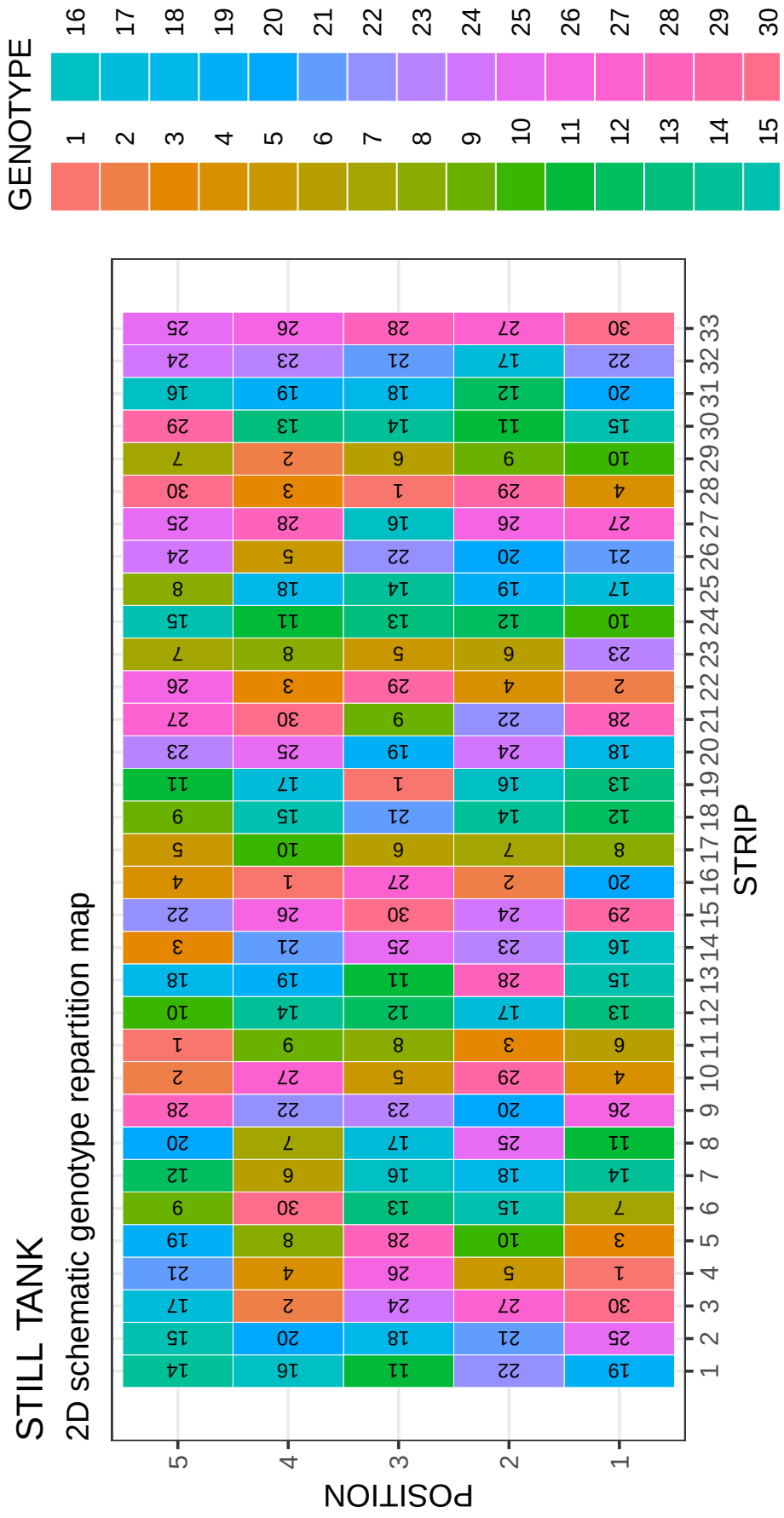


Figure 3.2: 2D schematic representation of the experimental design for 33 individual strips and 30 genotypes in the still tank.

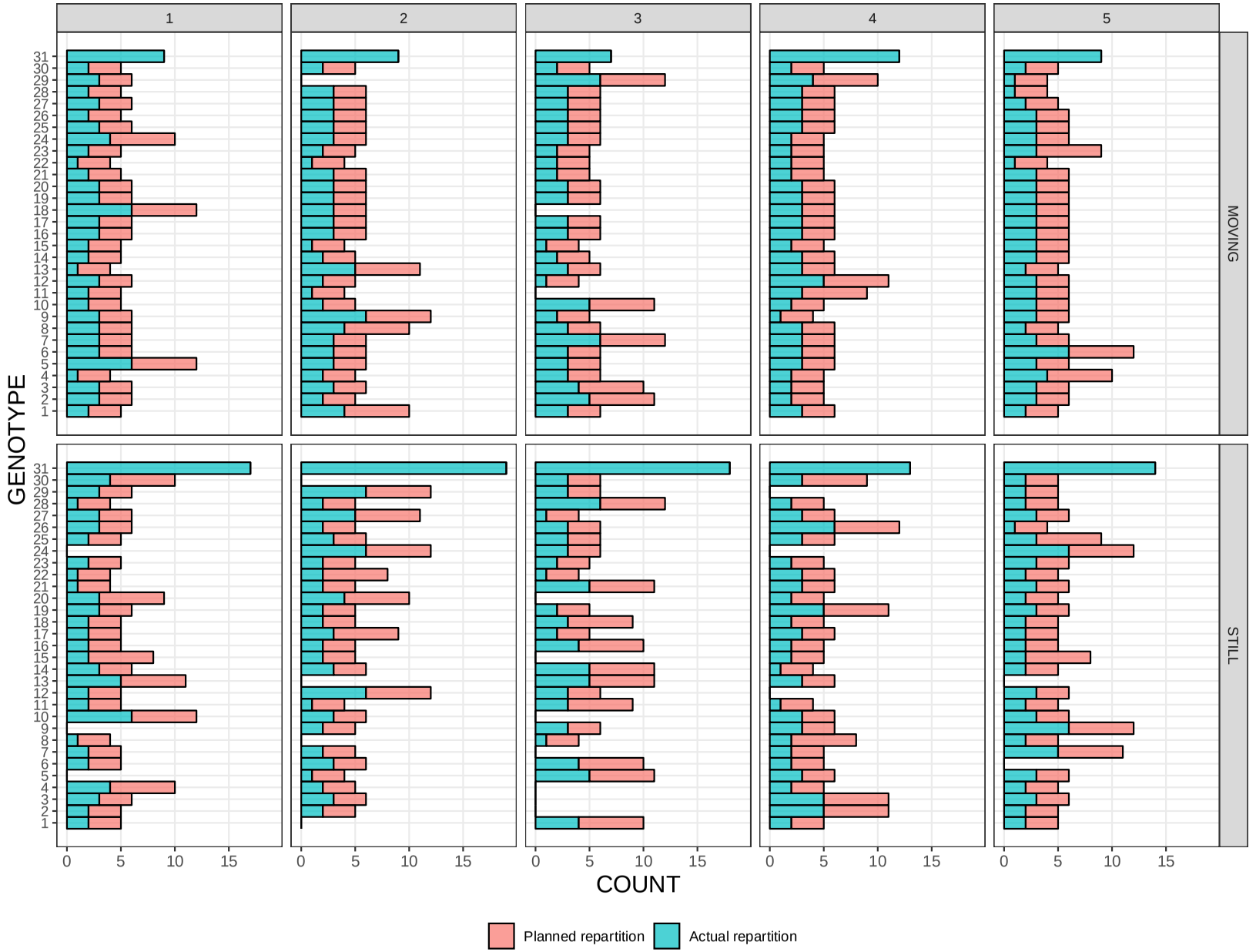


Figure 3.3: Genotype counts for each position in each tank. The red bars represent the planned repartition, while the blue bars represent the actual repartition.

3.2 Phenotyping experiment

The phenotyping experiment took place between February 25th and March 13th 2019. The seeds were first germinated and then transferred onto the platform. After the end of the experiment the plants were weighted, dried and weighted again to obtain dry and fresh weights.

3.2.1 Germination

Previous experiments in the greenhouses showed that germination of maize seeds on the platform often lead to asphyxiation of the seeds. Because of this, the seeds were germinated in an outside germination chamber and were later transferred onto the platform.

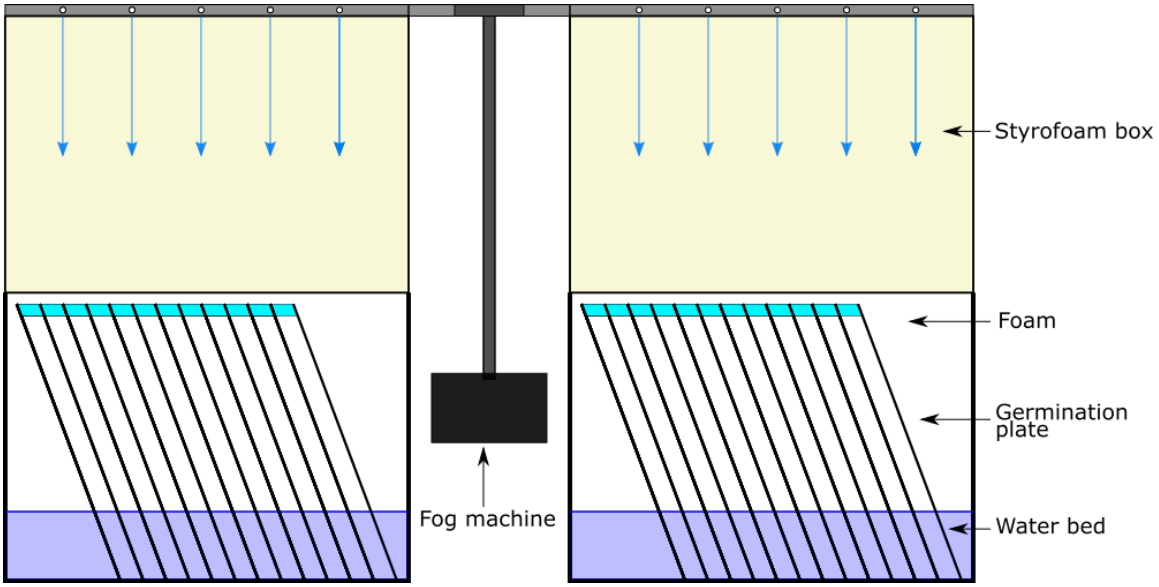
The seeds were placed inside a temperature-controlled germination chamber at 20°C for 3 days. The chamber consisted of 2 PVC trays to which an air-fog machine was connected to keep the seeds moist. Inside each tray, PVC plates were disposed diagonally and evenly spaced (fig. 3.4a). On those plates, the seeds were arranged on a filtering paper sheet with ledges to support the weight of the seeds (figure 3.4b and figure 3.4c). The bottom of the trays were filled with water to keep the filtering paper moist. There were 17 plates in total, 15 for the 30 genotypes and 2 for the border genotype (150 seeds dispatched on 2 plates).

3.2.2 Phenotyping platform

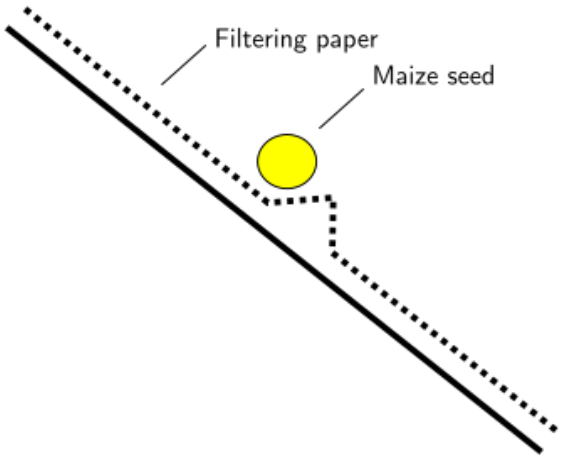
The phenotyping platform is located inside a greenhouse in the facilities of the UCLouvain (Louvain-la-Neuve, Belgium). It consists of two aeroponic tanks on which are arranged 99 styrofoam strips, each with five holes. At the end of each tank is a high definition camera that scans the root system of each plant individually, when it passes in front of it (fig. 3.5a). The strips rotate in a clockwise fashion in the tank and a full rotation is completed in 2 hours. Three sprinklers are placed regularly at the bottom of each side of the tank (fig. 3.5b). The sprinklers spray nutrient solution[3] at regular intervals, set by the operator. The spraying pattern (interval and duration) can be differentiated between day and night and can be modified at any moment of the experiment. In this case the pattern was 5 seconds of spraying ever 295 seconds all the time. During the experiment, the temperature of the greenhouse was set to 20°C at day and 18°C at night and the lights were on from 6 AM to 10 PM.

After 3 days in the germination chamber, the germinated seeds were transferred onto the platform following the created design, and the non-germinated seeds were discarded. The seeds were placed inside a foam cork and then placed inside a hole on a strip (fig 3.5c). They were placed at the bottom to allow the root system to grow freely. The corks were drilled vertically to let the leaf system develop with less resistance and allow a direct access to sunlight. More information about the platform is available in

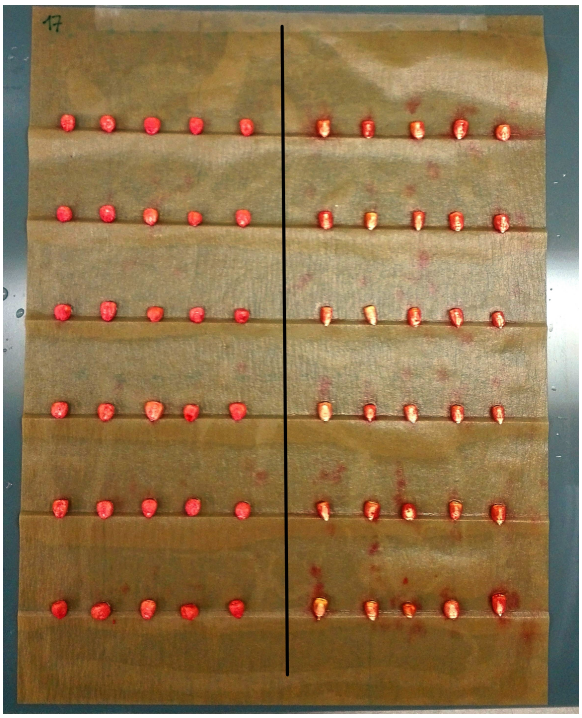
[3] In this experiment, we used Hoagland nutritive solution.



(a) Global schematic view of the germination chamber: a fog machine assure constant humidity in the germination chambers by creating fog at regular intervals (the blue arrows represent the path of the fog). The plates are placed at a 60° angle and 5 cm apart



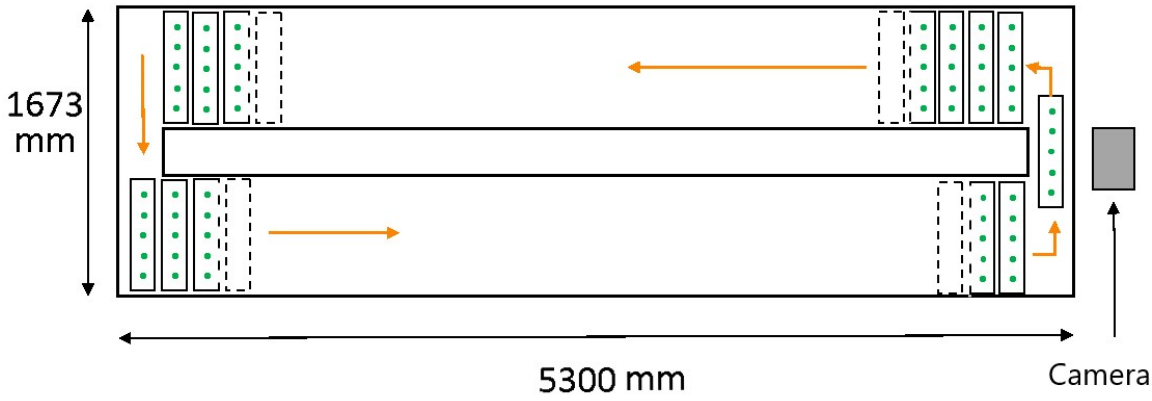
(b) Schematic view of a germination ledge on a PVC plate: each seed is fixed in position on the ledge by an additional drop of agar solution to avoid any fall-off



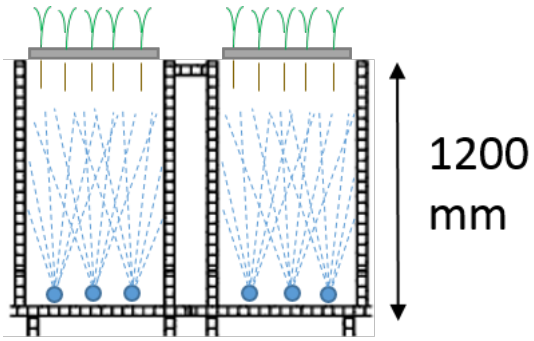
(c) 30 cm by 40 cm PVC plate with seeds on filtering paper (the black line represents the separation between the two genotypes on the plate). Each sheet had 6 rows of 10 seeds with one genotype on the left and one genotype on the right.

Figure 3.4: Germination chamber diagram with detailed view and pictures

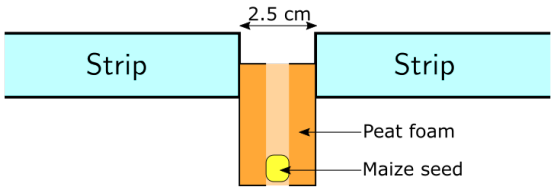
Appendix C.



(a) Schematic view of an aeroponic tank: plants are hold on strips, 5 plants per strip (green dots on layout). There are 99 strips in the tank for a total of 495 plants/tank. Strips move in the direction indicated by orange arrows



(b) Transversal schematic view of an aeroponic tank of the platform: at the bottom of each tank, sprinklers (represented in blue in the layout) are disposed at regular interval and spray nutritive solution



(c) Close up schematic view of a strip: inside each hole, seeds are placed inside a pierced peatfom cork to allow the root system to develop frelly

Figure 3.5: Detailed diagrams about the phenotyping platform

3.3 Data processing

After 15 days, the plants were considered fully grown and the experiment was stopped. The leaf and root systems were separated and weighted individually on scales precise to the milligram. They were then dried for 3 days in a 70°C oven and weighted again. At the end, four variables were kept for the spatial analysis:

- FRESH_RS: fresh weight of the root system
- FRESH_LS: fresh weight of the leaf system
- DRY_RS: dry weight of the root system

- DRY_LS: dry weight of the leaf system

Some seeds placed on the platform did not grow fully or correctly. Therefore some data points needed to be handled more carefully, as they do not represent the genotype growth correctly. However, an outlier is hard to distinguish from a representative data point because of the complex influence of the environment on the growth. Therefore, instead of choosing which plants were outliers in a binary way, we attributed individual weights to express the quality of the data. Those weights were established by reviewing the final root scan of each plant and checking for influential factors among the following:

- NO_RS: no additional root to the primary root
- NO_LS: no visible leaf system
- BAD_LS: leaf system grew under the strip (or abnormally in general)
- NO_SEED: no seed (or an empty cork) present on this position
- NOT_FG: plant not fully grown
- $OVERLAP$: leaf (or root) system of another plant overlaps on the root scan
- OK: no influential factors

An example of each influential factor is presented in figure 3.6. Some plants were attributed several influential factors, but OK, NO_SEED and NOT_FG were considered as exclusive. The factor attribution was ambiguous for some pictures, in those cases the plant were considered OK to avoid losing any data points. Following the determination of the factors, weights were attributed to each variable according to the weight matrix, presented in table 3.1.

