Exchange event between the European phenomic community and industry

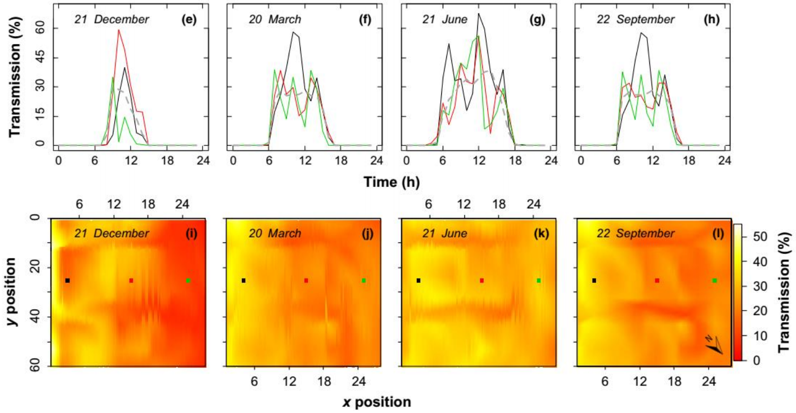
Progress on Design and analysis of phenotyping experiments across multiple platforms

Emilie Millet & Fred van Eeuwijk (WUR)

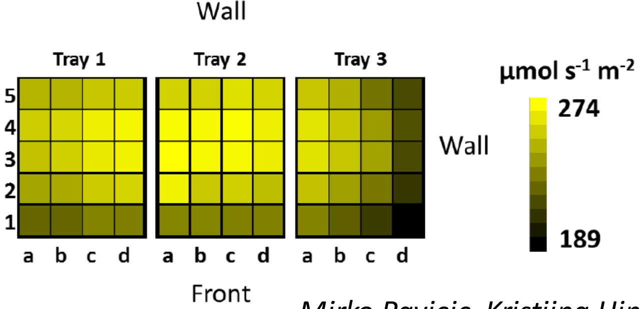
Background

Design the experiments and analyse the phenomics data with appropriate tools and methods

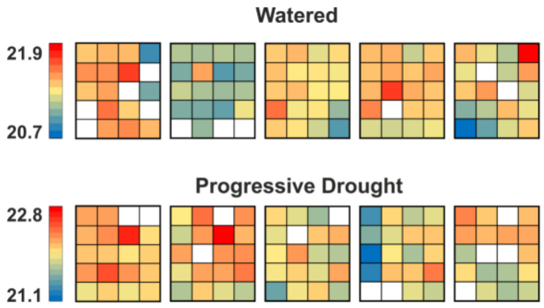
Platform = variability space and time



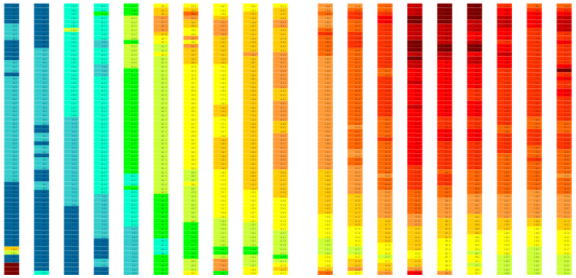
Cabrera-Bosquet et al. 2016



Mirko Pavicic, Kristiina Himanen



Finnish National Plant Phenotyping Infrastructure



Christian Jeudy, Christophe Salon

Background

Design the experiments and analyse the phenomics data with appropriate tools and methods

Platform = variability space and time
Platform = repeated measurement on thousands of units (plants, organs, etc.)

→ need to consider tools for **designing and analysing**

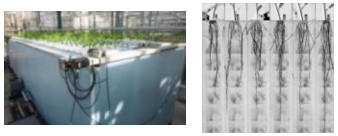
1. Include and correct for variability = increase precision of phenotype
2. Keep track of decisions = traceability (FAI**R**) + data re-use
3. Automation of the process = save time
4. New software/tools publicly available = adoption by the community

The analysis of phenomic data : new strategies

Design experiment

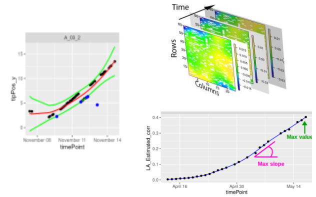
Expe map

Perform experiment & pictures analysis



Raw data

Raw data analysis



Calculated clean data

Integration of parameters/variables

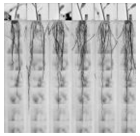
G×E, Crop model, GWAS, GP, etc.

