1. **Research question**

The main research question is : “how good is the SpATS model at estimating the variability in plant yield in a phenotyping platform”

How does the model estimate variability specifically attributable to our phenotyping platform?

How does the model compare to a standard spatial model?

How much of the total variability is accounted for by the genotypes and how well the model estimates it?

1. **Objectives**

From those questions, we derived 4 main objectives:

* Create an adequate design to answer our questions as precisely as possible
* Compare the efficiency of 2 spatial models in the conditions of the platform
* Estimate the main sources and trends of spatial variability in the platform
* Estimate the genotypic variability, see if it is relevant for in the model, and compare the estimates between models

1. **Methodology**

First, we created an experimental design. We implemented it onto the platform by setting up an experiment. We harvested the plants and measured the fresh and dry weights. Using those yield data, we ran 2 spatial model. A standard one, versus a recent model, called the SpATS model. From the output, we compared the fitted values, the mains spatial trends and the genotypes estimates.

1. **Phenotyping platform**

The platform is located in LLN. It consists of 2 tanks of 495 plants dispatched on 99 strips of 5 plants. Each tank has a camera that can take a picture of the root system of each plant every 2 hours. To better characterize environmental variability, 1tank was moving all the time, with pictures taken once a day and 1 tank was let still in the same position with only 1 rotation per day to take pictures. Since the experiment lasted 15 days, it resulted in 15 pictures of each plant. Each plant was identified using a specific QR CODE.

1. **Design factors**

There were 4 factors in our experimental design. The 2 tanks, the 99 strips, the 5 position

per strip and then finally the 30 genotypes. Since there was a hierarchical order in the

factors, it created blocking. The 2 tanks corresponded to 2 whole plots and the 99 strips

corresponded to 99 sub-plots.

The design was computed on JMP us D-optimality as criteria of choice. The design set up this

way was too complex to converge so instead we ran a design with 33 strips and triplicated it.

1. **Design output**

Here is a graphical representation of the design output we got for the first tank, each color

corresponds to a genotype. It is important to note that all genotypes were provided by the

national institute for agronomical research in France. They also provided an additional

genotype, called a “border genotype” to fill the gaps in case some seeds didn’t germinate.

Since its main function is to act as a back-up, this genotype wasn’t included in the design.

1. **Spatial models**

It is common to rewrite spatial models as mixed models. It useful for variance estimation but mostly because it allows the use of MM software’s and techniques. When we do that, two kinds of effects:

* Fixed: linked to environment and/or growing conditions (like sowing or tillage) and related to spatial positions in the field
* Random: Not linked to covariates and therefore harder to model. Fertility or moisture patches in the field and even genotypes effects

1. **Spatial variation**

Here you can see that when writing our model as a mixed model, there 3 main components

of the spatial variability:

* Global trends that are non-stationary, meaning they are spatially dependent and can be approximated by the row-columns effects.
* Locals trends are treated as spatially dependent random error
* Extraneous variation. These are variations that can be spatially dependent or not but are related to covariates. In our case it’s mostly going to be genotypic variability.

To model this spatial heterogeneity, there are 2 main approaches:

* Using a specific spatial variance-covariance structure
* Using smoothing techniques in 2 dimensions

The 2 models we will compare, use these 2 techniques

1. **Standard spatial model**

In this model, the global and extraneous trends are modelled using covariates.

The local trends are modelled as a two-dimensional spatially dependent error process with the random error.

They are often modelled using a variogram, but the AR(1) structure is similar and allows only 2 parameters to be estimated.

The AR structure can be elevated to a linear variance structure to be more robust to convergence issues

1. **SpATS model**

The second model uses a smooth bivariate surface to model both the local trends and the

global trends. Leaving the extraneous effects to be modelled by covariates.

This surface contains a linear part, a smoothing spline and row by column interactions.

This allows a finer depiction of the spatial grid and is suitable to highly heterogeneous environment.

1. **Goal of the analysis**

With our data analysis, we wanted to answer 4 main questions:

* Is there a significant difference between the moving and still tank, can the model capture it?
* Do the models account for spatial trends and which are the main trends?
* How different are the estimations of the two models in terms of fitted values and estimated coefficients?
* Is there a significant genotypic trend? Do the models estimate it as a genotypic effect?

To answer those, we use our weights as variables in the model.

1. **Tank difference**

Raw data show that there is a difference between tank. The moving tank seems to give higher weight values.

T-test for tank show it is significant

However paired t-test for genotypes are not significant for all, it is also because there were some genotypes with low number of replicates.

Some genotypes might be better adapted to moving/still conditions than others.

1. **Model comparison**

Raw data, fitted for the SpATS and for the BSS. Lighter colors depict higher values. Left is moving and right is still.

Both models clearly show a strong tank effect and have very similar fitted values.

The estimated spatial trends were mostly along the strips and not the position.

Higher heterogeneity in Tank A that was not picked up by the models.  
Suggests that maybe the spatial models are not suitable to moving environment.

1. **Genotypic effect**

The genotypes effect was very large in both model.  
It means it has a significant impact on the weights. It confirms the previous graph.

Both models gave vary similar estimates of the genotype effect.

1. Conclusions