Table 3.1: Weight attribution matrix for the different factors and variables. LS weight is both the fresh and dry weight for leaf system and RS weight is the same for the root system

Code	NO_RS	NO_LS	BAD_LS	NOT_FG	NO_SEED	OK	OVERLAP
Weight	1	2	3	4	0	5	5

(a) NO_RS: no additional root, other parent leaf system than the primary

(b) NO_LS: no apparent leaf system

(c) BAD_LS: the leaf system grew under the strip and deviated to the right (obstruction of another plant)

(d) NO_SEED

(e) NOT_FG: both system are present but underdeveloped compared to images

(f) $OVERLAP$: the leaf system of another plant hides the other root system of the plant

(g) OK

Figure 3.6: Example pictures of the influential factors chosen for the weight attribution.

3.4 SpATS model

In this section, the SpATS (Spatial Analysis of field Trials with Splines) model is introduced. For a more thorough treatment of the model and all its components, see Rodríguez-Álvarez, Boer, van Eeuwijk, *et al.* (2016).
Consider a field trial of n plots arranged in a rectangular grid, where the plot positions are collected in vectors of row $(\mathbf{r})$ and column $(\mathbf{c})$ coordinates. If $\mathbf{y}$ is the vector of plot data in field order, a common model for $\mathbf{y}$, to use as a starting point is

$$\boldsymbol{y} = \mathbf{1}_n\beta_0 + \boldsymbol{Z}_r\boldsymbol{c}_r + \boldsymbol{Z}_c\boldsymbol{c}_c + \varepsilon \tag{3.4.01}$$

were $\mathbf{1}_n$ is a column-vector of ones of length n, $\boldsymbol{c}_r$ and $\boldsymbol{c}_c$ are, respectively, the random effect coefficients for the rows and columns and associated matrices $\boldsymbol{Z}_r$ and $\boldsymbol{Z}_c$. To fully capture complex spatial patterns, a smooth bivariate surface jointly defined over the row and column positions is added to the model, which becomes

$$\boldsymbol{y} = f(\boldsymbol{u}, \boldsymbol{v}) + \boldsymbol{Z}_r\boldsymbol{c}_r + \boldsymbol{Z}_c\boldsymbol{c}_c + \boldsymbol{\varepsilon} \tag{3.4.02}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are, respectively, the vector of row and columns positions and where $f(.,.)$ represents the smooth bivariate function. Note that the intercept term, β_0 is embedded into $f(\boldsymbol{u}, \boldsymbol{v})$. To better understand this function, let us decompose it in a nested-ANOVA structure[4] (Lee, Durbán & Eilers 2013):

$$\begin{aligned} f(\boldsymbol{u}, \boldsymbol{v}) = & \underbrace{\mathbf{1}_n\beta_0 + \boldsymbol{u}\beta_1 + \boldsymbol{v}\beta_2 + \boldsymbol{u} \odot \boldsymbol{v}\beta_3}_{\text{Bilinear polynomial}} \\ & + \underbrace{f_u(\boldsymbol{u}) + f_v(\boldsymbol{v}) + \boldsymbol{u} \odot h_v(\boldsymbol{v}) + \boldsymbol{v} \odot h_u(\boldsymbol{u}) + f_{u,v}(\boldsymbol{u}, \boldsymbol{v})}_{\text{Smooth part}} \end{aligned} \tag{3.4.03}$$

where $\odot$ denotes the element-wise matrix product[5]. There are now two components to the model: a bilinear polynomial part (parametric) and a smooth part (non-parametric). The parametric part includes a linear trends along rows (β_1) and columns (β_2) as well as a linear interaction trend (β_3). The smooth part models the deviation from the compound linear trend, and can be decomposed in the following elements:

- $f_u(\boldsymbol{u})$ is a smooth trend along the rows, identical for all columns (i.e., a main smooth effect).
- $f_v(\boldsymbol{v})$ is a smooth trend along the columns, identical for all rows.
- $\boldsymbol{v} \odot h_u(\boldsymbol{u})$ and $\boldsymbol{u} \odot h_v(\boldsymbol{v})$ are linear-by- smooth interaction trends. For instance, $\boldsymbol{u} \odot h_v(\boldsymbol{v})$ is a varying coefficient surface trend, consisting of functions, linear in the rows, for each column, but with slopes that change smoothly along the columns, $h_v\boldsymbol{v}$).
- $f_{u,v}(\boldsymbol{u}, \boldsymbol{v})$ is a smooth-by-smooth interaction trend jointly defined over the row and column directions.

[4] Nested ANOVA (also called hierarchical ANOVA) are ANOVA models were there is a hierarchy in the factors. Usually the random factors are *nested* inside the fixed factors.

[5] See Appendix A.1 for details about the element-wise matrix product.

In total, six components are used to model the surface f. This may seem like a lot but this allows the translation of model 3.4.02 into a standard mixed model. An interesting property of this proposal is that since $\boldsymbol{u}$ and $\boldsymbol{v}$ are row and column position, it allows depicting the spatial trend in a grid finer than the number of rows and columns. Figure 3.7 shows an example of those six components in the context of a barley uniformity performed by Williams (1988). It shows clearly how the additional components, help to capture small variations in the spatial data.

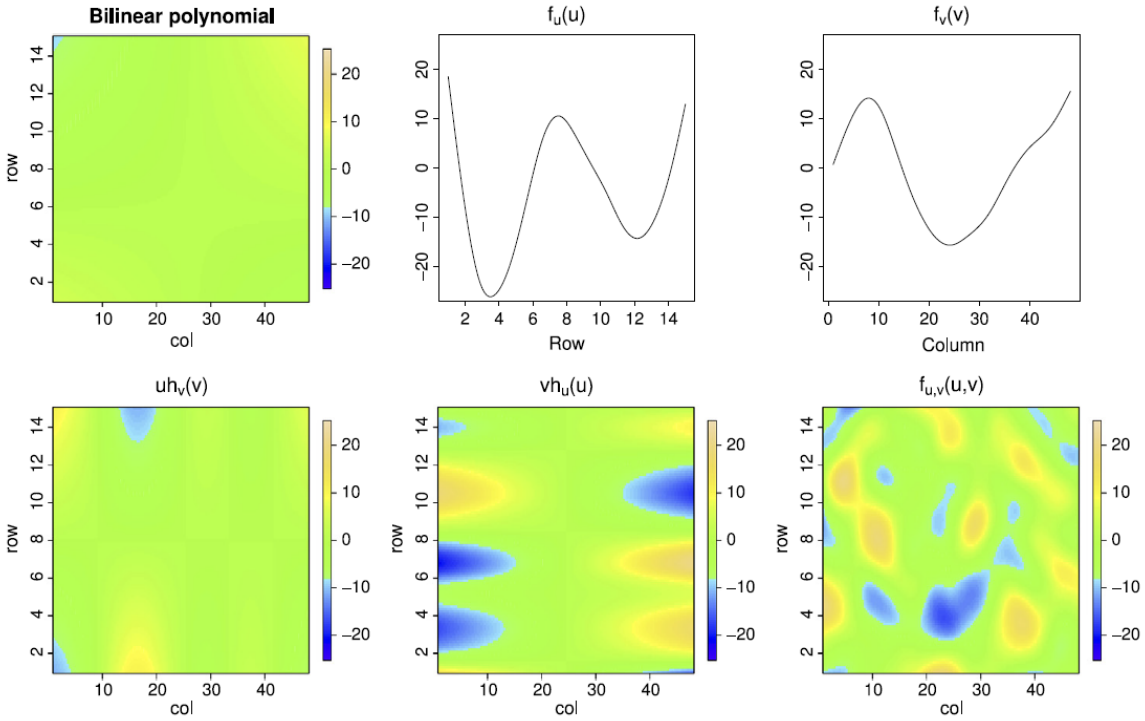


Figure 3.7: Bilinear and smooth components of the PS-ANOVA decomposition of the estimated spatial trend for the barley uniformity data from Rodriguez-Alvarez *et al.* (2018). This trial had 15 rows and 48 columns for a total of 720 data points.

3.4.1 Modelling using P-splines

The functions f_u, f_v, h_u and h_v are constructed with variations on one-dimensional P-splines, while $f_{u,v}$ is based on tensor products P-splines. For clarity's sake, let us consider a model only containing a smooth bivariate surface and an error term

$$y_i = f(u_i, v_i) + \varepsilon_i, \text{ with } \varepsilon_i \sim N\left(0, \sigma^2\right). \tag{3.4.11}$$

Lee, Durbán & Eilers (2013) show that this model can be represented using B-splines[6] and applying a penalty term to form P-splines. Let us first form two B-splines bases:

1. one for the columns, $\hat{\boldsymbol{B}}$ with $b_{il} = \hat{B}_l(u_i)$, where $\hat{B}_l(u_i)$ is the lth B-spline of the basis, evaluated at the ith column, u_i

[6]See Appendix A.3 for details about B-splines and P-splines.

2. and one for the rows, $\breve{\boldsymbol{B}}$ with $b_{ip} = \breve{B}_p(v_i)$, where $\breve{B}_l(v_i)$ is the pth B-spline of the basis, evaluated at the ith row, v_i.

Then, the smooth-by-smooth interaction can be written using those basis

$$f(u_i, v_i) = \sum_{l=1}^{L} \sum_{p=1}^{P} \hat{B}_l(u_i) \breve{B}_p(v_i) \alpha_{lp}, \tag{3.4.12}$$

where $\boldsymbol{\alpha} = (\alpha_{11}, \ldots, \alpha_{1P}, \ldots, \alpha_{LP})^t$ is a vector of unknown regression coefficients of dimension $(LP \times 1)$. Note that $\hat{\boldsymbol{B}}$ and $\breve{\boldsymbol{B}}$ are matrices of order $n \times L$ and $n \times P$ respectively, where L and P are the number of the B-spline basis functions. The values of L and P can be chosen according to the desired definition of the fitted surface f. However, they are often set to the number of columns and rows, respectively, to avoid unecessarly increasing computing time. Dierckx (1995) shows that, in the P-spline framework, the smooth-by-smooth interaction $f(u_i, v_i)$ is modelled by the tensor product of B-splines bases. Then, we can write, in matrix notation,

$$\boldsymbol{B} = \hat{\boldsymbol{B}} \boxdot \breve{\boldsymbol{B}} = \left(\hat{\boldsymbol{B}} \otimes \mathbf{1}_L^t\right) \odot \left(\mathbf{1}_P^t \otimes \breve{\boldsymbol{B}}\right), \tag{3.4.13}$$

where the operation $\boxdot$ is defined in terms of the Kronecker product[7] of two matrices (denoted by $\otimes$) and the element-by-element multiplication of two matrices (denoted by $\odot$). Therefore, model (3.4.11) can be written in matrix notation:

$$\boldsymbol{y} = \boldsymbol{B}\boldsymbol{\alpha} + \boldsymbol{\epsilon}. \tag{3.4.14}$$

The coefficients of this parametric model can be estimated by minimizing the sum of squares. The explicit solution is then:

$$\hat{\boldsymbol{\alpha}} = (\mathbf{B}^t\mathbf{B})^{-1}\mathbf{B}^t\mathbf{y} \tag{3.4.15}$$

To prevent over-fitting, Eilers & Marx (1996) propose to incorporate a discrete penalty on the coefficient associated to adjacent B-splines. As described in details in Appendix A.3, this penalty also determines the smoothness of the splines. The solution of equation (3.4.14) then becomes

$$\widehat{\boldsymbol{\alpha}} = \left(\boldsymbol{B}^t\boldsymbol{B} + \boldsymbol{P}\right)^{-1} \boldsymbol{B}^t\boldsymbol{y}, \tag{3.4.16}$$

where $\boldsymbol{P}$ is the penalty matrix. The details of this solution are presented in appendix A.4, but the important point to remember here, is that the smoothness of the bivariate surface is defined by the penalty matrix, which only depend on two tuning parameters $\hat{\lambda}$ (smoothing along the columns) and $\breve{\lambda}$ (smoothing along the rows).

3.4.2 Mixed model based smoothing parameter selection

As explained in Rodríguez-Álvarez, Boer, van Eeuwijk, *et al.* (2016), $\mathbf{P}$ is rank-deficient, meaning that the rank[8] is smaller than the number of rows and/or columns, and this

[7] See Appendix **??** for details about the Kronecker product.

[8] The rank of a matrix is the maximum number of linearly independent row-vectors.

causes numerical instability when applying mixed model estimation techniques. To obtain a full-rank penalty matrix, the key is to write model (3.4.14) as

$$\mathbf{B}\boldsymbol{\alpha} = \boldsymbol{X}_s\boldsymbol{\beta}_s + \boldsymbol{Z}_s\boldsymbol{c}_s. \tag{3.4.21}$$

This new way of writing the problem leads to another penalty matrix $\widetilde{\boldsymbol{P}}$, which uses the same tuning parameters ($\hat{\lambda}$ and $\breve{\lambda}$) as the old one $\mathbf{P}$, except it has a block diagonal structure, corresponding to the blocks in $\mathbf{Z}_s$.

Using reformulation (3.4.21), it is straightforward to see how our model only containing a smooth bivariate surface and an error term (3.4.11), can be rewritten as a mixed model:

$$\boldsymbol{y} = \boldsymbol{X}_s\boldsymbol{\beta}_s + \boldsymbol{Z}_s\boldsymbol{c}_s + \boldsymbol{\varepsilon}, \text{ with } \boldsymbol{\varepsilon} \sim N\left(\mathbf{0}, \sigma^2\boldsymbol{I}_n\right) \text{ and } \boldsymbol{c}_s \sim N\left(\mathbf{0}, \boldsymbol{G}_s\right), \tag{3.4.22}$$

where $\boldsymbol{G}_s$ also has a block diagonal structure, similar to that of $\mathbf{Z}_s$. This new definition has two advantages: first, it allows the estimation of our spatial mixed model with the tensor product of splines; then it allows the spatial surface ($f(\boldsymbol{u}, \boldsymbol{v})$) to be tuned by several distinct parameters, which provides flexibility to account for global and local variations. Details about the computation and selection of the smoothing parameters and about the structure of the $\boldsymbol{G}_s$ matrix are available in Appendix A.4 along with a diagram detailing the structure of all the matrices used in this section.

3.4.3 Spatial models for field trials

The tensor product of P-splines, presented in the previous section, constitutes the basis for the analysis of agricultural field trials because it allows the modelling of the random spatial variation typically present in a field. On top of this spatial field, we need to build up a more complex models in order to account for the genetic variation, the different tanks, strips and positions. From now on, we therefore consider the following linear mixed model

$$\boldsymbol{y} = \underbrace{\boldsymbol{X}_s\boldsymbol{\beta}_s + \boldsymbol{Z}_s\boldsymbol{c}_s}_{f(\boldsymbol{u},\boldsymbol{v})} + \boldsymbol{X}_d\boldsymbol{\beta}_d + \boldsymbol{Z}_d\boldsymbol{c}_d + \varepsilon, \text{ with } \boldsymbol{c}_s \sim N\left(\mathbf{0}, \boldsymbol{G}_s\right) \text{ and } \boldsymbol{c}_d \sim N\left(\mathbf{0}, \boldsymbol{G}_d\right), \tag{3.4.31}$$

where $\boldsymbol{X}_s$, $\boldsymbol{Z}_s$ and $\boldsymbol{G}_s$ are defined in the previous section and form the mixed model expression of the smooth spatial surface ($f(\boldsymbol{u}, \boldsymbol{v})$), and $\boldsymbol{X}_d$ and $\boldsymbol{Z}_d$ represent column-partitioned matrices, associated respectively with fixed and random components. Since we do not have any check genotypic varieties[9] or resolvable block effect, the only extra fixed effects are: an intercept (1_n) and the tank effect (T). We assume that the $\boldsymbol{X}_d$ matrix has full-rank. The position on the strip (P) and strip (St) variables are added as random effects in $\mathbf{Z}_d$. Therefore $\mathbf{X}_d = [\mathbf{X}_{1_n}, \mathbf{X}_T]$, $\boldsymbol{Z}_d = [\boldsymbol{Z}_P, \boldsymbol{Z}_{St}]$ and $\boldsymbol{G}_d =$ blockdiag $(\boldsymbol{G}_P, \boldsymbol{G}_{St})$.

[9] Check varieties are genotypes with known characteristics that are put in all the blocks/environments of an experiment to allow an easier comparison.

We further assume that $\boldsymbol{c}_s$ and $\boldsymbol{c}_d$ are independent. To keep the notation simple, we rewrite model (3.4.31) as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{Z}\boldsymbol{c} + \boldsymbol{\varepsilon}, \text{ with } \boldsymbol{c} \sim N(\boldsymbol{0}, \boldsymbol{G}) \text{ and } \boldsymbol{\varepsilon} \sim N\left(\boldsymbol{0}, \sigma^2 \boldsymbol{I}_n\right) \tag{3.4.32}$$

where $\boldsymbol{X} = [\boldsymbol{X}_s, \boldsymbol{X}_{1_n}, \mathbf{X}_T]$, $\boldsymbol{Z} = [\boldsymbol{Z}_s, \boldsymbol{Z}_P, \mathbf{Z}_{St}]$, and $\boldsymbol{G} =$ blockdiag $(\boldsymbol{G}_s, \boldsymbol{G}_P, \boldsymbol{G}_{St})$.

3.4.4 Model estimation

With all these specifications in mind, the model was fitted using cubic P-splines with second-order penalties. These settings are commonly used and allow flexibility of the model (Rodríguez-Álvarez, Boer, van Eeuwijk, *et al.* 2016; Rodriguez-Alvarez *et al.* 2018; Rodríguez-Álvarez, Lee, *et al.* 2015). We used 99 and 5 equally spaced knots for the P-splines, corresponding to the number of strips and positions, respectively. In this way there was approximately one knot for every plot. Thus there was a total of 362 model parameters to be estimated for the smooth surface (and 501 for the whole model). As Rodriguez-Alvarez *et al.* (2018) explains, the number of knots is not critical since the optimization of the fit to the data is essentially dependant on the smoothing parameters.

The SpATS model usually takes rows and columns coordinates as inputs for spatial position. To stay consistent with the platform notation, we replaced the rows by the strips and the columns by the positions. Given that we have 2 tanks (A and B), 99 strips and 5 positions, this coordinate replacement would have given us a field with 99 rows and 10 columns. In order to match the platform setup, we reshaped the data to have the tank side by side and the 99 strips divided in two columns, so that we would have a field with 50 rows and 20 columns. Figure 3.8 shows the reshaping of the positions. This new display of the data allows us to see the difference between tanks more clearly and to visualize the plant yields as they were in the greenhouse.

