Sensitivity analysis of dynamic crop models to assist crop science: assessing the impact of multiple traits on yield in Australian wheat

Pierre Casadebaig, Robert Faivre, Karine Chenu

Sensitivity analysis of dynamic crop models to assist crop science: assessing the impact of multiple traits on yield in Australian wheat.

Pierre Casadebaig (1), Robert Faivre (2), Karine Chenu (3)

1. "Agroecologies, Innovation, Ruralités", INRA, Toulouse
2. "Mathématiques et Informatique Appliquées de Toulouse", INRA, Toulouse
3. "Queensland Alliance for Agriculture and Food Innovation", University of Queensland, Australia

Systems approach to crop improvement [1, 2]

Taking advantage of genotypes x environment x management (GEM) interactions

Canopy viewed as a system of cultivar (G), pedo-climatic factors (E) and management practices (M).

Problem: improving crop yield with uncertain climate

- agronomy: improve management practices (choose G+M|E)
- **genetics**: improve genetic material (change G|E+M)

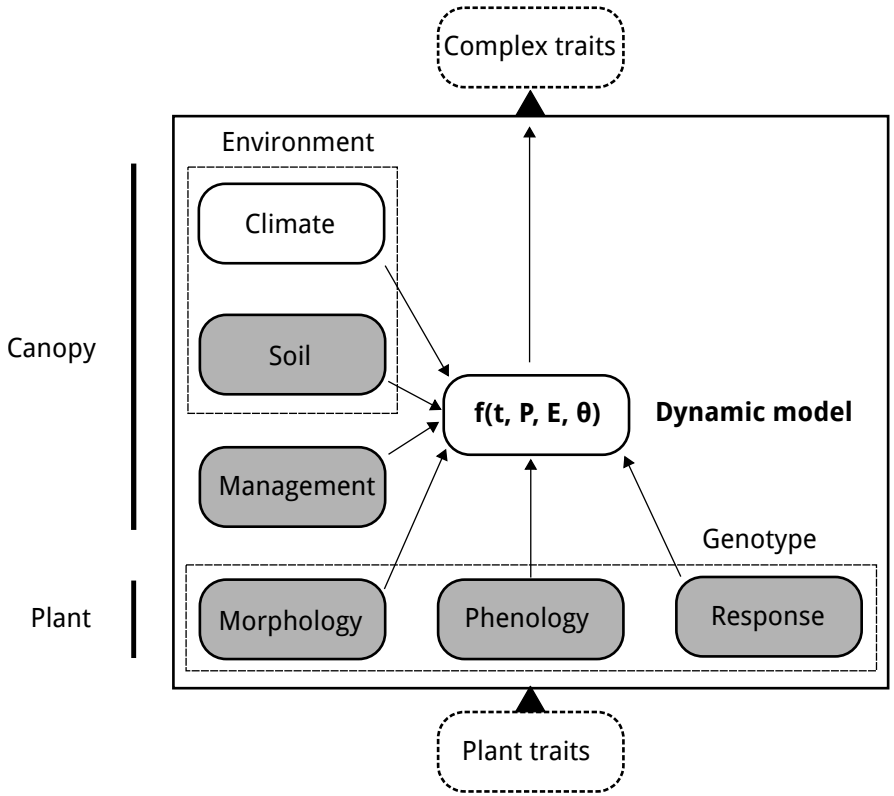
Tools

- experiments: hard to sample climatic variability
- models: genotypic determinism in crop models

Systems approach to crop improvement

A methodology for screening crop model parameters usable as plant traits for breeding

1. Model: APSIM platform [3]
   - dissect complex traits
2. Explore
   - **screen**: sensitivity analysis
     - factors and distribution
     - numerical design
     - simulation and index computation
   - **search**: variance analysis
     - parameter × environment
3. Optimization
   - parameter combination