The analysis of phenomic data : new strategies

Design experiment

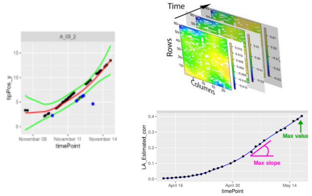
Expe map

Perform experiment & pictures analysis



Raw data

Raw data analysis



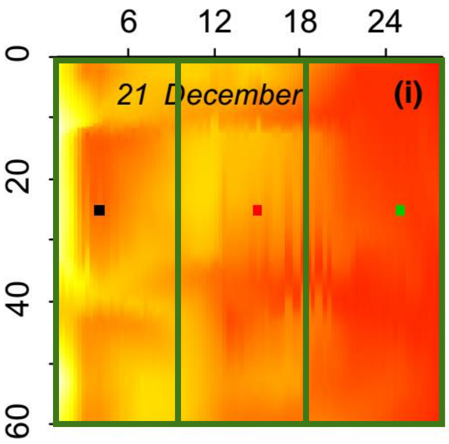
Calculated clean data

Integration of parameters/variables

G×E, Crop model, GWAS, GP, etc.

Designing phenomic experiment

The main goal of the experimental design: random allocation of treatment but not only...

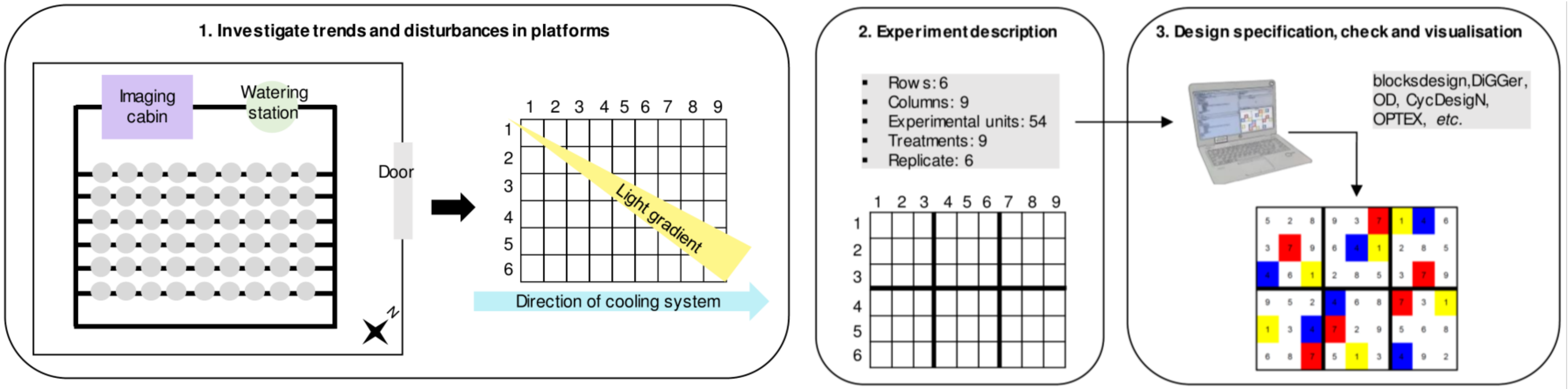


Expected environmental variability in 1 or 2 directions (rows and columns):

→ define **set of experimental units that are considered to be internally homogeneous: block.**

→ **restrict the randomization of a two-way layout** to assign the treatments to experimental units in a balanced way: **latinization**

Simple procedure to help choosing the design:



- Rows: 6
- Columns: 9
- Experimental units: 54
- Treatments: 9
- Replicate: 6

blocksdesign, DiGGer, OD, CycDesigN, OPTEX, *etc.*

5	2	8	9	3	7	1	4	6
3	7	9	6	4	1	2	8	5
4	6	1	2	8	5	3	7	9
9	5	2	4	6	8	7	3	1
1	3	4	7	2	9	5	6	8
6	8	7	5	1	3	4	9	2

Designing phenomic experiment

A user-friendly interface of a design generator has been developed (web app)