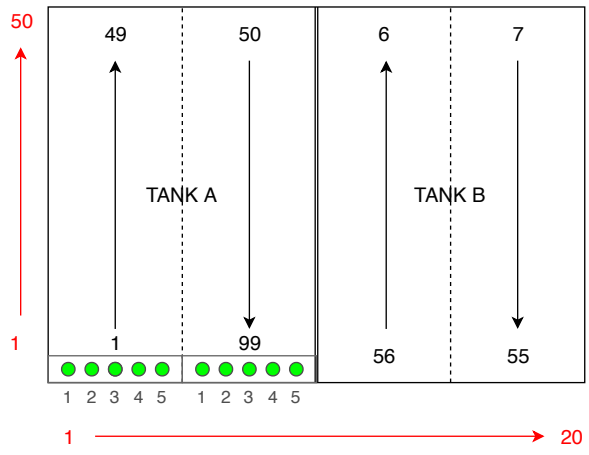


Figure 3.8: Reshaping of the data table to fit the disposition of the greenhouse. The black number indicates the original strip number, the grey numbers show the positions on the strip and the red numbers indicates the new numbers used in the spatial models.

The estimation procedure was performed using the R-package `SpATS` (Rodríguez-Álvarez, Boer, Eilers, *et al.* 2016). This package provides a REML-based estimation of the variances components and computes the best linear unbiased estimators (BLUEs) of the fixed effects and the empirical best linear unbiased predictors (BLUPs) of the random effects. A useful by-product of this computation is the effective dimension associated to each random effect.

Effective dimension

In the P-splines methodology, the effective dimension (ED) measures the complexity of the model components (Eilers, Marx & Durbán 2015). It is similar to the more common concept of effective degree of freedom (Buja *et al.* 1989). It is computed as the trace of the hat matrix $\mathbf{H}$[10] and is bounded between zero and a value depending on the number of knots (Rodríguez-Álvarez, Boer, van Eeuwijk, *et al.* 2016).

The total effective dimension ED_s expresses the number of parameters effectively involved in modelling the spatial surface. Consequently, it can be interpreted as a measure of the magnitude of field variations with larger values indicating more intense spatial patterns. Using the PS-ANOVA structure of the spatial model, the total effective dimension can be decomposed for each components of the smooth surface $f(\boldsymbol{u}, \boldsymbol{v})$. These partial effective dimensions are indicative of the relative contribution of each component to the fitted surface.

Generalized heritability

As we just said, the effective dimension is a useful tool to measure the relative importance of each spatial component, and to compare them. However, the effective dimension of the genetic component is harder to compare to the rest. This is where the concept of heritability comes in handy. In classical genetic models, the standard heritability is defined as the proportion of the total (phenotypic) variation that is attributable to the genetic component. Rodriguez-Alvarez *et al.* (2018) show the link between this definition and the genetic effective dimension and establish the following relationship, for experiments where there are no marker[11] information available about the the genotypes:

$$H_g^2 = \frac{ED_g}{n_g} = 1 - \frac{\overline{PEV}}{\sigma_g^2}. \tag{3.4.41}$$

In this equation, n_g is the number of genotypes and $\overline{PEV}$ is the prediction error variance for the genotype BLUPs. Given that our experiment does not incorporate a genetic relationship matrix, we'll use this definition to properly interpret the genetic effect in our model. We can take advantage of the right hand term of the equation to compute the heritability for the standard spatial model that does not have effetive di-

[10] The hat matrix $\mathbf{H}$ (also called the projection, or influence matrix) maps the vector of response values (denoted $\mathbf{y}$) to the vector of fitted values (denoted $\hat{\mathbf{y}}$), so that $\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}$

[11] (Molecular) Markers are fragments of DNA associated with a certain location in the genome. They are widely used in genotyping to identify useful traits and relationship between genotypes.

mensions Welham *et al.* (2010). Details about the effective dimensions are available in Rodriguez-Alvarez *et al.* (2018) and references therein.

3.5 Standard spatial models

In this section the $AR(1) \times AR(1)$ model, and its extension to the linear variance (LV) model, are presented. For more detailed information about the original $AR(1) \times AR(1)$ model, consult Gilmour *et al.* (1997). For information about the extensions of the model, see Piepho & Williams (2010) and Williams (1986).

As explained in the literature review chapter, the standard spatial models (SS models), such as the $AR(1) \times AR(1)$ model, use a spatially dependent error term to adjust for the local trends, whereas the SpATS model encompasses them in the smooth surface. To estimate the spatial covariance structure of this error process, Gilmour *et al.* (1997) advocate the use of variograms and more precisely, the use of first order autoregressive processes ($AR(1)$) to model those variograms. The general representation of an SS model is very similar to the SpATS model:

$$y = \mathbf{X}\boldsymbol{\beta} + \mathbf{X}_s\boldsymbol{\beta}_s + \mathbf{Z}\mathbf{c} + \mathbf{Z}_s\mathbf{c}_s + \boldsymbol{\xi} + \boldsymbol{\varepsilon}, \tag{3.5.01}$$

where $\mathbf{X}\boldsymbol{\beta}$ contains the fixed terms of the model, $\mathbf{Z}\mathbf{c}$ contains the random terms, $\boldsymbol{\varepsilon}$ is the vector of the residuals and $\boldsymbol{\xi}$ is a vector of spatially dependent residuals modelling the local trends. However, unlike in the SpATS model, $\mathbf{X}_s\boldsymbol{\beta}_s$ and $\mathbf{Z}_s\mathbf{c}_s$ do not model a smooth bivariate surface, but contain linear and one-dimensional cubic splines terms that account for the global variations in the data (see Verbyla *et al.* (1999) for details).

3.5.1 Variogram and $AR(1)$ process

As Gilmour *et al.* (1997) explain in their paper, they model an error process using a variogram. The error process is defined as the sum of the independent and spatially-dependent residuals: $\mathbf{e} = \boldsymbol{\xi} + \boldsymbol{\varepsilon}$, where the term of interest is the spatially dependant error term $\boldsymbol{\xi}$.

The variogram is a function that computes the evolution of the covariance of a spatially-dependent variable over the distance. It enables the visualization of the covariance structure of a variable as a function of spatial displacement. Examples of one- and two-dimensional variograms are presented in figure 3.9. In our case, let us consider the error process $\mathbf{e}$, being a stationary (direction-dependent) process over the rows $\boldsymbol{u}$ and the columns $\boldsymbol{v}$. For two distinct points in space $\boldsymbol{x}$ and $\boldsymbol{y}$, the covariance structure between these two points is modelled by the theoretical variogram (also called the semi-variogram because of the $1/2$ factor in its equation) as:

$$\omega(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{2}\operatorname{var}[\mathbf{e}(\boldsymbol{x}) - \mathbf{e}(\boldsymbol{y})] = \frac{1}{2}[\mathrm{V}(\boldsymbol{x}, \boldsymbol{x}) + \mathrm{V}(\boldsymbol{y}, \boldsymbol{y}) - 2\mathrm{V}(\boldsymbol{x}, \boldsymbol{y})], \tag{3.5.11}$$

where V is the covariance function of our error process $\mathbf{e}$. If we define the distance between the pair of points $(\mathbf{x}, \mathbf{y})$ as:

$$\mathbf{l} = \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} = \begin{bmatrix} |x_1 - y_1| \\ |x_2 - y_2| \end{bmatrix},$$

then Zimmerman & Harville (1991) show that the covariance function of the error process can be rewritten a directional exponential covariance (DEC) structure:

$$\mathrm{V}(\boldsymbol{x}, \boldsymbol{y}) = \exp(-\alpha_1 l_1 - \alpha_2 l_2). \tag{3.5.12}$$

For most field experiments, plots are arranged in regular grids and thus separated by equivalent distances. Therefore the distance between plots is expressed using the displacement vector, which takes values for l_1 of $0, d_1, 2d_1, \ldots, (r-1)d_1$ and for l_2 of $0, d_2, 2d_2, \ldots, (c-1)d_2$, where d_1 and d_2 are the plot dimensions. From that we can create an indexed displacement vector $\mathbf{l}^*$ with values for l_1^* of $0, 1, 2, \ldots, (r-1)$ and values for l_2^* of $0, 1, 2, \ldots, (c-1)$. The semi-variogram can then be rewritten as:

$$\begin{aligned} \omega(\mathbf{l}^*) &= \sigma_\varepsilon^2 + \sigma_\xi^2 \left[1 - \exp\left(-\alpha_1 d_1 l_1^* - \alpha_2 d_2 l_2^*\right)\right] && \mathbf{l}^* \neq 0 \\ &= 0 && \mathbf{l}^* = 0 \end{aligned} \tag{3.5.13}$$

A first order auto-regressive process, is a process in which the term at index i only depends on the value of the process at index $(i-1)$ and on a stochastic term (white noise): $X_i = c + \rho X_{i-1} + \varepsilon_i$. The only parameter in the model (if the intercept is ignored) is ρ. Cressie (1992) and Cullis & Gleeson (1991) show that by setting $\rho_1 = \exp(-\alpha_1 d_1)$ and $\rho_2 = \exp(-\alpha_2 d_2)$, the variogram (equation 3.5.13) becomes

$$\begin{aligned} \omega(\mathbf{l}^*) &= \sigma_\varepsilon^2 + \sigma_\xi^2 \left(1 - \rho_1^{l_1^*} \rho_2^{l_2^*}\right) && \mathbf{l}^* \neq 0 \\ &= 0 && \mathbf{l}^* = 0 \end{aligned}. \tag{3.5.14}$$

This formulation demonstrates the equivalence between the semi-variogram, the DEC model and the $AR(1) \times AR(1)$ process for field experiments. Figure 3.9 shows an example of a simulated variogram using an $AR(1) \times AR(1)$ process and actual variogram fitted using the same process.

3.5.2 Linear variance structure

The modification of the $AR(1) \times AR(1)$ model to include the linear variance (LV) structure, is based on the one-dimensional LV structure proposed by Williams (1986) and its elevation to two dimensions by Piepho & Williams (2010). Let us first consider the case of LV in one dimension (a single position over several strips in our case) for n=99 plots. The baseline covariance structure is written as:

$$\boldsymbol{V}_0 = \sigma^2 \boldsymbol{I}_n + \eta \boldsymbol{J}_n, \tag{3.5.21}$$

where $\boldsymbol{I}_n$ is the n-dimensional identity matrix and $\boldsymbol{J}_n$ is a n by n matrix of ones everywhere, σ^2 is the residual variance and η is the block variance. Williams (1986) propose to augment this baseline covariance model by a spatial component $\boldsymbol{V}_s$ such that plots

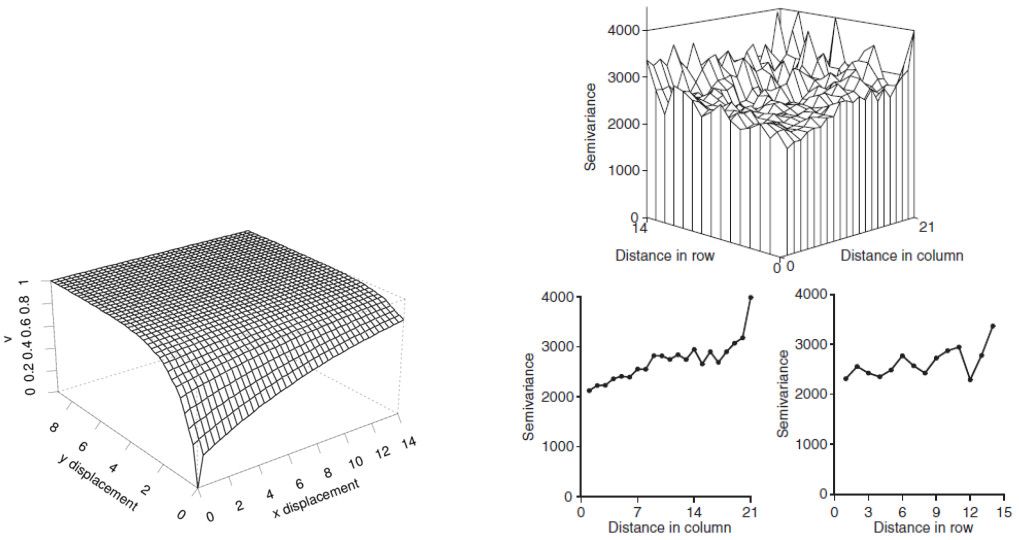


Figure 3.9: Examples of variogram using $AR(1)\times AR(1)$ processes.(Leftmost panel) Variogram for a standardized $AR(1) \times AR(1)$ process, with $\rho_x = 0.9$, $\rho_y = 0.4$ and $\sigma^2 = 0.3$ from Gilmour *et al.* (1997). (Rightmost panel) Variogram fitted to wheat data of Gilmour *et al.* (1997) with the one-dimensional associated variograms in the rows and columns direction (Piepho & Williams 2010).

within the same block obey a correlation structure that decays linearly with distance. For two plots x and y, they define the matrix $\boldsymbol{L}_n$ as a $n \times n$ matrix where the (x, y)-th element is equal to $|x - y|$. From that statement, the decaying correlation structure can be expressed in matrix form as $\boldsymbol{V}_S = \kappa \left(\boldsymbol{J}_n - \tilde{\phi}\boldsymbol{L}_n\right)$, where $\tilde{\phi}$ is a slope parameter determining the decaying process and κ is a spatial variance. To avoid confounding effects, Williams (1986) rewrite the complete, augmented model as a linear structure:

$$\boldsymbol{V} = \boldsymbol{V}_0 + \boldsymbol{V}_S = \sigma^2\boldsymbol{I}_n + \eta\boldsymbol{J}_n + \phi\boldsymbol{M}_n, \tag{3.5.22}$$

where $\boldsymbol{M}_n = (n-1)\boldsymbol{J}_n - \boldsymbol{L}_n$ and $\phi = \kappa\tilde{\phi}$. The components of the model have the following interpretation:

- $\sigma^2\boldsymbol{I}_n$ is the residual error (or nugget effect in variogram-specific terms).
- $\eta\boldsymbol{J}_n$ is the block effect that captures extraneous variations on the field.
- $\phi\boldsymbol{M}_n$ models a spatial correlation within a block by a LV–covariance structure where correlation decays linearly with distance.

Piepho & Williams (2010) present two main ways to extend the model to two dimensions. The first one is a superimposition of the two structures: a LV structure along the columns but independent between the rows ($\boldsymbol{V} \otimes \boldsymbol{I}$) and a LV structure along the columns but independent between the rows ($\boldsymbol{IV}$). The second one is a product of the two LV structures: $\boldsymbol{V} \otimes \boldsymbol{V}$. In a desire of clarity, the full notation of these structure is not presented here, but are detailed in Piepho & Williams (2010).

The similarity between the LV structure and the $AR(1)$ process comes from the fact that when the spatial correlation $\rho^l \approx 1$ then $\rho^{|x-y|} = \exp\left[\log(\rho)\,|x-y|\right] \approx 1 - \tilde{\phi}\,|x-y|$

with $\tilde{\phi} = -\log(\rho)$. Thus, when the correlation is close to one (which is quite common in practice (Pilarczyk 2007)) distances $|x - y|$ are not large, $AR(1)$ and LV models are expected to yield similar results. The main advantages of the LV model is that it is more robust to convergence issues when correlation is close to unity and also that it has one parameter instead of two, for the spatial component. To visualize the different spatial covariance structures of the different models, Figure 3.10 shows an example of each of those structures.

3.5.3 Best standard spatial model

As explained in the previous section, several spatial covariance structures can be used in the standard spatial (SS) model to account for the spatial variations. Comparing the fit to the data of all the different models, would be a tedious process. Instead, all the SS models will be compared in term of AIC. This goodness-of-fit statistic is commonly used to compare statistical models (Gilmour *et al.* 1997; Piepho & Williams 2010; Velazco *et al.* 2017). It is based on the value of the likelihood function, $\mathcal{L}$:

$$AIC = 2k - 2ln(\mathcal{L}). \tag{3.5.31}$$

The main advantage of the AIC is that it takes into account the number of parameters of the model k, to avoid overfitting. The best standard spatial model (BSS) was retained as the one with the lowest AIC. This model was used with the data from the platform and its results were compared to the one from the SpATS model.

3.5.4 Model estimation

The models were implemented using the `proc mixed` of the SAS software (SAS Institute, Cary, NC, USA). The spatial covariance structure was specified in the `type` option of the `random` statement of the procedure. The $AR(1)$ models were fitted using `type = AR(1)`, the $AR(1) \times AR(1)$ were fitted using `type = sp(POWA)` and the LV models were fitted using `type = lin(q)`.

$$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{pmatrix}$$

$\boxed{AR(1)}$

$$\sigma^2 \begin{pmatrix} 1 & \rho_r^{d(1,2,r)}\rho_c^{d(1,2,c)} & \rho_r^{d(1,3,r)}\rho_c^{d(1,3,c)} & \rho_r^{d(1,4,r)}\rho_c^{d(1,4,c)} \\ \rho_r^{d(2,1,r)}\rho_c^{d(2,1,c)} & 1 & \rho_r^{d(2,3,r)}\rho_c^{d(2,3,c)} & \rho_r^{d(2,4,r)}\rho_c^{d(2,4,c)} \\ \rho_r^{d(3,1,r)}\rho_c^{d(3,1,c)} & \rho_r^{d(3,2,r)}\rho_c^{d(3,2,c)} & 1 & \rho_r^{d(3,4,r)}\rho_c^{d(3,4,c)} \\ \rho_r^{d(4,1,r)}\rho_c^{d(4,1,c)} & \rho_r^{d(4,2,r)}\rho_c^{d(4,2,c)} & \rho_r^{d(4,3,r)}\rho_c^{d(4,3,c)} & 1 \end{pmatrix}$$

$\boxed{AR(1) \times AR(1)}$

$$\sigma^2 \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \eta \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} +$$

$$(n-1)\phi \begin{pmatrix} 1 & 1-d(1,2) & 1-d(1,3) & 1-d(1,4) \\ 1-d(2,1) & 1 & 1-d(2,3) & 1-d(2,4) \\ 1-d(3,1) & 1-d(3,2) & 1 & 1-d(3,4) \\ 1-d(4,1) & 1-d(4,2) & 1-d(4,3) & 1 \end{pmatrix}$$

$\boxed{\text{LV}}$

Figure 3.10: Spatial covariance structures of the standard spatial models. The $AR(1) \times AR(1)$ model is presented as anisotropic (the auto-correlation parameter is directionally dependent). In isotropic conditions (auto-correlation is similar in every direction), the correlation value would be the same along the rows and the columns ($\rho_c = \rho_r$). The distance is represented in units, as the indexed displacement. here $d(i, ij)$ represents the indexed displacement between plot i and plot j.

3.6 Model comparison

The SpATS model was compared with the BSS model in terms of quality of fit to the data and ability to detect genotype differences, by using plots of the fitted values and residuals. Similar parameters to those used by Velazco *et al.* (2017), were considered also for the comparison:

- Genetic variance (σ_g^2): precision of the genotype effect estimation. The generalized heritability will also be used to asses the same parameter.
- BLUPs of the genotypes: similarity between the estimates of the effect of each genotype for the two models.
- Estimates of the tank effect: similarity between the estimates of the fixed tank effect for the two models.

It is interesting to note that Rodriguez-Alvarez *et al.* (2018) also use the Pearson correlations of predicted genotypes values between environments (i.e. field trials) as a way of comparing models. Since only one trial was studied in this thesis, this correlation cannot be used.

Chapter 4

Results and discussion

4.1 Descriptive statistics

Using the individual weights attributed to each plant, we compute the weighted mean and weighted standard deviation for all variables. Even though we have four variables in our analysis, it is important to notice that there is a high correlation between the variables (see Table 4.1) and that in a future trial, the analysis of other less-correlated variables (non-measured here) could be interesting.

Table 4.1: Correlation matrix of the four variables.

	DRY_LS	DRY_RS	FRESH_RS	FRESH_LS
DRY_LS	1.00	0.70	0.84	0.93
DRY_RS	0.70	1.00	0.76	0.68
FRESH_RS	0.84	0.76	1.00	0.93
FRESH_LS	0.93	0.68	0.93	1.00

Because of germination problems on the platform and inside the germination chamber, not all genotypes were similarly represented in the experiment. Table 4.2 presents the effective germination rates for each genotype, i.e. the number of seed actually kept for the spatial analysis over the number of seeds placed on the platform. This table is interesting because germination rate is a genotypic feature. We see clear discrepancies between genotypes, some have high germination rates (e.g. genotypes 29 and 22 both havee germination rate higher than 90%) while 6 genotypes have a germination rate lower than 50% (below the dashed line on Table 4.2), even though more than 15 seeds per genotype were placed on the platform. This indicates that all the genotypes, used in this experiment, may not be well suited to aeroponic growth.

Using the weighted means and standard deviations, we created boxplots ordered by descending mean value for tank A, presented in Figure 4.1. The numerical values of these results are presented in Table B.1, in Appendix B.1. It is clear that values for tank A are almost always higher than for tank B expect for some genotypes (e.g. genotype 12 on all variables and genotype 11 on the dry weights), even if there seems to be more

Table 4.2: Effective on-platform germination rates (GR) with the number of seeds kept for data analysis (NS kept) and the number of seeds actually placed on the platform (NS placed) for each genotype. The dotted line represents the 50% germination rate limit.

Genotype	NS placed	NS kept	GR
25	29	27	93.1
23	22	20	90.9
16	27	23	85.2
18	26	22	84.6
3	29	24	82.8
17	27	21	77.8
19	30	23	76.7
1	24	18	75.0
12	28	21	75.0
14	26	19	73.1
28	26	19	73.1
9	29	20	69.0
6	29	19	65.5
7	29	19	65.5
4	23	15	65.2
⋮	⋮	⋮	⋮

Genotype	NS placed	NS kept	GR
⋮	⋮	⋮	⋮
27	29	18	62.1
21	26	16	61.5
5	30	18	60.0
24	30	18	60.0
10	29	17	58.6
20	26	15	57.7
22	18	10	55.6
29	28	15	53.6
13	27	14	51.9
15	18	9	50.0
2	26	12	46.2
11	19	8	42.1
26	29	12	41.4
8	21	8	38.1
30	23	3	13.0

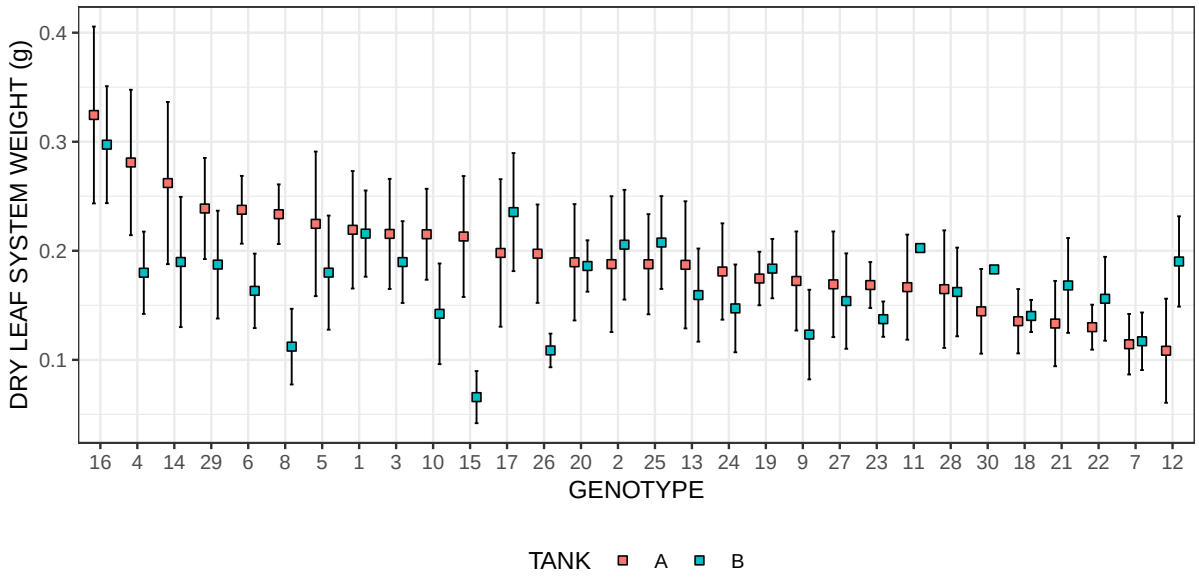
variation in tank A. A t-test proved this difference to be highly significant[1] (see Table 4.3). However, the value for tank B is higher for some genotypes (e.g. genotype 12 for all variables and genotype 11 for the dry weights). This shows that those genotypes might be more suited to a still growing environment.

Table 4.3: Mean weight value (g) for each tank with associated p-value from a t-test for the difference of means between the two tanks

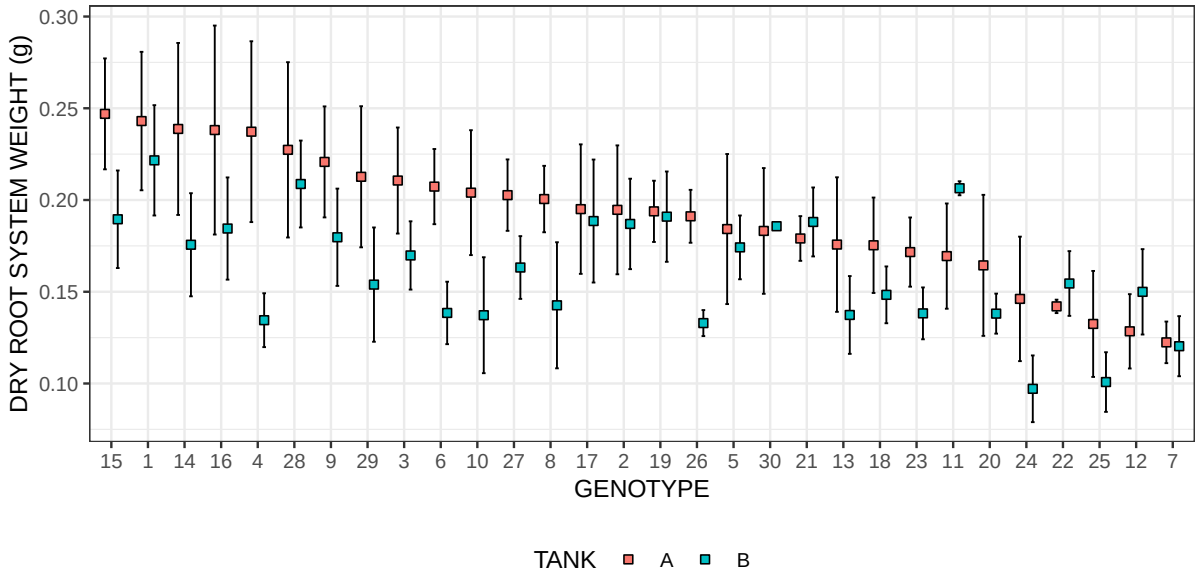
	FRESH_LS	FRESH_RS	DRY_LS	DRY_RS
Mean value in tank A	2.604	3.09	0.1963	0.1964
Mean value in tank B	1.755	1.79	0.1725	0.1666
p-value	$<$0.001	$<$0.001	0.003	$<$0.001

To further assess difference in means between the two tanks, we performed a paired t-test of the means difference between the two tanks, for all variables. The resulting p-values are presented in table B.2 in Appendix B.2. We see that the difference in mainly significant for the fresh weights (which is also visible on Figure 4.1 **a** and **b**) and only for a handful of genotypes (6, 23, 10, 26, 4 and 8). We chose to use multiple univariate t-tests, because multivariate t-tests are extremely sensitive. Any violation to the multivariate normality assumption and equality of variance-covariance matrices will lead a significant result.

[1] The test were performed with a significance level, α = 0.05.

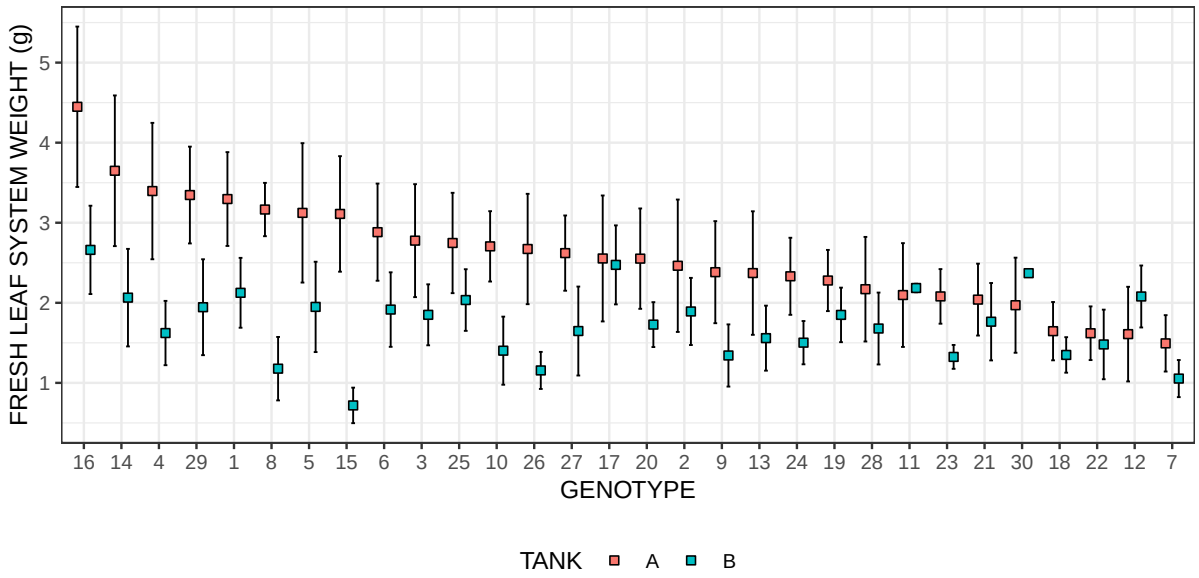


(a) Dry leaf weight (DRY_LS)

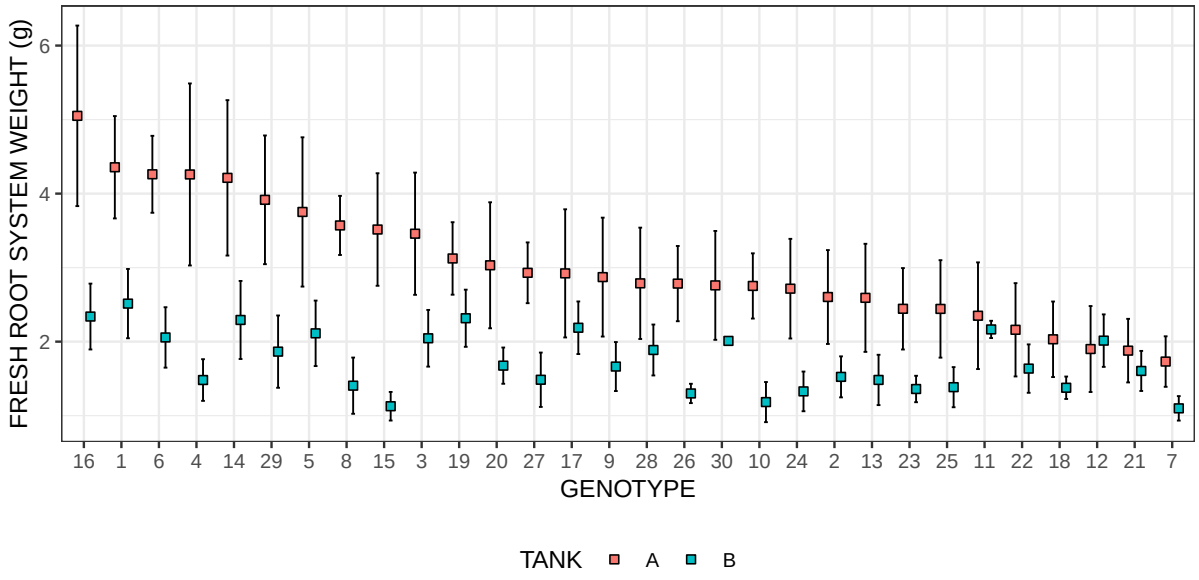


(b) Dry root weight (DRY_RS)

Figure 4.1: Boxplot displaying mean weight (□) and associated standard deviation (—), grouped by tanks and ordered by descending mean value for tank A.



(c) Fresh leaf weight ($FRESH\_LS$)



(d) Fresh root weight ($FRESH\_RS$)

Figure 4.1: Boxplot displaying mean weight (□) and associated standard deviation (—), grouped by tanks and order by descending mean value for tank A.

4.2 SpATS analysis

For the SpATS model, we analysed the four weight variables using the following model:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_T + \mathbf{X}_s\boldsymbol{\beta}_s + \mathbf{Z}_s\boldsymbol{s} + \mathbf{Z}_u\boldsymbol{u} + \mathbf{Z}_v\boldsymbol{v} + \mathbf{Z}_g\boldsymbol{g} + \boldsymbol{\varepsilon}, \tag{4.2.01}$$

where:

- $\mathbf{X}\beta_T$ is the fixed term for the tanks,
- $\mathbf{X}_s\boldsymbol{\beta}_s + \mathbf{Z}_s\boldsymbol{s}$ is the mixed model representation of the bivariate smooth surface $f(\boldsymbol{u}, \boldsymbol{v})$,
- $\mathbf{Z}_u\boldsymbol{u}$ is the random effect of the rows,
- $\mathbf{Z}_v\boldsymbol{v}$ is the random effect for the columns,
- $\mathbf{Z}_g\boldsymbol{g}$ is the random effect for the genotypes, and
- $\boldsymbol{\varepsilon}$ is the residual error.

The analysis of the results is split in 2 parts:

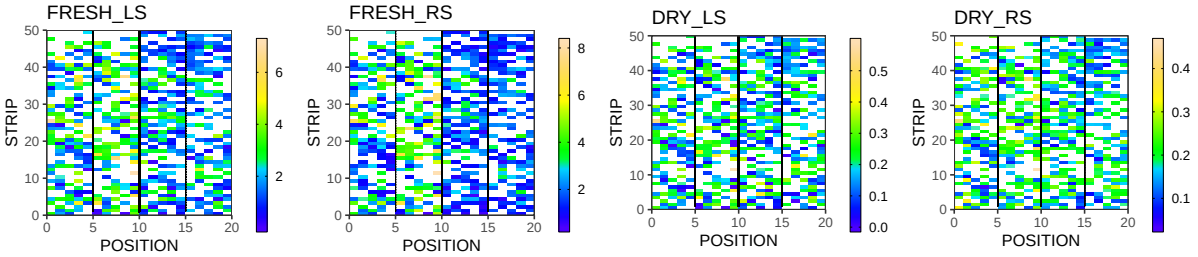
1. A visual analysis of the fitted values and the residuals compared to the raw data; to see if the spatial patterns have been accounted for by the spatial model.
2. An analysis of the individual contribution of each term of the bivariate smooth surface $f(\boldsymbol{u}, \boldsymbol{v})$, using the effective dimensions.

The analysis of the estimated coefficient for the genotypes and for the tanks is presented in a following section, with the estimates of the BSS model.

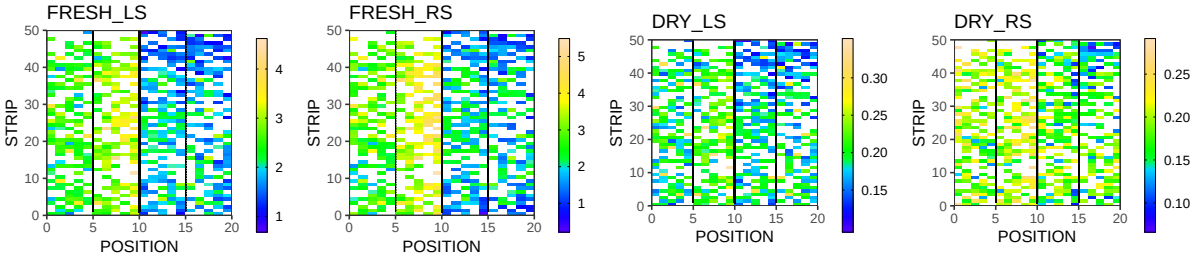
4.2.1 Visual analysis

The raw data, the fitted data and the spatial model residuals are presented, for all variables, in Figure 4.2. The difference between tanks is more pronounced for the fresh weights than for the dry weights, just as stated in the previous section. However, this difference is well represented in the fitted data, where tank B (on the right in each picture) consistently display lower weight values. A strip effect is also noticeable for tank B, where the last strips have lower values than the first ones. This might be an effect of the still growing environment, since this effect is not present in tank A, which was constantly moving.

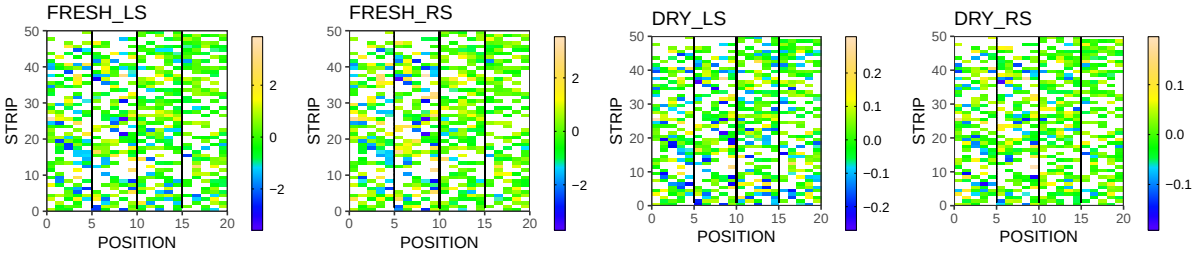
No clear pattern emerges from the residuals, expect the fact that tank A seems to have more single data point with lower values that were not picked up by the spatial surface. This means that tank A had more local heterogeneities. Indeed, when we consider the movement in tank A, it makes sense not to see clear spatial trends revealed in the data, and rather having a wider variations among the residuals. A analysis of the residuals distribution revealed no violation to the normality assumption.



(a) Raw data



(b) Fitted spatial trend



(c) Residuals' spatial plot

Figure 4.2: Raw data, fitted spatial trend and residuals' plot for each variable.