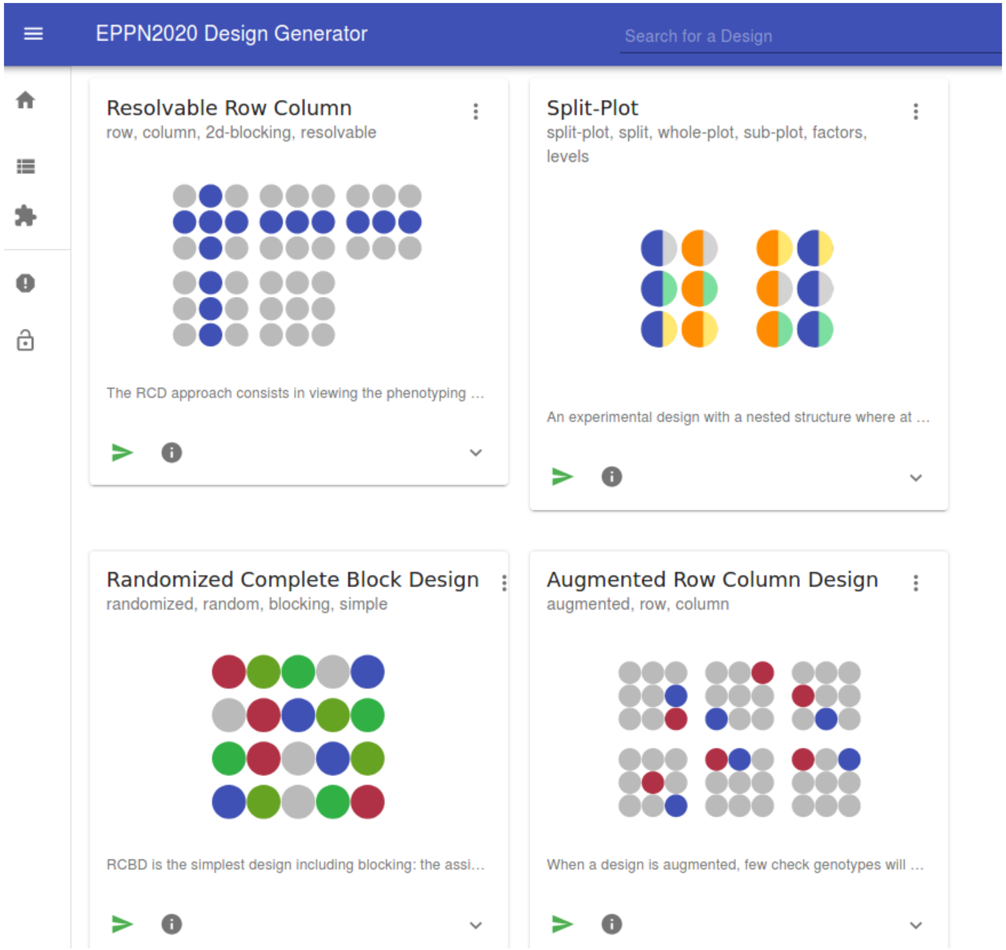
Robert Horne,
Darren Murray

Available designs:

- Resolvable Row-Column (augmented)
- Randomized Complete Block
- Augmented Row-Col
- Split-Plot
- Partially replicated (p-rep)

Create a «facility» based on the layout of his own platform.

⇒ the user specify the design details/layout (number of genotypes, the number of replicates, etc.)
⇒ gives warnings when mis-specified
⇒ allows visualisation of the 2D map of the experiment



Resolvable Row Column
row, column, 2d-blocking, resolvable
The RCD approach consists in viewing the phenotyping …

Split-Plot
split-plot, split, whole-plot, sub-plot, factors, levels
An experimental design with a nested structure where at …

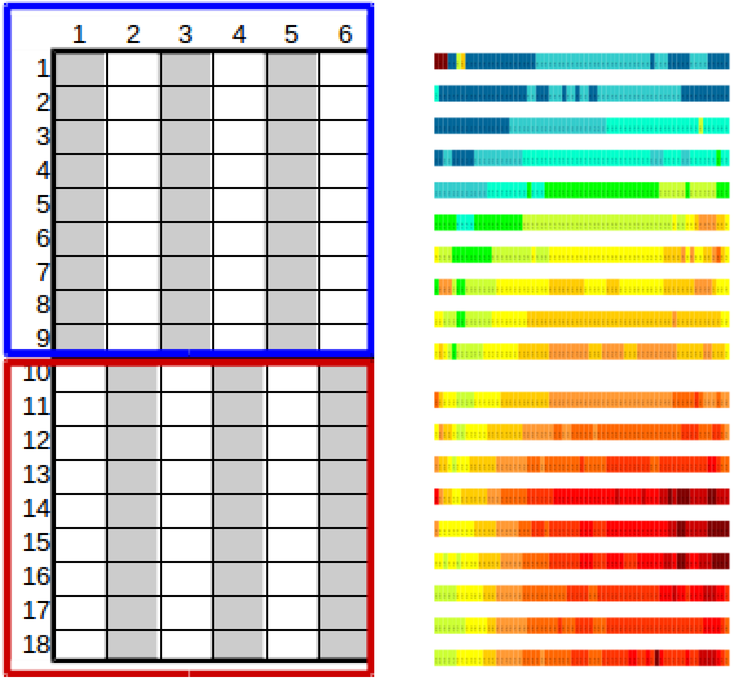
Randomized Complete Block Design
randomized, random, blocking, simple
RCBD is the simplest design including blocking: the assi…