Furthermore, a comparison of the scales of the fitted values and the raw data on Figure 4.2 shows us that the smooth surface could not account for the extreme values present in the raw data. For example the scale of the weight values for FRESH_RS goes from 0 to 8, whereas it only goes from 0 to to 5 for the fitted surface. However, the scale of the residuals is lower than the one of the fitted values, indicating that the smooth surface accounts for a good amount of the spatial variation.

4.2.2 Smooth surface terms analysis

The smooth part of the smooth bivariate surface can be decomposed in the following terms:

$$f(\boldsymbol{u}, \boldsymbol{v}) = f_u(\boldsymbol{u}) + f_v(\boldsymbol{v}) + \boldsymbol{u} \odot h_v(\boldsymbol{v}) + \boldsymbol{v} \odot h_u(\boldsymbol{u}) + f_{u,v}(\boldsymbol{u}, \boldsymbol{v}), \tag{4.2.21}$$

that are linear terms, linear-by smooth and smooth-by-smooth interaction terms (see section 3.4 for more details). The contribution of each term to the whole surface can

be assessed using the effective dimension of each component. Table 4.4 presents the contribution of each term for each variable as the percentage of the total effective dimension of the smooth surface.

For the fresh weights, the smooth-by-smooth interaction ($f_{u,v}(\boldsymbol{u}, \boldsymbol{v})$) represents almost 70% of the total surface variation. This emphasizes the complexity of the spatial patterns. The smooth-by-linear along the strips interaction ($\boldsymbol{u} \odot h_v(\boldsymbol{v})$) and the smooth along the strips ($f_u(\mathbf{u})$) terms account for the rest of the variation. This confirms that there is more variation along the strips than along the positions.

For the dry weights, the variation is shared among all the terms except for smooth along the position that accounts for nothing. This means that there were less heterogeneous spatial variation and more smooth gradients than for the fresh weights. Indeed, if we look back at the fitted residuals on Figure 4.2, we see clearly that there is a gradient among the strips and the positions, instead of random patches of high yield.

The total effective dimension is also displayed in the table, because it is an indicator of the overall complexity of the spatial patterns that the smooth surface models. Since the total is lower for the dry weights, it confirms again that the spatial variability was less complex for those variables. As explained in the previous chapter, the effective dimension of a parameter, is proportional to its variance. Therefore, a table presenting the variance of all the random components of the models is presented in Appendix B.3.1

Table 4.4: Contribution of each component of the spatial smooth surface as a percentage of the total effective dimension of the surface. Here $\boldsymbol{v}$ represents the columns, i.e. the position on the strip; and $\boldsymbol{u}$ represents the rows, i.e. the strip itself.

Model components	FRESH_LS	FRESH_RS	DRY_LS	DRY_RS
$f_v(\mathbf{v})$	0[†]	0	0	0
$f_u(\mathbf{u})$	8.73	14,10	17,26	17,26
$\boldsymbol{u} \odot h_v(\boldsymbol{v})$	19,95	16,40	24,77	24,77
$\boldsymbol{v} \odot h_u(\boldsymbol{u})$	2,56	0	18,36	18,36
$f_{u,v}(\boldsymbol{u}, \boldsymbol{v})$	68,76	69,5	39,62	39,62
Total	9,21 (100%)	16,15 (100%)	5,93 (100%)	4,91 (100%)

[†] All values inferior to 0.0001% were marked as 0 in the table.

4.3 Standard spatial model analysis

For the standard spatial model analysis,we fitted a baseline model, only considering a fixed effect for the tanks and a random effect for the genotypes and spatially independent residuals:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta}_T + \mathbf{Z}_g\mathbf{g} + \mathbf{e}. \tag{4.3.01}$$

The model was then augmented separately for each variable, by adding linear regression terms on the rows and columns and one of the following covariance structure:

- $AR(1)$ process along the rows or columns
- $AR(1) \times AR(1)$ process
- LV process along the rows or columns
- Superimposed row and column structure $LV + LV$
- Separable process along the rows or columns $LV \otimes J$
- Separable process along the rows and columns $LV \times LV$

At each step, the AIC was computed and used to select the best model (a lower value is preferred). Table 4.5 gives the structure of the final selected model for each variable.

Table 4.5: Best standard spatial (BSS) model selected for each of the four variables. All the models contain an intercept and a fixed effect for the tank and a random effect for the genotypes. P represents a random effect for the positions (columns), S a random effect for the strips (rows) and n represent the spatially independent residuals.

Variable	BSS
FRESH_LS	$S + AR(1) \times AR(1)$
FRESH_RS	$S + AR(1) \times AR(1)$
DRY_LS	$S + P + LV \times LV$
DRY_RS	$S + P + LV + LV$

First, we see that, for the fresh weights, the models only contains a random effect on the strips (rows) and not the positions (columns). This is similar to the interpretation of the effective dimensions of the SpATS model from Table 4.4, that highlighted the strong strip effect and the almost non-existing position effect for the fresh weights. However, all models used a strips-by-positions covariance structure, illustrating the fact that the spatial trends display complex patterns that cannot be accounted for by a one dimensional process only.

Then, we see that the fresh weights were best represented by an auto-regressive process whereas the dry weights needed a linear covariance structure. Piepho & Williams (2010) explain that both structures give similar results with large auto-correlation (ρ) values, which is the case here (see table 4.6) but that the LV structure is more robust to convergence issues. Since we encountered some convergence problems when fitting the $AR(1)$ structure on the dry weight models, it might explain the better AIC value of the LV models.

Finally, both dry weight models use a linear covariance structure but DRY_LS uses the superimposed structures whereas DRY_LS uses the separable model. The only difference between those models is the way they model the pairwise variances for plots

on different strips (rows) and different positions (columns), but it is not relevant here, as they yield similar results in most cases (Piepho & Williams 2010).

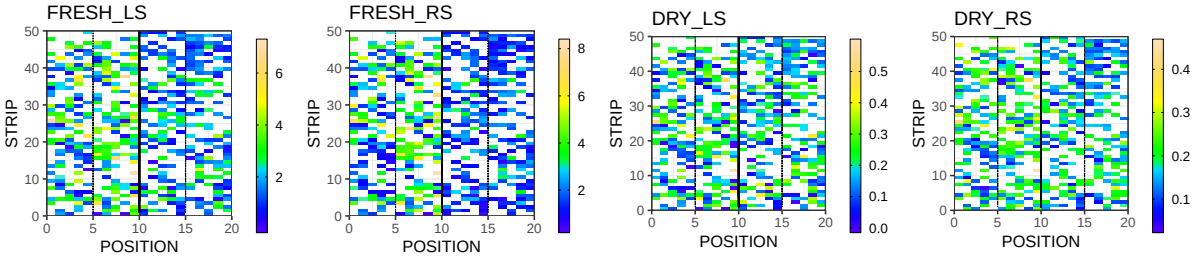
Table 4.6 presents the auto-correlation values for the two fresh weight models (since the dry weights models do not use an auto-regressive covariance structure). We see that both models exhibit high auto-correlation, however, it is slightly less pronounced over the positions than over the strips. These high values mean that there is a high spatial variation for both variables, which is similar to the conclusions from the SpATS model. However, Piepho, Möhring, *et al.* (2015) suggest that auto-correlation values close to one indicates confounding between the trends and the rows/strips. This might explain the absence of a random term for position in the BSS models for the fresh weights (Table 4.5).

Table 4.6: Auto-correlation values along the strips (ρ_s) and the positions (ρ_p), for the two BSS models using an $AR(1) \times AR(1)$ spatial covariance structure.

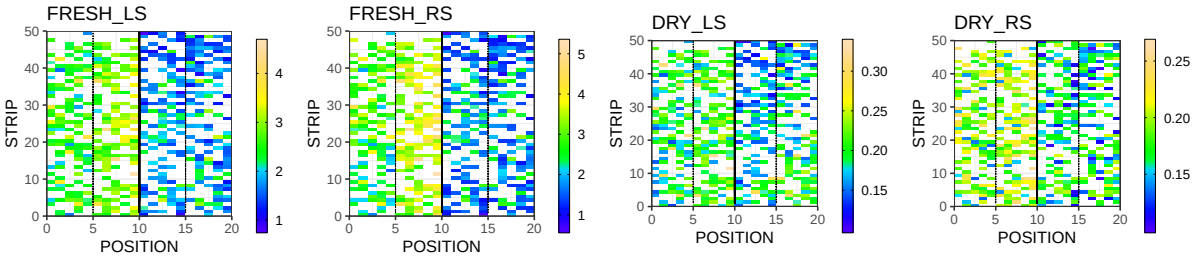
Variable	ρ_s	ρ_p
FRESH_LS	0,998	0,992
FRESH_RS	0,993	0,952

4.3.1 Visual analysis

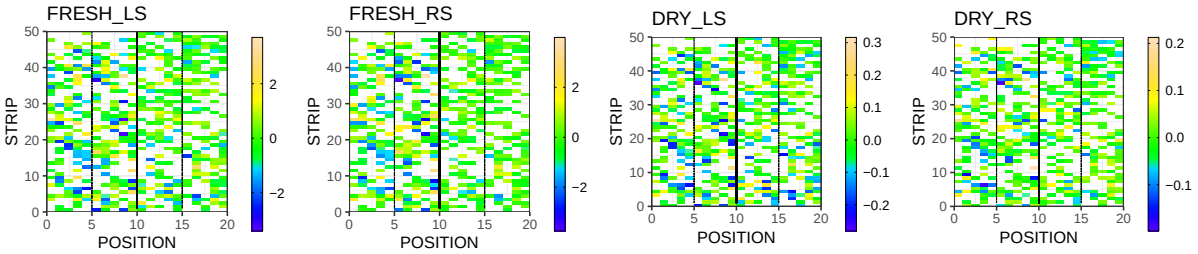
Figure 4.3 presents the raw data, the fitted values and the residuals of each variables, in a similar fashion than for the SpATS model. We see that the fitted data are very similar to those of the SpATS model and capture the spatial trends correctly. Furthermore, the range of the residuals is similar to the SpATS model, meaning that there are no visual clues indicating that the BSS model is better or worse than the SpATS model.



(a) Raw data



(b) Fitted spatial trend



(c) Residuals' spatial plot

Figure 4.3: Raw data, fitted spatial trend and residuals' plot for each variable.

4.4 Model comparison

4.4.1 Estimation performances

In this section we compare the two models on three different points:

- The estimates of the fixed tank effect
- The genetic variance and the heritability
- The estimates of the genotypic random effects

Table 4.7 presents the estimates of the fixed tank effects for each model. The estimates from the BSS model are consistently lower than the ones from the SpATS model, but univariate t-tests for the difference in means using the mean value and the standard deviation, showed no significant differences[2] between the estimates of the two models. Furthermore, both values are close to the weighted means by tank, obtained from the raw results (see Table 4.3). Finally, the tank effect was significant for all variables and in each model, confirming the clear tank effect in this platform.

Table 4.7: Comparison of the estimated fixed tank effect for both models.

		FRESH_LS	FRESH_RS	DRY_LS	DRY_RS
Tank A	SpATS	2,9527	3,7282	0,2164	0,2223
	BSS	2,7016	3,2975	0,2075	0,1959
Tank B	SpATS	1,3343	1,1445	0,1622	0,1523
	BSS	1,3056	1,1185	0,1660	0,1606

Table 4.8 presents the genetic variability (σ_g^2) and the heritability (H^2) of all variables for both models. The SpATS model seem to be slightly more effective at estimating genotypic effects since the genetic variance is always lower than for the BSS model.

Table 4.8: Comparison of both models in term of genetic variance (σ_g^2) and heritability (H^2)

	σ_g^2		H^2	
	SpATS	BSS	SpATS	BSS
FRESH_LS	0,1709	0,1829	0.72	0.7
FRESH_RS	0,2618	0,2644	0.79	0.75
DRY_LS	1,267E-03	1,339E-03	0.73	0.71
DRY_RS	6,902E-04	7,350E-04	0.83	0.77

The heritability used in this table is defined as the part of phenotypic variation that is due to the genotype (see section 3.4.4 of chapter 3 for more details). Just as for the

[2] All test were performed with a significance level, α = 0.05.

variances, the heritability is higher (and therefore better) for the SpATS model than for the BSS model. This is especially true for the DRY_RS variable, where the difference in heritability is doubled compared to the other variables. However, when we look at the estimates of the genotypic random effects between models on Figure 4.4, we see that these differences are not relevant since estimates from the two models are correlated to more than 99%. These high correlations were expected, given the similarity between the fitted values displayed in both models (see Figure 4.2 and Figure 4.3).

Furthermore, the genotype orders given by the models are very similar to the ones exposed in the sorted means boxplots (Figure 4.1), with the first two genotypes being 16 and 14. They clearly stand out in the comparison plots (Figure 4.4). The distinction of genotypes order in the DRY_RS model is less obvious, but it is still similar to the pattern picked up in the analysis of the weighted means. Without any additional information on the genotypes, it is hard to interpret this ranking in a meaningful way.

In conclusion of this comparison, both models yield very similar results in terms of fitted values. This similarity is reflected in the estimated tank effects and in the genotypic variances. However those variances were sligthly lower for the SpATS model, but that did not impact the estimates of the genotypic effects.

4.4.2 Parametrization

In terms of parametrization, both models are similar in their formulation of the spatial model as a mixed model. They include the same fixed effects and the same basic random effects (such as genotype). The major difference between both models is in the way they account for extra spatial variation. The BSS model fits a variogram over spatially dependant and independent residuals,using either an auto-regressive process or a linear variance structure, whereas the SpATS model uses tensor product of P-splines.

A main disadvantage of the BSS model is that it uses the basic mixed model to account for global trends and the fitted variogram to account for the local trends. This distinction between global and local can lead to variance underestimation issues (Zimmerman & Harville 1991) or local trends overestimation issues (Gumpertz & Brownie 1991), when the global trends are not well accounted for. Conversely, SpATS models all the field trends (both global and local) in a single continuous process, thus avoiding the distinction between global and local trends.

Another disadvantage is the process that the BSS is modelling. It contains both the random and the spatially correlated data as a single term. Theoretically, only the spatially dependent data should be captured by the model and the remaining variation should be random. However, several authors showed that a part of the random error is often modelled as spatially correlated data (Cullis, Gogel, *et al.* 1998; Piepho, Möhring, *et al.* 2015). This is especially true for models with high auto-correlation values like ours, where there is a possibility of confounding between rows and columns and spatial covariance structure. Nevertheless, these issues are dampened when using a linear variance structure, as we did with the dry weights.

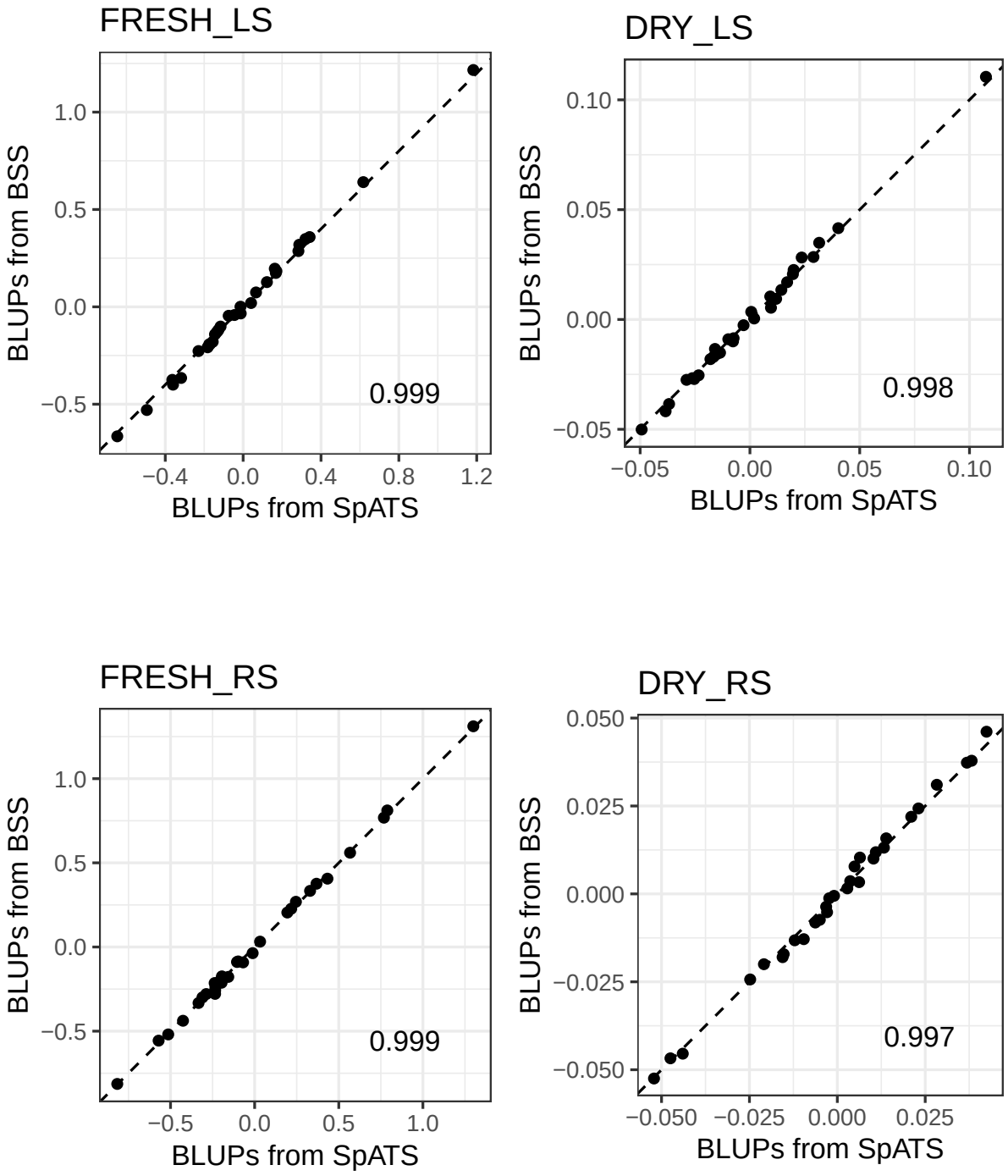


Figure 4.4: Comparison of the genotype BLUPs from the SpATS model and the BSS model, with the Pearson correlation (bottom right corner of each panel).

However, the weakness of the SpATS models resides in its covariance structure. The BSS model allows for more diverse way of modelling the spatial covariance structure, even using non-linear terms, whereas the SpATS model only has a linear structure. Since no studies have been conducted on this matter, it is hard to estimate how much it affects spatial modelling in field trials (Velazco *et al.* 2017).

Chapter 5

Conclusion

This thesis aimed to evaluate the efficiency of spatial modelling techniques on the UCLouvain phenotyping platform. More precisely, we wanted to assess the differences between a still and a moving growing tank in the platform, and to estimate the genotypic effects as precisely as possible with the use of spatial models for field trials. In parallel, we also wanted to evaluate the performance of a classical spatial model against the SpATS model, a new technique, that recently showed promising results in the framework of field trials analysis.

To achieve this goal, we created a custom experimental design, fitted to the platform and set up a phenotyping experiment. The results indicated a strong difference in yield between the tanks. This was confirmed by the results of both models, proving this effect to be highly significant. This indicated that the moving environment of the phenotyping platform was more profitable to plant growth, because they were less sensitive to local spatial variations. Moreover, this moving environment allowed a more precise and continuous characterization of the plants. The results also hinted at an important genotypic effect. Some genotypes expressed much lower yield than other, even more so, in the moving environment where the spatial influence is supposed to be less pronounced. These genotype features were also retrieved by both spatial models, that gave similar results.

In terms of performance, both spatial models gave similar results. They exposed the complexity of the spatial trends present in the data and were both able to account for it with the same magnitude. However, these models mainly differed in their parametrization and interpretation. The SpATS model was more versatile because it accounted for both local and global trends in a single process and did not require a stepwise selection process for the best model, like most classical linear models do. The interpretation of the parameters is also simpler with SpATS. The concept of effective dimensions introduced with the model allowed to easily see the contribution of each spatial component to the whole fitted spatial surface. Conversely the standard spatial model gave interpretable parameters, but the interpretation might not be as straightforward. Moreover, the SpATS model was also more flexible to pick up spatial patterns in highly heterogeneous environments, whereas the standard models might consider this heterogeneity as random error. Overall the SpATS model seemed to be a better choice for spatial modelling in the context of this experiment.

However, spatial modelling may not be the best tool to analyse data from a moving phenotyping platform. The very concept of row and column position is hard to translate to plants that are constantly shifting positions. Even if some spatial trends remain influential with those changing locations, the relative distance between plants is changing, making it hard to model spatial correlation correctly. Depending on the goal of the platform, the modelling strategy could be rethought to incorporate these challenges in a meaningful way.

This reflection opens interesting perspective for future research. If the identification of spatial trends is secondary, and the evaluation of the genotypes is the main objective, it could be interesting to compare the results of a generalized linear model (GLM) against the spatial models used here. Especially given the specificity of the moving environment. Furthermore, we could even more capitalize on the scan of the plants realized during the experiment. The root pictures could reveal interesting information that were not present, or not detected in the weights of the plants. The main perspective remains to incorporate the evolution of root growth in the model. Since several pictures of each plant were taken during the experiment, we have longitudinal data of each plant. This untapped potential could reveal even more insights on the difference between genotypes.

In conclusion, this thesis had both an agronomical interest, to estimate the genotypes effects and a statistical interest, to compare the efficiency of different spatial models. From a phenotyping point of view, this thesis has shown that this type of platform are helpful in identifying competitive genotypes. They could help relieving the current bottleneck and help the future of plant breeding. From the statistical angle, both models yielded similar results, but SpATS was the superior in terms of parametrization. This new model seems to be a very promising tool for spatial analysis of field trials on phenotyping platform.

Bibliography

1. Atkinson, A. C. Optimal Design. *Wiley StatsRef: Statistics Reference Online,* 1–17 (2014).
2. Atkinson, A. C. & Bailey, R. A. One hundred years of the design of experiments on and off the pages of Biometrika. en. *Biometrika* **88,** 53–97 (Feb. 2001).
3. Atkinson, A. C. & Donev, A. N. The construction of exact D-optimum experimental designs with application to blocking response surface designs. *Biometrika* **76,** 515–526 (1989).
4. Bohachevsky, I. O., Johnson, M. E. & Stein, M. L. Generalized simulated annealing for function optimization. *Technometrics* **28,** 209–217 (1986).
5. Brien, C. J., Berger, B., Rabie, H. & Tester, M. Accounting for variation in designing greenhouse experiments with special reference to greenhouses containing plants on conveyor systems. *Plant Methods* **9,** 5 (Feb. 2013).
6. Buja, A., Hastie, T., Tibshirani, R., *et al.* Linear smoothers and additive models. *The Annals of Statistics* **17,** 453–510 (1989).
7. Cabrera-Bosquet, L. *et al.* High-throughput estimation of incident light, light interception and radiation-use efficiency of thousands of plants in a phenotyping platform. en. *New Phytologist* **212,** 269–281 (Oct. 2016).
8. Cressie, N. Statistics for spatial data. *Terra Nova* **4,** 613–617 (1992).
9. Cullis, B. R. & Gleeson, A. C. Spatial Analysis of Field Experiments-An Extension to Two Dimensions. *Biometrics* **47,** 1449–1460 (1991).
10. Cullis, B. R., Smith, A. B. & Coombes, N. E. On the design of early generation variety trials with correlated data. en. *Journal of Agricultural, Biological, and Environmental Statistics* **11,** 381 (Dec. 2006).
11. Cullis, B., Gogel, B., Verbyla, A. & Thompson, R. Spatial Analysis of Multi-Environment Early Generation Variety Trials. *Biometrics* **54,** 1–18 (1998).
12. Currie, I. D. & Durban, M. Flexible smoothing with P-splines: a unified approach. en. *Statistical Modelling* **2,** 333–349 (Dec. 2002).
13. Davidoff, B., Lewis, J. W. & Selim, H. M. A method to verify the presence of a trend in studying spatial variability of soil temperature. English. *Soil Science Society of America journal (USA)* (1986).
14. Dierckx, P. *Curve and Surface Fitting with Splines* en (Clarendon Press, 1995).
15. Durban, M., Currie, I. D. & Kempton, R. A. Adjusting for fertility and competition in variety trials. en. *The Journal of Agricultural Science* **136,** 129–140 (Mar. 2001).

16. Eilers, P. H. C. & Marx, B. D. Flexible smoothing with B-splines and penalties. en. *Statistical Science* **11,** 89–121 (May 1996).
17. Eilers, P. H. C. & Marx, B. D. Multivariate calibration with temperature interaction using two-dimensional penalized signal regression. *Chemometrics and Intelligent Laboratory Systems* **66,** 159–174 (June 2003).
18. Eilers, P. H., Marx, B. D. & Durbán, M. Twenty years of P-splines. *SORT: statistics and operations research transactions* **39,** 0149–186 (2015).
19. Fagroud, M. & Van Meirvenne, M. Accounting for Soil Spatial Autocorrelation in the Design of Experimental Trials. en. *Soil Science Society of America Journal* **66,** 1134–1142 (July 2002).
20. Fahrmeir, L., Kneib, T., Lang, S. & Marx, B. *Regression: Models, Methods and Applications* en (Springer-Verlag, Berlin Heidelberg, 2013).
21. Federer, W. T. Recovery of interblock, intergradient, and intervariety information in incomplete block and lattice rectangle designed experiments. *Biometrics,* 471–481 (1998).
22. Fedorov, V. V. *Theory of optimal experiments* eng. Open Library ID: OL18496755M (Academic Press, New York, 1972).
23. Fiorani, F. & Schurr, U. Future Scenarios for Plant Phenotyping. en. *Annual Review of Plant Biology* **64,** 267–291 (Apr. 2013).
24. Furbank, R. T. & Tester, M. Phenomics – technologies to relieve the phenotyping bottleneck. en. *Trends in Plant Science* **16,** 635–644 (Dec. 2011).
25. Gilmour, A. R., Cullis, B. R. & Verbyla, A. P. Accounting for Natural and Extraneous Variation in the Analysis of Field Experiments. *Journal of Agricultural, Biological, and Environmental Statistics* **2,** 269–293 (1997).
26. Goos, P. & Jones, B. *Optimal Design of Experiments: A Case Study Approach* en. Google-Books-ID: EMWYkYd3sPoC (John Wiley & Sons, June 2011).
27. Gumpertz, M. & Brownie, C. Raleigh, North Carolina (1991).
28. Heredia-Langner, A., Carlyle, W. M., Montgomery, D. C., Borror, C. M. & Runger, G. C. Genetic algorithms for the construction of D-optimal designs. *Journal of Quality Technology* **35,** 28–46 (2003).
29. Heredia-Langner, A., Montgomery, D. C., Carlyle, W. M. & Borror, C. M. Model-robust optimal designs: A genetic algorithm approach. *Journal of Quality Technology* **36,** 263–279 (2004).
30. Houle, D., Govindaraju, D. R. & Omholt, S. Phenomics: the next challenge. en. *Nature Reviews Genetics* **11,** 855–866 (Dec. 2010).
31. Iqbal, J., Thomasson, J. A., Jenkins, J. N., Owens, P. R. & Whisler, F. D. Spatial Variability Analysis of Soil Physical Properties of Alluvial Soils. en. *Soil Science Society of America Journal* **69,** 1338 (2005).
32. Johnson, M. E. & Nachtsheim, C. J. Some guidelines for constructing exact D-optimal designs on convex design spaces. *Technometrics* **25,** 271–277 (1983).
33. Jung, J. S. & Yum, B. J. Construction of exact D-optimal designs by tabu search. *Computational Statistics & Data Analysis* **21,** 181–191 (1996).

34. Lado, B. *et al.* Increased Genomic Prediction Accuracy in Wheat Breeding Through Spatial Adjustment of Field Trial Data. en. *G3: Genes, Genomes, Genetics* **3,** 2105–2114 (Dec. 2013).

35. Lee Hwang, D.-J. Smoothing mixed models for spatial and spatio-temporal data. eng (May 2010).

36. Lee, D.-J. & Durbán, M. P-spline ANOVA-type interaction models for spatio-temporal smoothing. *Statistical Modelling* **11,** 49–69 (Feb. 2011).

37. Lee, D.-J., Durbán, M. & Eilers, P. Efficient two-dimensional smoothing with P-spline ANOVA mixed models and nested bases. *Computational Statistics & Data Analysis* **61,** 22–37 (May 2013).

38. Lobet, G. & Draye, X. Novel scanning procedure enabling the vectorization of entire rhizotron-grown root systems. en. *Plant Methods* **9,** 1 (2013).

39. Lobet, G., Draye, X. & Périlleux, C. An online database for plant image analysis software tools. *Plant Methods* **9,** 38 (Oct. 2013).

40. Lobet, G., Pagès, L. & Draye, X. A Novel Image-Analysis Toolbox Enabling Quantitative Analysis of Root System Architecture. en. *Plant Physiology* **157,** 29–39 (Sept. 2011).

41. Meyer, R. K. & Nachtsheim, C. J. Constructing Exact D-Optimal Experimental Designs by Simulated Annealing. *American Journal of Mathematical and Management Sciences* **8,** 329–359 (Feb. 1988).

42. Meyer, R. K. & Nachtsheim, C. J. The coordinate-exchange algorithm for constructing exact optimal experimental designs. English (US). *Technometrics* **37,** 60–69 (Jan. 1995).

43. Mooney, S. J., Pridmore, T. P., Helliwell, J. & Bennett, M. J. Developing X-ray Computed Tomography to non-invasively image 3-D root systems architecture in soil. en. *Plant and Soil* **352,** 1–22 (Mar. 2012).

44. Nielsen, D., Biggar, J. & Erh, K. Spatial variability of field-measured soil-water properties. English. *Hilgardia* **42,** 215–259 (Nov. 1973).

45. Oakey, H., Verbyla, A., Pitchford, W., Cullis, B. & Kuchel, H. Joint modeling of additive and non-additive genetic line effects in single field trials. en. *Theoretical and Applied Genetics* **113,** 809–819 (Sept. 2006).

46. Patterson, H. D. & Hunter, E. A. The efficiency of incomplete block designs in National List and Recommended List cereal variety trials. en. *The Journal of Agricultural Science* **101,** 427–433 (Oct. 1983).

47. Piepho, H., Möhring, J., Pflugfelder, M., Hermann, W. & Williams, E. Problems in parameter estimation for power and AR(1) models of spatial correlation in designed field experiments. *Communications in Biometry and Crop Science* **10,** 3–16 (2015).

48. Piepho, H. & Williams, E. Linear variance models for plant breeding trials. *Plant breeding* **129,** 1–8 (2010).

49. Pieruschka, R., Schurr, U., *et al.* Plant Phenotyping: Past, Present, and Future. *Plant Phenomics* **2019,** 7507131 (2019).

50. Pilarczyk, W. *The extent and prevailing shape of spatial relationship in Polish variety testing trials on cereals* in *Proceedings of the International Symposium "Agricultural Field Trials–Today and Tomorrow". Stuttgart. Grauer-Verlag* (2007), 153–159.

51. Pound, M. P. *et al.* RootNav: Navigating Images of Complex Root Architectures. en. *PLANT PHYSIOLOGY* **162,** 1802–1814 (Aug. 2013).

52. Risser, M. D. Nonstationary Spatial Modeling, with Emphasis on Process Convolution and Covariate-Driven Approaches. *arXiv preprint arXiv:1610.02447* (2016).

53. Rodríguez-Álvarez, M. X., Boer, M. P., van Eeuwijk, F. A. & Eilers, P. H. C. Spatial Models for Field Trials. en. *arXiv:1607.08255 [stat].* arXiv: 1607.08255 (July 2016).

54. Rodriguez-Alvarez, M. X., Boer, M. P., van Eeuwijk, F. A. & Eilers, P. H. Correcting for spatial heterogeneity in plant breeding experiments with P-splines. *Spatial Statistics* **23,** 52–71 (2018).

55. Rodríguez-Álvarez, M. X., Lee, D.-J., Kneib, T., Durbán, M. & Eilers, P. Fast smoothing parameter separation in multidimensional generalized P-splines: the SAP algorithm. *Statistics and Computing* **25,** 941–957 (2015).

56. Rodríguez-Álvarez, M., Boer, M., Eilers, P. & van Eeuwijk, F. *SpATS: spatial analysis of field trials with splines. R package version 1.0–4* (2016).

57. Rodríguez, M., Jones, B., Borror, C. M. & Montgomery, D. C. Generating and Assessing Exact G-Optimal Designs. *Journal of Quality Technology* **42,** 3–20 (Jan. 2010).

58. Schiml, S. & Puchta, H. Revolutionizing plant biology: multiple ways of genome engineering by CRISPR/Cas. *Plant methods* **12,** 8 (2016).

59. Tardieu, F., Cabrera-Bosquet, L., Pridmore, T. & Bennett, M. Plant Phenomics, From Sensors to Knowledge. en. *Current Biology* **27,** R770–R783 (Aug. 2017).

60. Tester, M. & Langridge, P. Breeding technologies to increase crop production in a changing world. *Science* **327,** 818–822 (2010).

61. Van Es, H. M. 1.2 Soil Variability. *Methods of Soil Analysis: Part 4 Physical Methods,* 1–13 (2002).

62. Van Es, H. M. Spatial Nature of Randomization and Its Effect on the Outcome of Field Experiments. en. *Agronomy Journal* **85,** 420–428 (1993).

63. Van Es, H. M., Gomes, C. P., Sellmann, M. & van Es, C. L. Spatially-Balanced Complete Block designs for field experiments. *Geoderma. Pedometrics 2005* **140,** 346–352 (Aug. 2007).

64. Vargas, M., van Eeuwijk, F. A., Crossa, J. & Ribaut, J.-M. Mapping QTLs and QTL × environment interaction for CIMMYT maize drought stress program using factorial regression and partial least squares methods. en. *Theoretical and Applied Genetics* **112,** 1009–1023 (Apr. 2006).

65. Velazco, J. G. *et al.* Modelling spatial trends in sorghum breeding field trials using a two-dimensional P-spline mixed model. en. *Theoretical and Applied Genetics* **130,** 1375–1392 (July 2017).

66. Verbyla, A., Cullis, B., Kenward, M. & Welham, S. The analysis of designed experiments and longitudinal data by using smoothing splines. *Journal of the Royal Statistical Society. Series C: Applied Statistics* **48,** 269–311 (1999).

67. Virlet, N., Sabermanesh, K., Sadeghi-Tehran, P. & Hawkesford, M. J. Field Scanalyzer: An automated robotic field phenotyping platform for detailed crop monitoring. en. *Functional Plant Biology* **44,** 143 (2017).

68. Wand, M. P. Smoothing and mixed models. en. *Computational Statistics* **18,** 223–249 (May 2003).

69. Watson, S. Spatial dependence and block designs in spaced plant herbage trials. en. *The Journal of Agricultural Science* **134,** 245–258 (May 2000).

70. Welham, S. J., Gogel, B. J., Smith, A. B., Thompson, R. & Cullis, B. R. A comparison of analysis methods for late-stage variety evaluation trials. *Australian & New Zealand Journal of Statistics* **52,** 125–149 (2010).

71. Wilkinson, G. N., Eckert, S. R., Hancock, T. W. & Mayo, O. Nearest Neighbour (NN) Analysis of Field Experiments. *Journal of the Royal Statistical Society. Series B (Methodological)* **45,** 151–211 (1983).

72. Williams, E. R. A neighbour model for field experiments. en. *Biometrika* **73,** 279–287 (Aug. 1986).

73. Williams, E. & Luckett, D. The use of uniformity data in the design and analysis of cotton and barley variety trials. en. *Australian Journal of Agricultural Research* **39,** 339–350 (1988).

74. Williams, K. A. Contemporary Ergonomics 1987: Proceedings of the Ergonomics Society's 1987 Annual Conference, Swansea, Wales, 6–10 April 1987. *American Journal of Occupational Therapy* **42,** 545–545 (1988).

75. Yates, F. THE COMPARATIVE ADVANTAGES OF SYSTEMATIC AND RANDOMIZED ARRANGEMENTS IN THE DESIGN OF AGRICULTURAL AND BIOLOGICAL EXPERIMENTS. en. *Biometrika* **30,** 440–466 (Jan. 1939).

76. Zimmerman, D. L. & Harville, D. A. A Random Field Approach to the Analysis of Field-Plot Experiments and Other Spatial Experiments. *Biometrics* **47,** 223–239 (1991).

Appendices

Appendix A

Additional informations on computation

A.1 Element-wise product

The element-wise product between two matrix $\mathbf{A}$ and $\mathbf{B}$ is noted $\mathbf{A} \odot \mathbf{B}$ and is defined in the following way:
For two matrices $\mathbf{A}$, $\mathbf{B}$ of same dimensions $n \times m$, the element-wise product is a $n \times m$ matrix where the elements are defined by:

$$(\mathbf{A} \odot \mathbf{B})_{i,j} = (\mathbf{A})_{i,j} \cdot (\mathbf{B})_{i,j}$$

The product is undefined for matrices of different dimensions

A.2 Kronecker product

The Kronecker product of two matrix $\mathbf{A}$ and $\mathbf{B}$ of respective dimensions $n \times m$ and $p \times q$ is a $np \times mq$ block matrix where the elements are defined by:

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \cdots & a_{mn}\mathbf{B} \end{bmatrix}$$

A.3 Polynomials splines

Fahrmeir *et al.* (2013) state that a function $f : [a, b] \to \mathbb{R}$ is called a polynomial spline of degree $l \geq 0$ with knots $a = \kappa_1 < \ldots < \kappa_m = b$, if it fulfills the following conditions:

1. $f(z)$ is $(l-1)$ times continuously differentiable. The special case of $l = 1$ corresponds to $f(z)$ being continuous (but not differentiable). We do not state any smoothness requirements for $f(z)$ when $l = 0$.

2. $f(z)$ is a polynomial of degree l on intervals $[\kappa_j, \kappa_{j+1})$ defined by the knots.