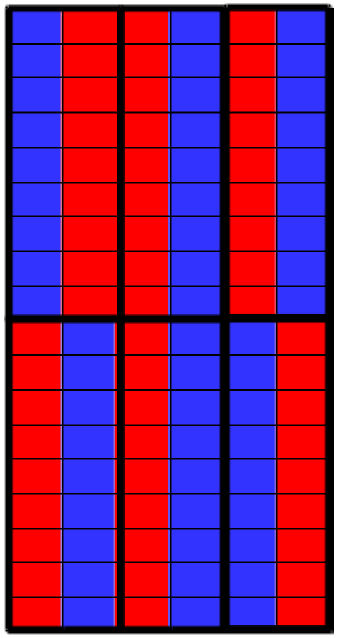
Augmented Row Column Design
augmented, row, column
When a design is augmented, few check genotypes will …

Designing phenomic experiment

9 genotypes x 2 water level x 6 repetitions

→ **How to design this?**





Split the treatment in two part of the platform and randomize the genotypes
-> What if there is environmental variability?
-> Cannot assess the water effect

Split into main block (water) and subplot (genotypes) and latinized the genotypes
-> What software?

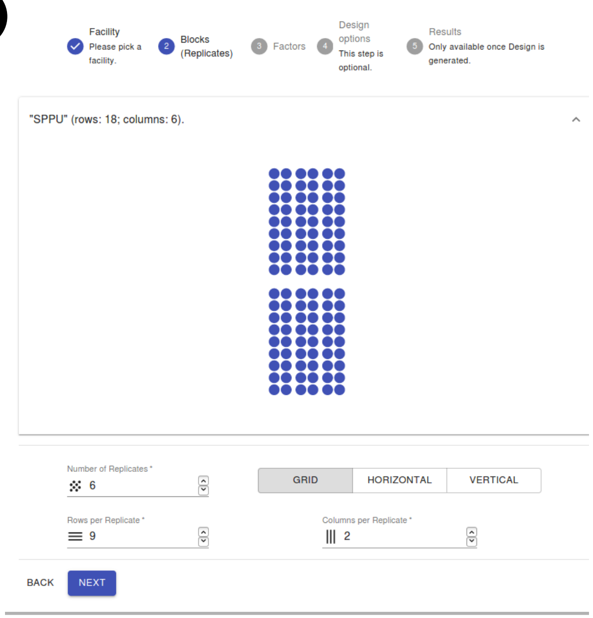
Designing phenomic experiment

The design generator web app

VSNi

Robert Horne,
Darren Murray

❶



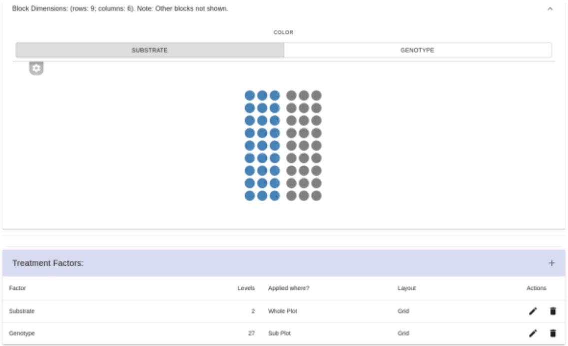
"SPPU" (rows: 18; columns: 6).

Number of Replicates * 6

Rows per Replicate * 9

Columns per Replicate * 2

❷

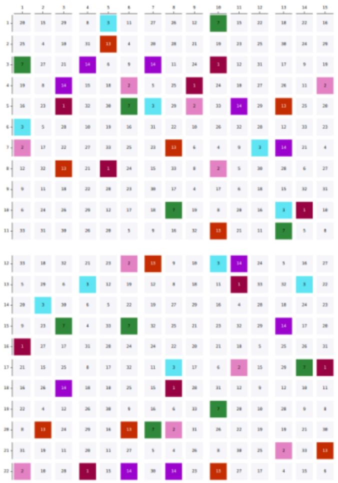


Block Dimensions: (rows: 9; columns: 6). Note: Other blocks not shown.

Treatment Factors:

Factor	Levels	Applied where?	Layout	Actions
Substrate	2	Whole Plot	Grid	✎ 🗑
Genotype	27	Sub Plot	Grid	✎ 🗑

❸



The analysis of phenomic data : new strategies

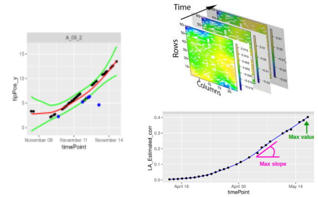
Design experiment

Expe map

Perform experiment & pictures analysis

Raw data

Raw data analysis



Calculated clean data

Integration of parameters/variables

G×E, Crop model, GWAS, GP, etc.