Moreover, it can be shown that each polynomial spline of degree l with knots $\kappa_1 < \ldots < \kappa_m$ can be uniquely determined as a linear combination of the $d = l + m - 1$ functions $B_1, \ldots, B_d$, called the *basis functions*, since we can uniquely represent all polynomials splines by using these functions.

A.3.1 B-splines

B-splines are polynomial splines with specific basis functions. B-spline basis functions are constructed from piecewise polynomials that are fused smoothly at the knots to achieve the desired smoothness constraints. More specifically, a B-spline basis function consists of $(l + 1)$ polynomial pieces of degree l , which are joined in an $(l - 1)$ continuously differentiable way. All B-spline basis functions are set up based on a given knot configuration. Using the complete basis, the function $f(z)$ can again be represented through a linear combination of $d = m + l - 1$ basis functions, i.e.,

$$f(z) = \sum_{j=1}^{d} \gamma_j B_j(z).$$

The B-splines of order $l = 0$ can be written as

$$B_j^0(z) = \begin{cases} 1 & \kappa_j \leq z < \kappa_{j+1} \\ 0 & \text{otherwise} \end{cases} \qquad j = 1, \ldots, d-1$$

and the B-splines for higher order l can be written as

$$B_j^l(z) = \frac{z - \kappa_{j-l}}{\kappa_j - \kappa_{j-l}} B_{j-1}^{l-1}(z) + \frac{\kappa_{j+1} - z}{\kappa_{j+1} - \kappa_{j+1-l}} B_j^{l-1}(z).$$

The estimation of a polynomial spline in B-spline representation can be traced back to the estimation of a linear model with a large number of parameters and design matrix

$$\mathbf{Z} = \begin{pmatrix} B_1^l(z_1) & \ldots & B_d^l(z_1) \\ \vdots & & \vdots \\ B_1^l(z_n) & \ldots & B_d^l(z_n) \end{pmatrix}.$$

The linear combination of basis functions can then be written in matrix form

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\gamma}$$

where the coefficient matrix, $\boldsymbol{\gamma}$ can be estimated using least squares.
The estimation of a B-spline fit can be summarized in three steps:

1. We calculate a complete B-spline basis for a given number of knots.
2. The least squares estimate $\hat{\boldsymbol{\gamma}}$ yields an amplitude $\hat{\gamma}_j$ for the scaling of every basis function.
3. We obtain the final estimate by summing the scaled basis function.

A.3.2 Penalized splines

We clearly see that the quality of the estimation by polynomials splines highly depends on the number of knots and that this can easily lead to an over-fitting issue. To overcome this problem, *penalized splines (P-splines)* introduce a roughness penalty term that prevents over-fitting and minimize a *penalized least squares (PLS) criterion* instead of the usual least squares criterion.
To characterize the smoothness of any type of function, the use of (squared) derivatives is appropriate, since these represent measures for the variability of a function. Therefore penalties based on the second derivative, such as

$$\lambda \int \left(f''(z)\right)^2 dz,$$

are particularly attractive since they measure the curvature of a function. Since we know that the first derivative of a B-spline can be written as a function of the first differences of the corresponding coefficient vector, we can use differences of a higher order r if we aim at a smooth function in terms of rth-order derivatives. This leads to the penalized residual sum of squares

$$\mathrm{PLS}(\lambda) = \sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{d} \gamma_j B_j\left(z_i\right) \right)^2 + \lambda \sum_{j=r+1}^{d} \left(\Delta^r \gamma_j\right)^2,$$

where Δ^r denotes the rth-order differences.The smoothing parameter $\lambda \geq 0$ controls the compromise between fidelity to the data and smoothness of the resulting function estimate. The PLS criterion can be rewritten using matrix notation

$$\mathrm{PLS}(\lambda) = (\mathbf{y} - \mathbf{Z}\boldsymbol{\gamma})'(\mathbf{y} - \mathbf{Z}\boldsymbol{\gamma}) + \lambda \boldsymbol{\gamma}' \mathbf{K_r} \boldsymbol{\gamma}$$

where $\mathbf{K_r}$ is the rth-order difference penalty matrix, and can be decomposed as $\mathbf{D_r}\prime\mathbf{D_r}$ with D_r the rth-order difference matrix. The smoothing parameter $\lambda \geq 0$ controls the compromise between fidelity to the data and smoothness of the resulting function estimate. The PLS estimate of the coefficient matrix is then

$$\hat{\boldsymbol{\gamma}} = \left(\mathbf{Z}'\mathbf{Z} + \lambda \mathbf{K}\right)^{-1} \mathbf{Z}'\mathbf{y}.$$

For more detailed information about polynomials splines, please refer to Fahrmeir *et al.* (2013) and Eilers & Marx (1996)

A.4 Penalized form of the solution and smoothing parameter selection

Let us consider the model only containing a bivariate smooth surface and an error term:

$$y_i = f\left(u_i, v_i\right) + \varepsilon_i, \text{ with } \varepsilon_i \sim N\left(0, \sigma^2\right).$$

It can be rewritten, in matrix notation, as the tensor product of B-splines:

$$\boldsymbol{y} = \boldsymbol{B}\boldsymbol{\alpha} + \boldsymbol{\epsilon}.$$

Since the model is purely parametric, it can be estimated by minimizing the residual sum of squares (with explicit solution $\hat{\boldsymbol{\alpha}} = (\mathbf{B}^t\mathbf{B})^{-1}\mathbf{B}^t\mathbf{y}$). To prevent over-fitting, Eilers & Marx (1996) propose to incorporate a discrete penalty on the coefficient associated to adjacent B-splines. For the two-dimensional case, the vector $\boldsymbol{\alpha}$ can be seen as an $(L \times P)$ matrix of coefficients, $\mathbf{A} = [\alpha_{lp}]$. Now the rows and columns of $\mathbf{A}$ correspond to the regression coefficients in the v and u direction, respectively. In anisotropic (direction-dependant) P-splines, a different amount of smoothing is assumed along the u and v directions. It leads to two penalties: one on all rows of $\mathbf{A}$, the other on all of its columns; and the penalized least squares objective function becomes (Eilers & Marx 2003)

$$\begin{aligned} S^* = & \underbrace{\|\boldsymbol{y} - \boldsymbol{B}\boldsymbol{\alpha}\|^2}_{\text{Original objective function}} \\ & + \underbrace{\hat{\lambda}\|\hat{\boldsymbol{D}}A\|_F^2}_{\text{Penalty along the columns}} \\ & + \underbrace{\breve{\lambda}\left\|\boldsymbol{A}\breve{D}^t\right\|_F^2}_{\text{Penalty along the rows}} \\ = & \|\boldsymbol{y} - \boldsymbol{B}\boldsymbol{\alpha}\|^2 + \boldsymbol{\alpha}^t\boldsymbol{P}\boldsymbol{\alpha}, \end{aligned}$$

where $\boldsymbol{P} = \hat{\lambda}\left(\boldsymbol{I}_P \otimes \hat{\boldsymbol{D}}^t\hat{\boldsymbol{D}}\right) + \breve{\lambda}\left(\breve{\boldsymbol{D}}^t\breve{\boldsymbol{D}} \otimes \boldsymbol{I}_L\right)$ is the penalty matrix, $\hat{\lambda}$ and $\breve{\lambda}$ are the smoothing parameters acting, respectively, on the columns and rows of $\mathbf{A}$, and $\hat{\mathbf{D}}$ and $\breve{\mathbf{D}}$ are the matrices that form differences of order d_u and d_v respectively. The minimizer of the starting equation then becomes

$$\widehat{\boldsymbol{\alpha}} = \left(\boldsymbol{B}^t\boldsymbol{B} + \boldsymbol{P}\right)^{-1}\boldsymbol{B}^t\boldsymbol{y},$$

which is the penalized form of the solution. However, $\boldsymbol{P}$ is rank-deficient. To have a full rank penalty matrix, the key is to rewrite the model and setting

$$\mathbf{B}\boldsymbol{\alpha} = \boldsymbol{X}_s\boldsymbol{\beta}_s + \boldsymbol{Z}_s\boldsymbol{c}_s.$$

There are now two bases: $\mathbf{X}_s$, with coefficients that are not penalized at all, and $\mathbf{Z}_s$, with a size penalty on its coefficients. This decomposition follows the proposal by Lee & Durbán (2011), based on eigenvalue decomposition which gives rise to a diagonal penalty matrix.

The two bases have the following structures:

$$\boldsymbol{X}_s = [\mathbf{1}_n, \boldsymbol{u}, \boldsymbol{v}, \boldsymbol{u} \odot \boldsymbol{v}] \quad \text{and} \quad \mathbf{Z}_s = [\mathbf{Z}_v, \mathbf{Z}_u, \mathbf{Z}_v \boxdot \mathbf{u}, \mathbf{v} \boxdot \mathbf{Z}_u, \mathbf{Z}_v \boxdot \mathbf{Z}_u],$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are still, respectively, the vectors of row and column positions. Here $\mathbf{Z}_u$ and $\mathbf{Z}_v$ are penalized version of the B-splines basis $\breve{\mathbf{B}}$ (rows) and $\hat{\mathbf{B}}$ (columns). This new way of writing the problem leads to another penalty matrix $\widetilde{\boldsymbol{P}}$, which is a block diagonal matrix. Each block of $\widetilde{\boldsymbol{P}}$ corresponds to a block in $\mathbf{Z}_s$. Similarly to $\boldsymbol{P}$, the penalty matrix of the previous section, this new penalty matrix only depends on the two tuning parameters $\hat{\lambda}$ (smoothing along the columns) and $\breve{\lambda}$ (smoothing along the rows). Figure A.1 presents a diagram clarifying the structures and relations of the different

matrices presented throughout this section.

This reformulation provides the ANOVA type decomposition discussed in the previous section (3.4.03), and explains how the bilinear smooth surface can be modelled using P-splines and tensor products of P-splines. The block structure of $\mathbf{X}_s$ and $\mathbf{Z}_s$ implies

$$\begin{aligned} f(\boldsymbol{u},\boldsymbol{v}) &= \boldsymbol{X}_s\boldsymbol{\beta}_s + Z_s\boldsymbol{c}_s \\ &= \mathbf{1}_n\beta_0 + \boldsymbol{u}\beta_1 + \boldsymbol{v}\beta_2 + \boldsymbol{u}\odot\boldsymbol{v}\beta_3 \\ &+ \underbrace{f_v(\boldsymbol{v})}_{\boldsymbol{Z}_v\boldsymbol{c}_{s1}} + \underbrace{f_u(\boldsymbol{u})}_{\boldsymbol{Z}_u\boldsymbol{c}_{s2}} + \underbrace{\boldsymbol{u}\odot h_v(\boldsymbol{v})}_{[\boldsymbol{Z}_v\square\boldsymbol{u}]\boldsymbol{c}_{s3}} + \underbrace{\boldsymbol{v}\odot h_u(\boldsymbol{u})}_{[\boldsymbol{v}\square\boldsymbol{Z}_u]\boldsymbol{c}_{s4}} + \underbrace{f_{u,v}(\boldsymbol{u},\boldsymbol{v})}_{[\boldsymbol{Z}_v\square\boldsymbol{Z}_u]\boldsymbol{c}_{s5}}, \end{aligned}$$

where $\boldsymbol{c}_{sk}$ $(k = 1,\ldots,5)$ contains the elements of $\boldsymbol{c}_s$ that correspond to the kth block of $\boldsymbol{Z}_s$, i.e. $\boldsymbol{c}_s = (\boldsymbol{c}_{s1}^t,\ldots,\boldsymbol{c}_{s5}^t)^t$. The details about the specific block component of $\boldsymbol{Z}_s$ and the computation of the new penalty matrix are available in Rodriguez-Alvarez *et al.* (2018) and the appendices therein.

Therefore, using this new notation, model (3.4.11) that only contains a smooth bivariate surface and an error term can be rewritten in the following way:

$$\boldsymbol{y} = \boldsymbol{X}_s\boldsymbol{\beta}_s + \boldsymbol{Z}_s\boldsymbol{c}_s + \boldsymbol{\varepsilon},\ \text{with}\ \boldsymbol{\varepsilon}\sim N\left(\mathbf{0},\sigma^2\boldsymbol{I}_n\right)\ \text{and}\ \boldsymbol{c}_s\sim N\left(\mathbf{0},\boldsymbol{G}_s\right),$$

where $\boldsymbol{G}_s = \sigma^2\widetilde{\boldsymbol{P}}^{-1}$ and also has a block diagonal structure, similar to that of $\widetilde{\boldsymbol{P}}$ (this structure is also represented on figure A.1). However, $\boldsymbol{G}_s$ depends on two different parameters, $\breve{\sigma}^2 = \sigma/\breve{\lambda}$ and $\hat{\sigma}^2 = \sigma/\hat{\lambda}$, which are variance parameters. As shown in the diagram in Figure A.1, the same variance parameters control the smoothness of the both the main effects and interactions terms. This prevents the use of standard mixed models software for estimation since $\mathbf{G}_s$ has its last block depending on both $\breve{\sigma}^2$ and $\hat{\sigma}^2$, but in a non-linear way. Even though Rodríguez-Álvarez, Lee, *et al.* (2015) presented a specialized algorithm to deal with this issue, here the PS-ANOVA decomposition approach (Lee, Durbán & Eilers 2013) is used to allow the use of standard mixed model estimation procedures. Lee, Durbán & Eilers (2013) therefore propose to use a different variance component for each smooth component in $\mathbf{G}_s$, thus redefining this matrix as a linear function of variance parameters:

$$\boldsymbol{G}_s = \bigoplus_{k=1}^{5}\boldsymbol{G}_{sk} = \ \text{blockdiag}\ \left(\boldsymbol{G}_{s1},\boldsymbol{G}_{s2},\boldsymbol{G}_{s3},\boldsymbol{G}_{s4},\boldsymbol{G}_{s5}\right),$$

where $\boldsymbol{G}_{sk}$ is the kth block of the $\mathbf{G}_s$ matrix, depending on the specific variance component σ^2_{sk}. In other words, here the tensor product P-splines mixed model is represented as the sum of 5 sets of mutually independent Gaussian random components $\mathbf{c}_{sk}$, each depending on one variance σ^2_{sk} $(k = 1,\ldots,5)$.

Within this mixed model framework, the smoothing parameters, defined earlier as the ratio between the residual variance and the corresponding variance effect $\lambda_{sk} = \sigma^2_e/\sigma^2_{sk}$, are determined by restricted maximum likelihood (REML). Therefore the smoothness of the spatial surface is tuned by five distinct parameters, applying anisotropic (direction-dependant) smoothing. This parametrization provides flexibility to account for both global and local variations in the field. Furthermore, the decomposition of

$f(\boldsymbol{u}, \boldsymbol{v})$ enables a more explicit interpretation of the main patterns of spatial variation (Rodriguez-Alvarez *et al.* 2018).

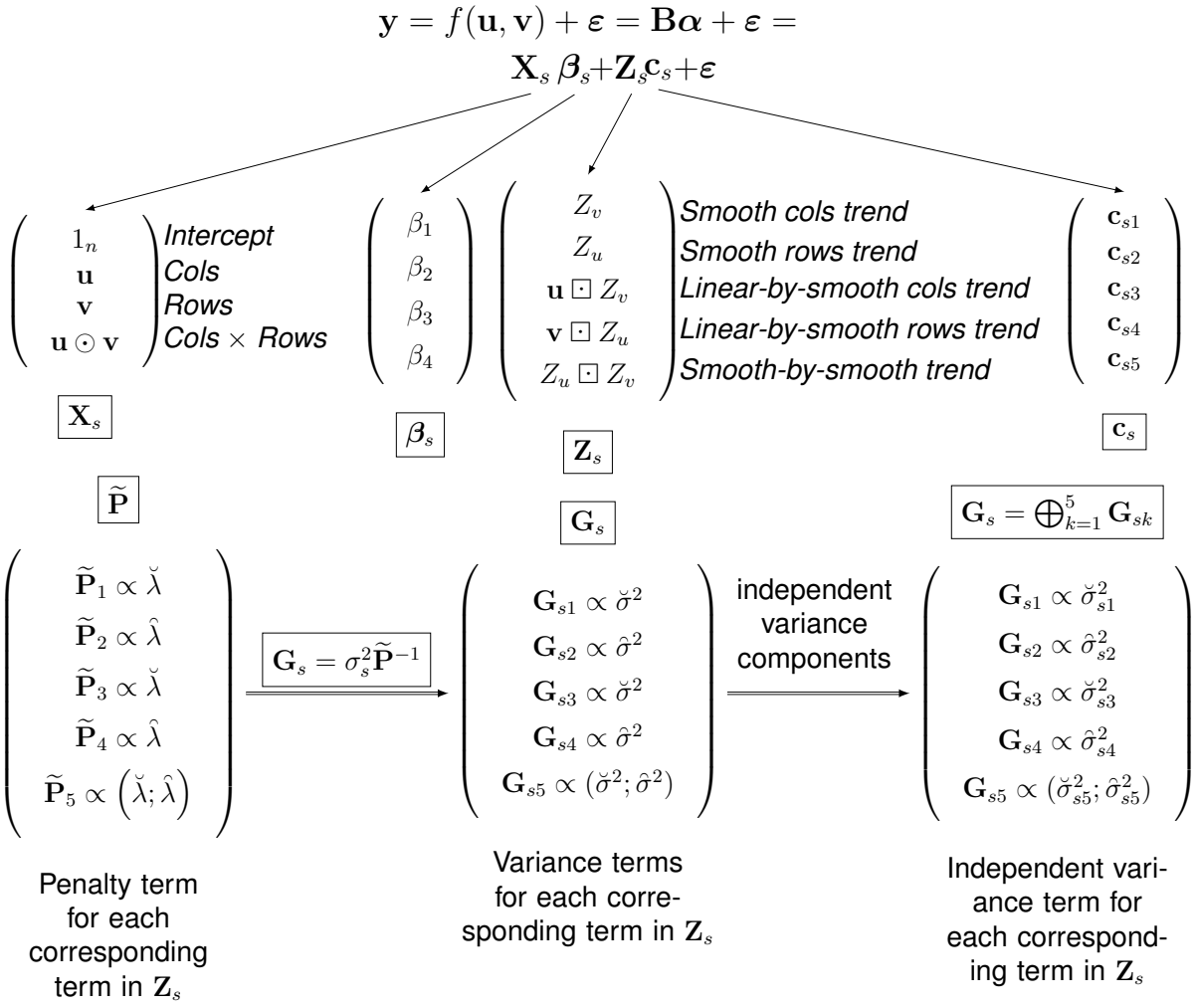


Figure A.1: Diagram detailing the structure of the matrices used in this section. All matrices are block diagonal matrix with each element represented on the diagram, being an individual block. The symbol $\propto$ shows how each block of the $\widetilde{\mathbf{P}}/\mathbf{G}_s$ matrix relates to the tuning/variance parameters. The last block of both the $\widetilde{\mathbf{P}}$ and $\mathbf{G}_s$ matrices depends on both parameters but in a non-linear way.