The analysis of phenomic data : new strategies

In classical field trial = multiple environments + one time measurement (yield)

1. correction for spatial variability per environment
2. combining environments

= stage-wise analysis

Or 1 + 2 at once

=> Two methods can be equivalent if weights (error) are carried on from stage 1 to stage 2

In phenomics = multiple experimental series + multiple time measurement (+ multiple scales)

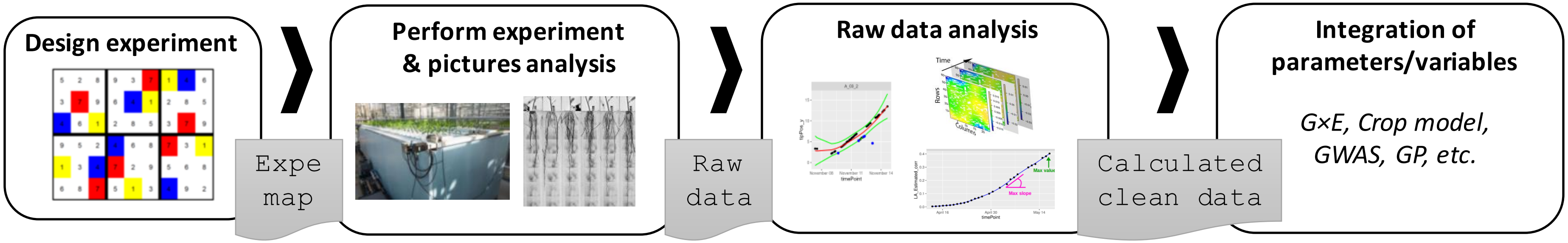
→ **How to model spatial and temporal trends in high-dimensional data?**

- Time + Space in 3D models?
- Longitudinal modelling → spatial correction?
- Spatial correction → time modelling?
- How to detect outliers?

Advantages of stage-wise:

- making diagnostic of individual experiments
- clean individual experiment data
- sometimes required to obtained starting values of one-stage analysis

The analysis of phenomic data : new strategies



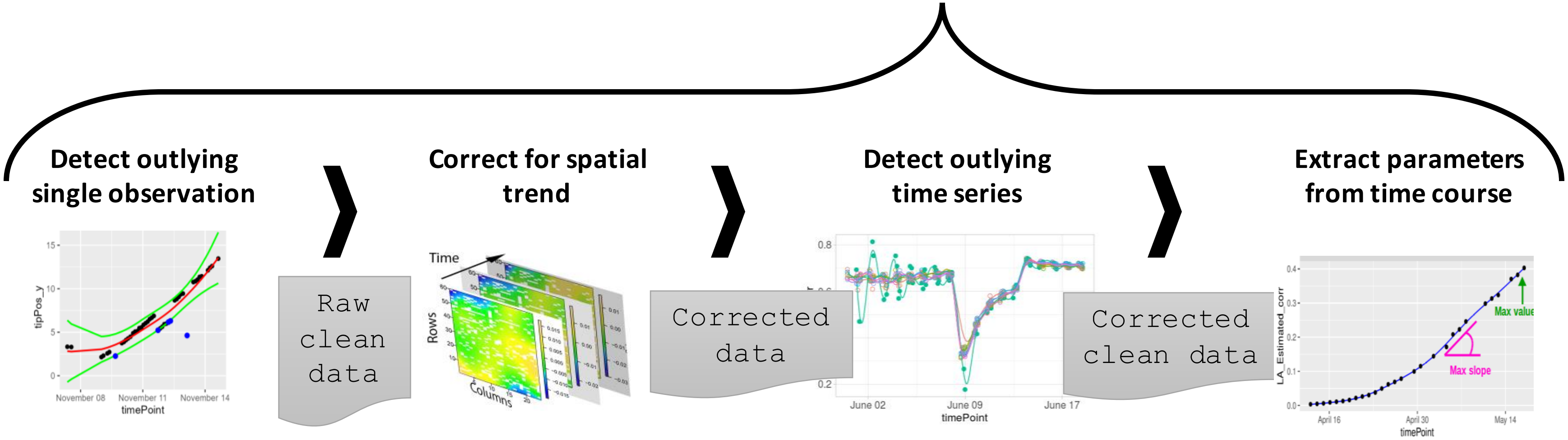
Design experiment ❯ Expe map

Perform experiment & pictures analysis ❯ Raw data

Raw data analysis ❯ Calculated clean data

Integration of parameters/variables

G×E, Crop model, GWAS, GP, etc.



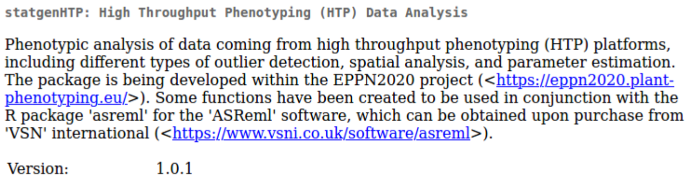
Detect outlying single observation ❯ Raw clean data

Correct for spatial trend ❯ Corrected data

Detect outlying time series ❯ Corrected clean data

Extract parameters from time course

R pakage: statgenHTP



`statgenHTP: High Throughput Phenotyping (HTP) Data Analysis`

Phenotypic analysis of data coming from high throughput phenotyping (HTP) platforms, including different types of outlier detection, spatial analysis, and parameter estimation. The package is being developed within the EPPN2020 project (<https://eppn2020.plant-phenotyping.eu/>). Some functions have been created to be used in conjunction with the R package 'asreml' for the 'ASReml' software, which can be obtained upon purchase from 'VSN' international (<https://www.vsni.co.uk/software/asreml>).

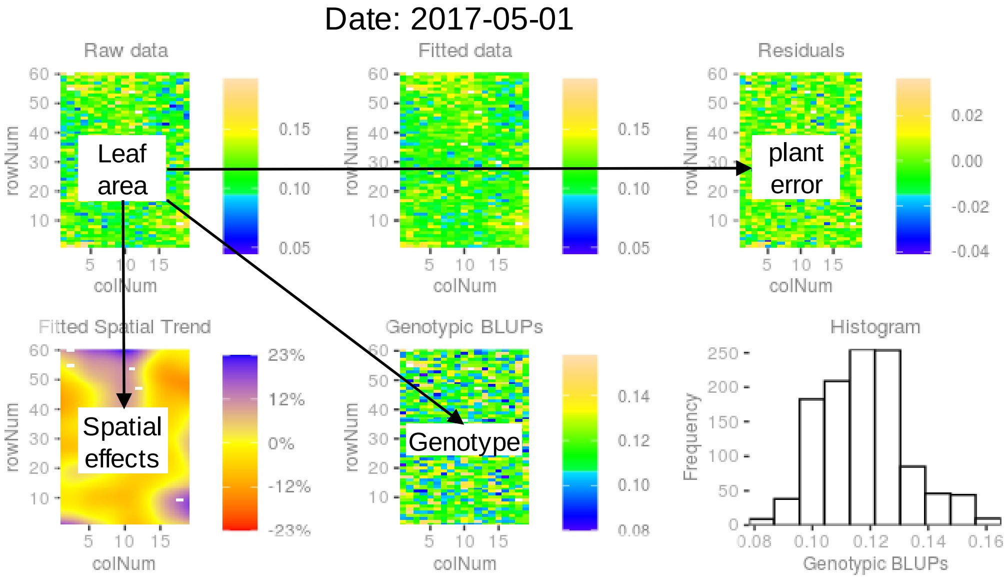
Version:	1.0.1

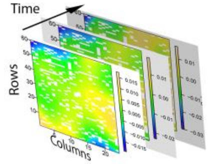
Bart-Jan van Rossum
Martin Boer

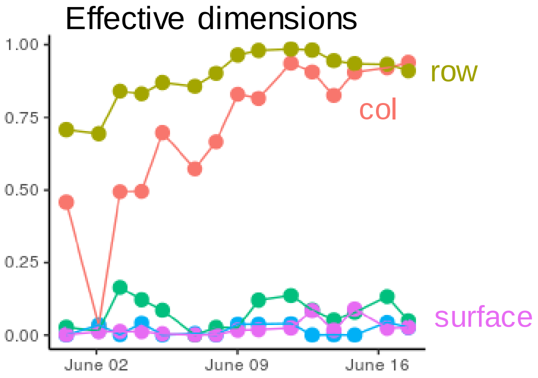
Nadine Hilgert
Isabelle Sanchez

Correction of spatial trends and design factors

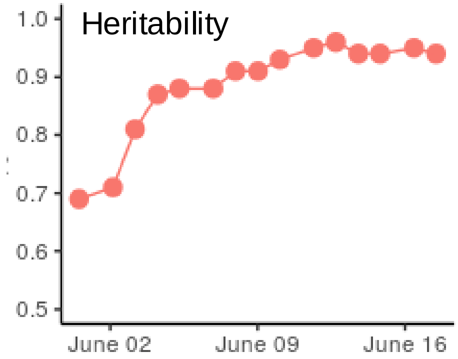
Accurately separate the genetic (treatment) effects from the spatial effects at each time point using flexible 2-dimensional P-spline surfaces (SpATS in R):