Appendix B

Additional figures and tables

B.1 Descriptive statistics

Table B.1: Weighted mean and standard deviation for each genotype. $\mathbf{DRY_{LS}}$ represents the dry weight of the leaf system; $\mathbf{DRY_{RS}}$, the dry weight for the root system; $\mathbf{FRESH_{LS}}$, the fresh weight for the leaf system and $\mathbf{FRESH_{RS}}$, the fresh weight for the root system. All the results are presented as mean $\pm$ standard deviation (g)

Genotype	DRY_{LS}	DRY_{RS}	$FRESH_{LS}$	$FRESH_{RS}$
1	0.2267 $\pm$ 0.0869	0.2354 $\pm$ 0.0698	2.7231 $\pm$ 1.1612	3.4447 $\pm$ 1.4431
2	0.2113 $\pm$ 0.0993	0.1964 $\pm$ 0.06	2.4058 $\pm$ 1.254	2.2725 $\pm$ 1.1119
3	0.2132 $\pm$ 0.0747	0.1939 $\pm$ 0.0474	2.406 $\pm$ 1.0814	2.8146 $\pm$ 1.3663
4	0.227 $\pm$ 0.1148	0.1824 $\pm$ 0.0861	2.45 $\pm$ 1.5508	2.7773 $\pm$ 2.1926
5	0.2126 $\pm$ 0.113	0.1829 $\pm$ 0.0606	2.6521 $\pm$ 1.5044	3.0241 $\pm$ 1.7336
6	0.2024 $\pm$ 0.0739	0.1747 $\pm$ 0.051	2.4244 $\pm$ 1.1691	3.2168 $\pm$ 1.4545
7	0.126 $\pm$ 0.0441	0.1251 $\pm$ 0.0232	1.4118 $\pm$ 0.5677	1.5669 $\pm$ 0.5783
8	0.186 $\pm$ 0.0805	0.1798 $\pm$ 0.0567	2.3614 $\pm$ 1.1696	2.6894 $\pm$ 1.273
9	0.1559 $\pm$ 0.0865	0.2064 $\pm$ 0.0574	1.9704 $\pm$ 1.1782	2.3862 $\pm$ 1.3949
10	0.1885 $\pm$ 0.0875	0.1769 $\pm$ 0.0693	2.1228 $\pm$ 1.046	2.029 $\pm$ 1.0451
11	0.1789 $\pm$ 0.0811	0.1783 $\pm$ 0.0521	2.1608 $\pm$ 1.0747	2.339 $\pm$ 1.2172
12	0.1684 $\pm$ 0.0893	0.1469 $\pm$ 0.0417	2.0166 $\pm$ 0.9281	2.0985 $\pm$ 0.9015
13	0.1927 $\pm$ 0.0794	0.1639 $\pm$ 0.054	2.1326 $\pm$ 1.077	2.1176 $\pm$ 1.1582
14	0.2438 $\pm$ 0.1288	0.2173 $\pm$ 0.0796	3.0901 $\pm$ 1.6735	3.486 $\pm$ 1.8646
15	0.1175 $\pm$ 0.0885	0.2097 $\pm$ 0.0591	1.48 $\pm$ 1.2962	1.8417 $\pm$ 1.3372
16	0.3244 $\pm$ 0.1276	0.2208 $\pm$ 0.0879	3.7094 $\pm$ 1.8196	3.9028 $\pm$ 2.2598
17	0.2357 $\pm$ 0.1111	0.2019 $\pm$ 0.065	2.7519 $\pm$ 1.2418	2.8555 $\pm$ 1.4
18	0.1427 $\pm$ 0.039	0.1653 $\pm$ 0.0442	1.5621 $\pm$ 0.5627	1.7781 $\pm$ 0.8075
19	0.1785 $\pm$ 0.0505	0.1926 $\pm$ 0.04	2.0917 $\pm$ 0.7454	2.7726 $\pm$ 0.9668
20	0.1901 $\pm$ 0.0721	0.1506 $\pm$ 0.0497	2.0641 $\pm$ 0.9693	2.2429 $\pm$ 1.2908
21	0.1649 $\pm$ 0.0786	0.1859 $\pm$ 0.0329	1.9682 $\pm$ 0.8601	1.7689 $\pm$ 0.6407
22	0.1482 $\pm$ 0.0669	0.1508 $\pm$ 0.0296	1.522 $\pm$ 0.7799	1.793 $\pm$ 0.838
23	0.156 $\pm$ 0.04	0.1574 $\pm$ 0.0372	1.7621 $\pm$ 0.6537	1.9874 $\pm$ 0.9962
24	0.1668 $\pm$ 0.0827	0.1236 $\pm$ 0.0575	1.9357 $\pm$ 0.8693	2.0506 $\pm$ 1.2218
25	0.2013 $\pm$ 0.0856	0.1195 $\pm$ 0.0481	2.4338 $\pm$ 1.0937	1.9605 $\pm$ 1.1428
26	0.1463 $\pm$ 0.075	0.1573 $\pm$ 0.0365	1.7995 $\pm$ 1.1967	1.9256 $\pm$ 1.0053
27	0.1682 $\pm$ 0.0894	0.188 $\pm$ 0.0401	2.2662 $\pm$ 1.0855	2.3647 $\pm$ 1.0382
28	0.1753 $\pm$ 0.0865	0.2244 $\pm$ 0.066	2.0357 $\pm$ 1.0699	2.441 $\pm$ 1.1785
29	0.215 $\pm$ 0.0934	0.1823 $\pm$ 0.0663	2.4861 $\pm$ 1.3776	2.6743 $\pm$ 1.5787
30	0.1573 $\pm$ 0.0592	0.184 $\pm$ 0.0485	2.1033 $\pm$ 0.8712	2.51 $\pm$ 1.1265

B.2 T-tests

Table B.2: P-values resulting from the individual t-tests of the differences between means for all genotypes. All values inferior to 0.05 are colored in green, while value between 0.05 and 0.1 are in yellow and the rest is in red. P-value was not computed for genotype 30 since there were only 3 values available.

Genotype	FRESH_LS	FRESH_RS	DRY_LS	DRY_LS
6	0,067045	8,18E-05	0,024774	0,000941
23	0,005072	0,009236	0,076901	0,036332
10	0,007526	0,001057	0,10752	0,054917
26	0,068053	0,029225	0,092141	0,007438
4	0,034853	0,024024	0,109144	0,032899
8	0,009049	0,00777	0,035141	0,201765
14	0,04338	0,023285	0,256493	0,09208
9	0,038485	0,039673	0,220818	0,123382
24	0,042817	0,015859	0,408412	0,079817
3	0,07032	0,020974	0,492672	0,060329
29	0,066428	0,055044	0,348994	0,183319
25	0,086712	0,013015	0,564199	0,091173
16	0,015628	0,002877	0,637848	0,164959
27	0,068433	0,001215	0,728604	0,036351
15	0,249287	0,264329	0,299309	0,383055
18	0,25502	0,054048	0,806946	0,14867
13	0,278165	0,130413	0,634507	0,287053
5	0,112019	0,047269	0,439077	0,742416
12	0,302533	0,792875	0,051195	0,271333
1	0,027344	0,004086	0,93751	0,512452
20	0,178574	0,110094	0,943514	0,447728
19	0,170245	0,038213	0,684042	0,87518
7	0,119486	0,017251	0,915835	0,878908
28	0,360779	0,127756	0,954177	0,606738
21	0,574376	0,503482	0,42371	0,570469
2	0,453042	0,077753	0,790318	0,828287
11	0,875426	0,772947	0,404975	0,179616
22	0,793643	0,55334	0,510492	0,40223
17	0,887499	0,192592	0,493846	0,835214
30	.	.	.	.

B.3 Variance tables

B.3.1 SpATS variance table

Table B.3: Individual variances of all the components of the SpATS model.

	FRESH_LS	FRESH_RS	DRY_LS	DRY_RS
$\mathbf{c}_g$	0.171	0.262	1.27×10^{-3}	6.9×10^{-4}
$\mathbf{c}_v$	3.24×10^{-3}	5.08×10^{-3}	1.95×10^{-7}	8.47×10^{-17}
$\mathbf{c}_u$	1.86×10^{-4}	2.25×10^{-5}	2.42×10^{-7}	3.93×10^{-8}
$f_v(\mathbf{v})$	2.2	11	1.06×10^{-4}	0.265
$f_u(\mathbf{u})$	2.75×10^{-5}	5.36×10^{-5}	5.4×10^{-8}	1.55×10^{-8}
$\boldsymbol{u} \odot h_v(\boldsymbol{v})$	5.60×10^{-52}	1.76×10^{-38}	1.51×10^{-52}	1.44×10^{-13}
$\boldsymbol{v} \odot h_u(\boldsymbol{u})$	1.81×10^{-9}	4.83×10^{-4}	6.75×10^{-10}	1.43×10^{-6}
$f_{u,v}(\boldsymbol{u}, \boldsymbol{v})$	0.398	0.252	6.13×10^{-5}	7.13×10^{-4}
ϵ	4.707	5.014	0.03364	0.01219

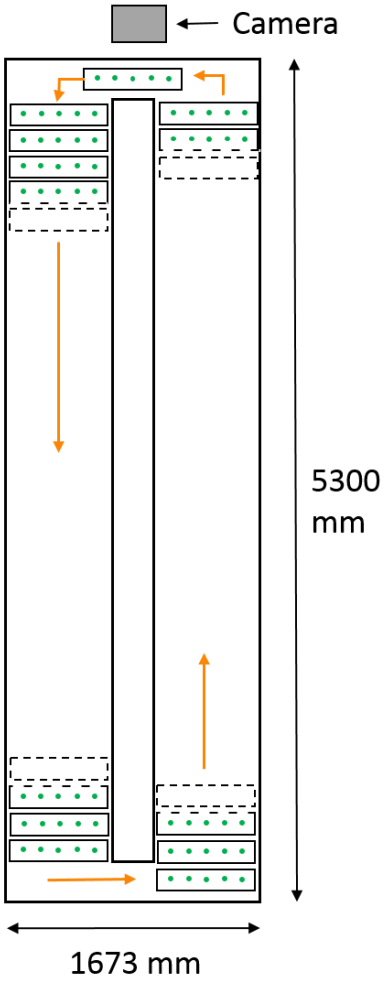
Appendix C

Phenotyping platform information file

Platform name

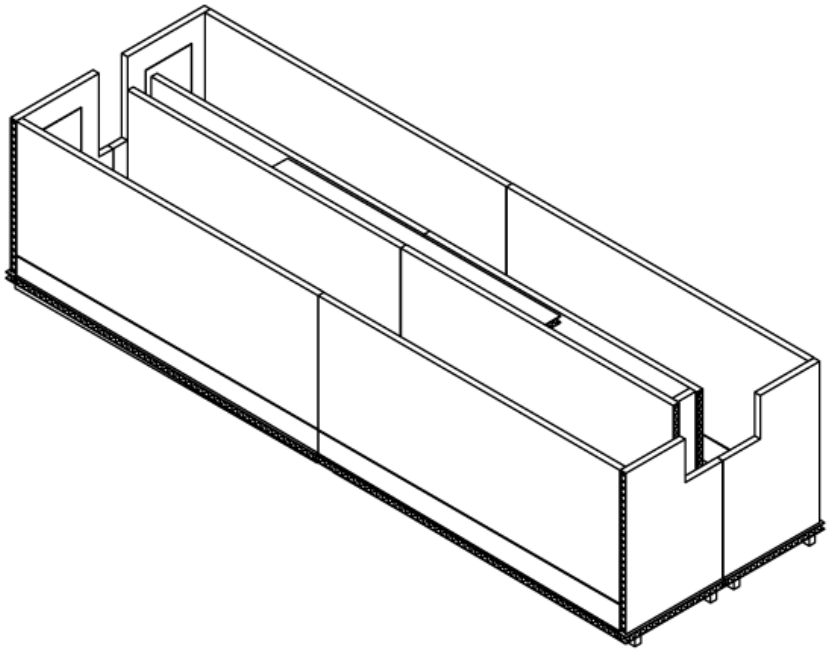
Partner site	UCL
Site and installation	Site: Louvain-la-Neuve, Installation: Aeroponics
Contact person(s)	Xavier Draye xavier.draye@uclouvain.be

Description of the platform structure

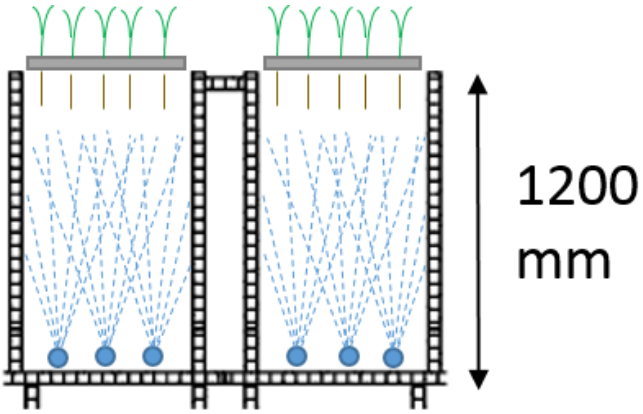


JRA2 - Jan. 2018

Aeroponic tank: plants are hold on strips, 5 plants per strip (green dots on layout). There are 99 strips in the tank for a total of 495 plants/tank. Strips move in the direction indicated by orange arrows. A full revolution takes 2 hours. When strips pass in front of the camera, at the top of the layout, plants are imaged individually.



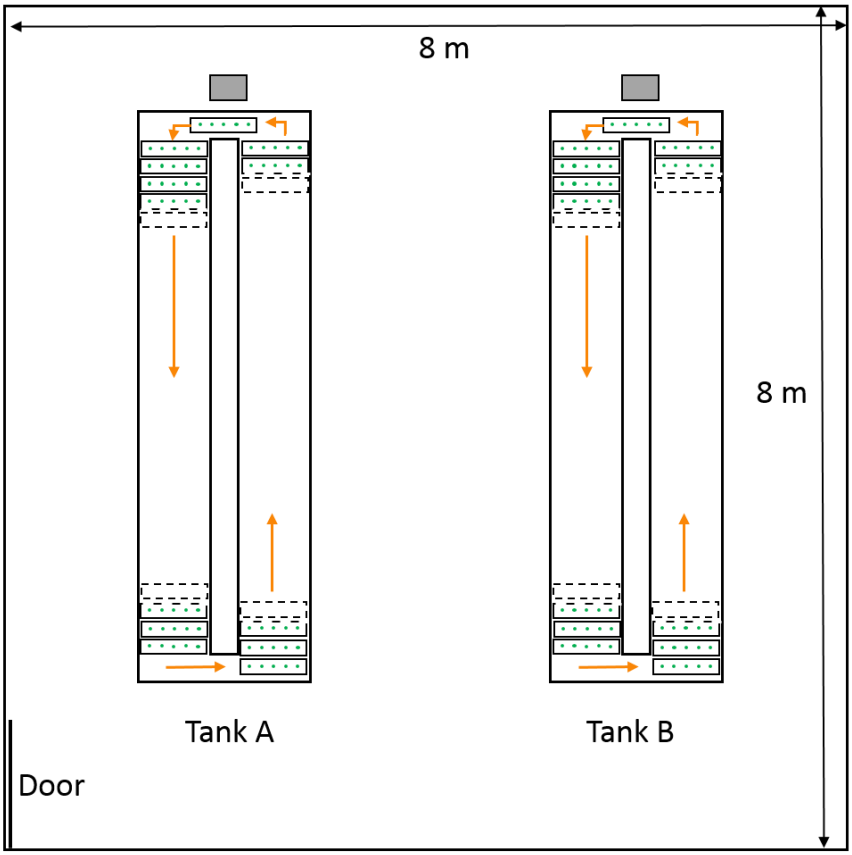
3D view of one tank, without the strips.



Transversal view of the aeroponic tank: 3 sprinklers are placed regularly in the bottom of each side of the tank. The sprinklers spray nutrient solution at regular interval, set by the operator. The spraying

pattern (interval and duration) can be differentiated between day and night and can be modified at any moment of the experiment.

2 identical tanks are available in the installation, located next to each other in the same greenhouse.

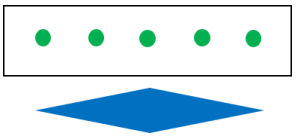


Sources and directions (if known) of environmental variations in the installation

1) Between the 2 tanks.

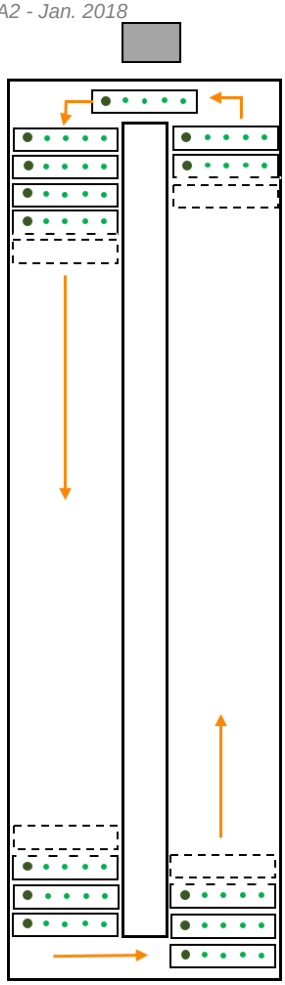
2) The side of the tank placed along the greenhouse wall may be warmer than the side near the centre of the greenhouse because of the presence of heating pipes along the walls.

3) Inside each tank, between plants that grow in the middle of the strip as compared to plants growing at the border of the strip. We suppose that the plants at the extremity of the strip may receive a bit less water than the others.



Layout of a strip with supposed variation of water availability: more water in the middle and less in the border

4) Last year, we observed that the plants growing on the left side of the strips were growing faster than the ones growing on the right side. We understood that the lamps were not exactly centred in the middle of the tank. We moved the lamps to put them exactly at the centre of each tank but we haven't done any new experiment yet.



Layout representing the plants that grow faster on the left side of the strips. The plants keep moving inside the tank but the left/right distinction is maintained during the whole experiment.

As strips keep moving within each tank, we don't expect to observe environmental variation between the different strips of each tank.

Description of experimental design and randomization and a motivation for the design and the randomization

- Design

Completely randomized design: individual plants are located in a strip and at a position randomly with Excel.
+ 2 treatments (eg: shadow, change of nutrient solution properties…) corresponding to the 2 tanks
OR 2 blocks corresponding to the 2 tanks

- Design specifications

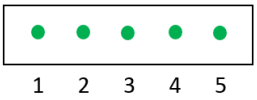
- Motivation

How plant positions are defined and recorded in the experiment

How are the pot positions defined according to the design, i.e. how are the spatial coordinates defined (see example 6)?

QR code associated to each plant

Number of the QR code:
Ex: B_76_5
B: tank id (A or B)
76: strip id (from 1 to 99)
5: position in the strip (from 1 to 5)



If pots are rearranged during the experiment, how is the change in spatial position recorded?

All strips move at the same pace. Each plant passes every 2 hour in front of the camera, where a picture is taken. The time of the picture enables to record the moment at which each plant passes in front of the camera. It would be possible to compute the pathway the plant had in the tank between two pictures.
No changes between the two tanks or within each strip (position 1 to 5)

If repeated measurements are taken, at what times are these taken?

Every 2 hours, 24h a day

Leuven Statistics Research Centre (LStat)
Celestijnenlaan 200 B
3001 HEVERLEE, BELGIË
tel. +32 16 32 88 75
https://lstat.kuleuven.be/contact