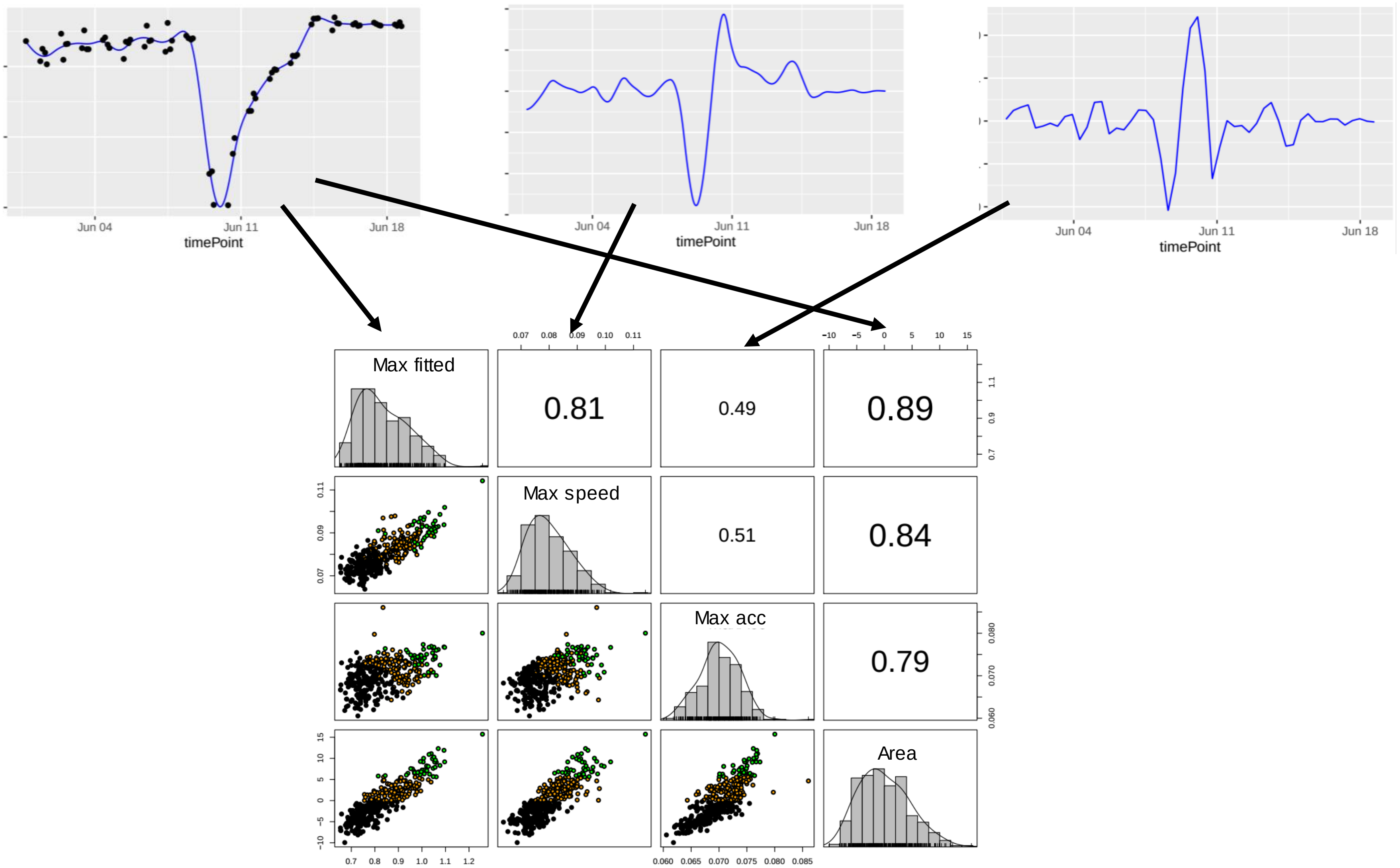
(1) Correction following the partition by Welham et al. 2004
→ data at the **experimental/observational unit level**
or
(2) Prediction
→ data at the **treatment (genotype) level**



Diana Perez
Maria Xose Rodriguez Alvarez

Estimation of time-independent variables

Fit splines on corrected data:

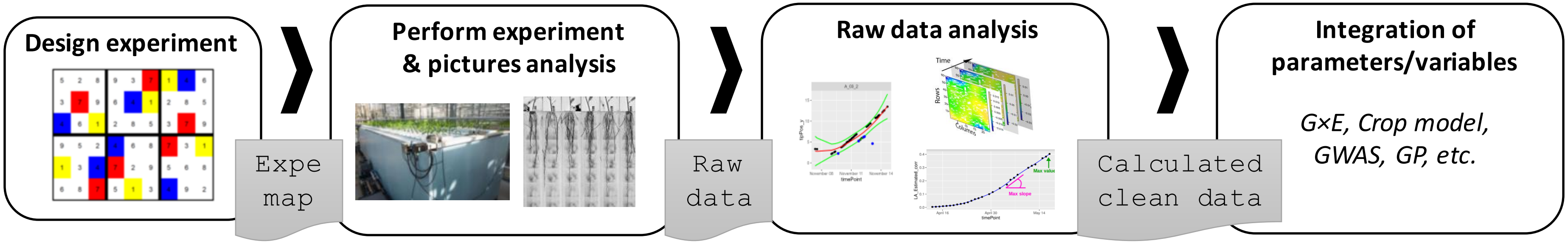


fitted values

1st derivatives = speed
e.g. growth rate

2nd derivatives = acceleration

The analysis of phenomic data : new strategies



Design experiment

Expe map

Perform experiment & pictures analysis

Raw data

Raw data analysis

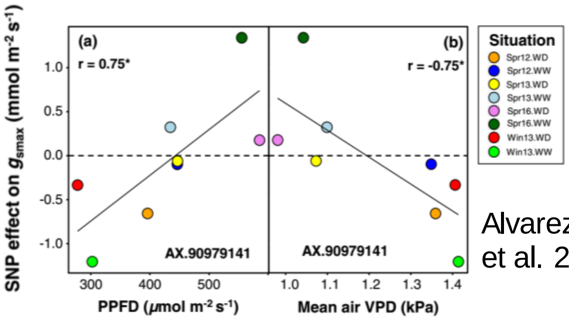
Calculated clean data

Integration of parameters/variables

G×E, Crop model, GWAS, GP, etc.

Integration of multiple experiments/scales

Multiple experiments = G×E and QTL×E in platform too
→ allelic effect vs. environmental cov
→ multi-trait QTL

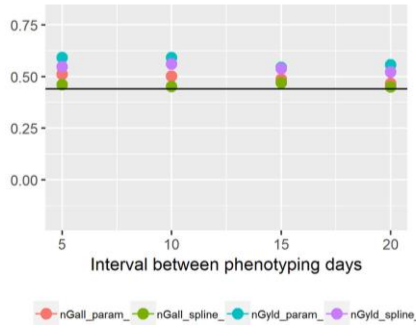


Alvarez Prado et al. 2017

Using phenomic traits helps better predict the target trait (yield)

	Calibration	Test
Target trait	Simulated with CGM	To predict
Secondary trait	Simulated with CGM	Pred with CGM

Bustos-Korts et al. 2019



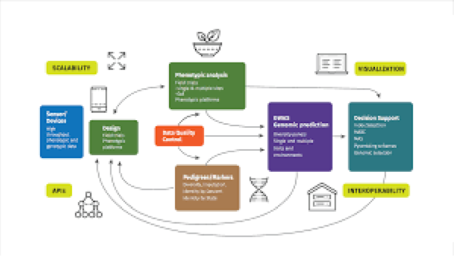
	Calibration	Test	Test	Test
Target trait	Observed	To predict	To predict	To predict
Secondary trait	Observed	Obs. or not	Pred with CGM	Pred with GP
		Velazco et al. 2019	Robert et al. 2020	Arouisse et al. *under review*

redrawn from Robert et al. 2020

Summary

Development of new technology = big data in phenomics
→ Need to adapt existing methods and to streamline the analysis/cleaning process

- Experimental design also in platform → Design generator + procedures
- Strategy for data analysis in multiple stages → statgenHTP R package
- Part of a pipeline of analysis with multiple modules



Collaborative work, multi-disciplines

- Dialog with platform managers/users
- Platform user decision/expertise
- Link with IS and data re-use
- EMPHASIS standards

Ackowledgement

WAGENINGEN UNIVERSITY & RESEARCH

F van Eeuwijk
W Kruijer
M Boer
B van Rossum
M Aarts
R Boesten

VSNi

R Horne
D Murray

JÜLICH Forschungszentrum

F Fiorani
O Muller
H Porter

bcam basque center for applied mathematics

C Rodriguez
D Perez

Mistea

I Sanchez
N Hilgert

PRIFYSGOL ABERYSTWYTH UNIVERSITY

J Doonan
G Garzon Martinez

SLOVAK UNIVERSITY OF AGRICULTURE NITRA

K Olšovská
M Kovár

lepse Montpellier

F Tardieu
C Welcker
L Cabrera-Bosquet

UCLouvain

X Draye
S Passot

University of Nottingham UK | CHINA | MALAYSIA

T Pridmore
D Wells
C Sturrock

GDEC

R Rincent
V Allard
J Legouis

VIB

S Dhondt
V Storme

HELSINGIN YLIOPISTO HELSINGFORS UNIVERSITET UNIVERSITY OF HELSINKI

K Himanen
M Pavicic