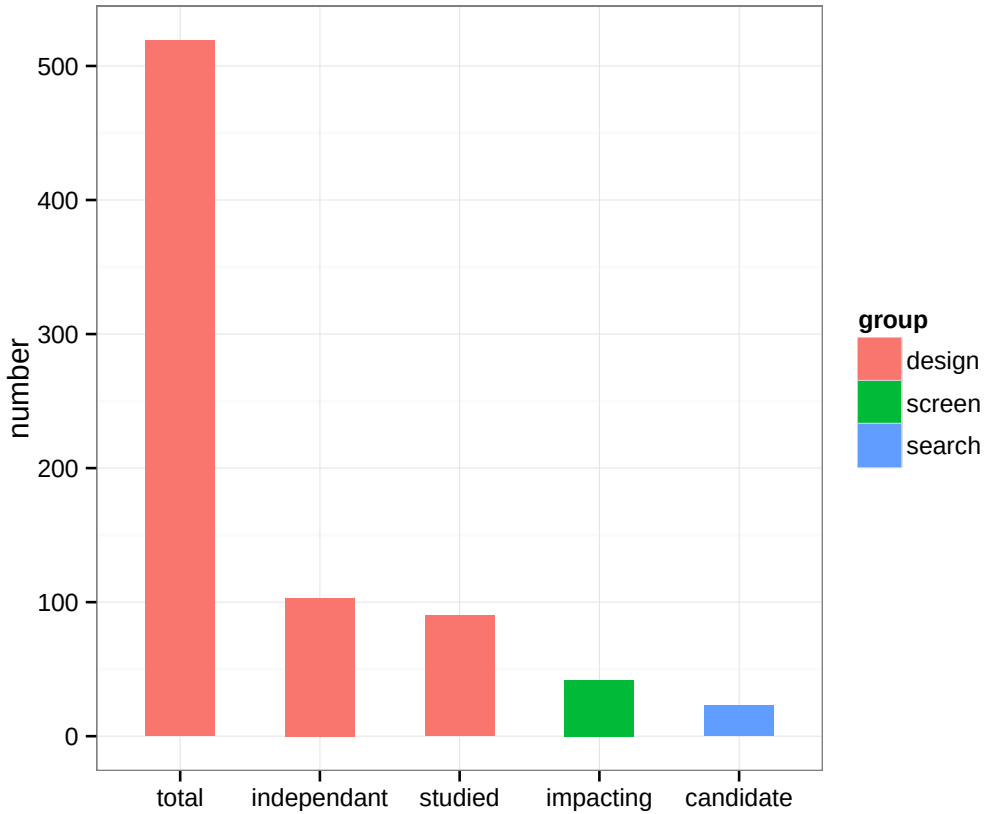


Defining input factors

How many parameters in a typical crop model ?

APSIM-wheat model

- $\sim$ 500 parameters
- plant-related parameters
- 103 independant
  - 62 values
  - 41 functions
- 90 studied after grouping

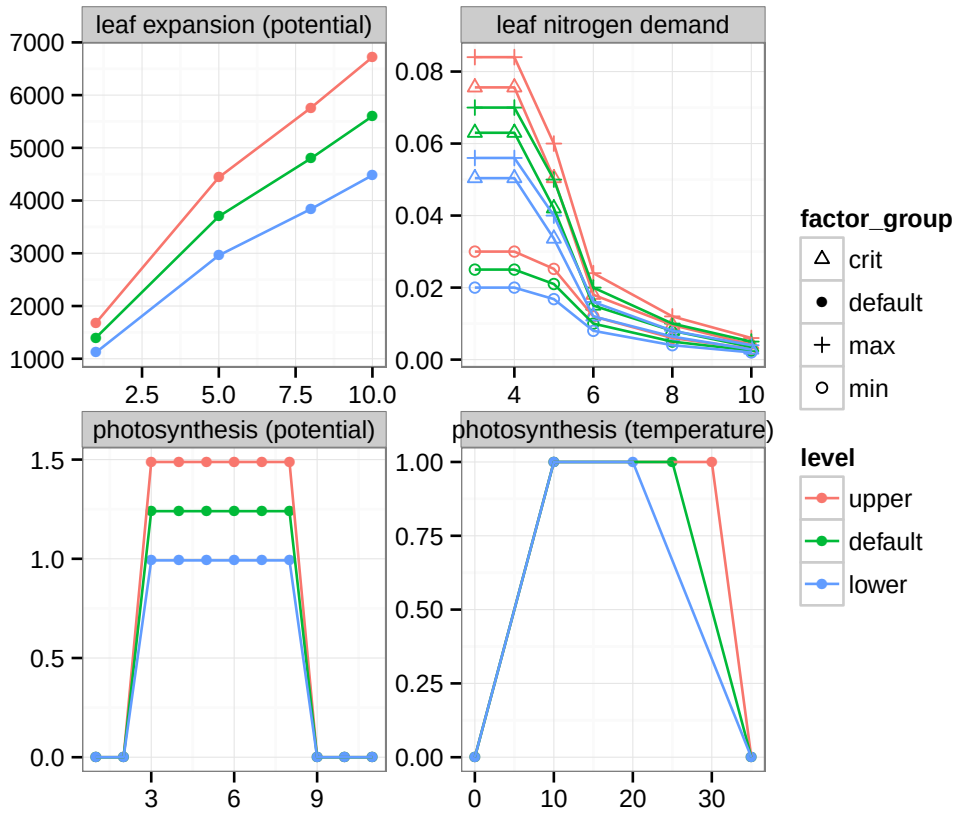


Setting the variation range

How to keep biological meaning without going into 90 special cases ?

Consensus

- 40% variation range
- 3 rules:
  1. scale single value
  2. scale y vector
  3. scale point in x or y vector

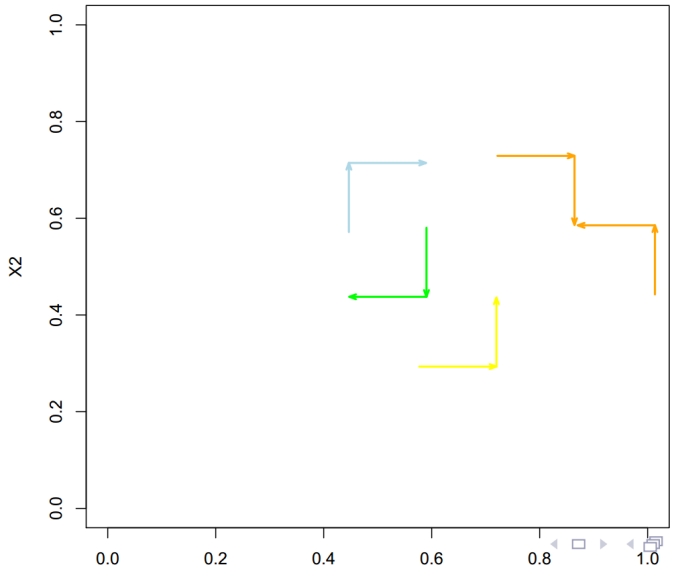


Sampling method for parameter and environmental space

How to sample GEM landscape for wheat in Australia ?

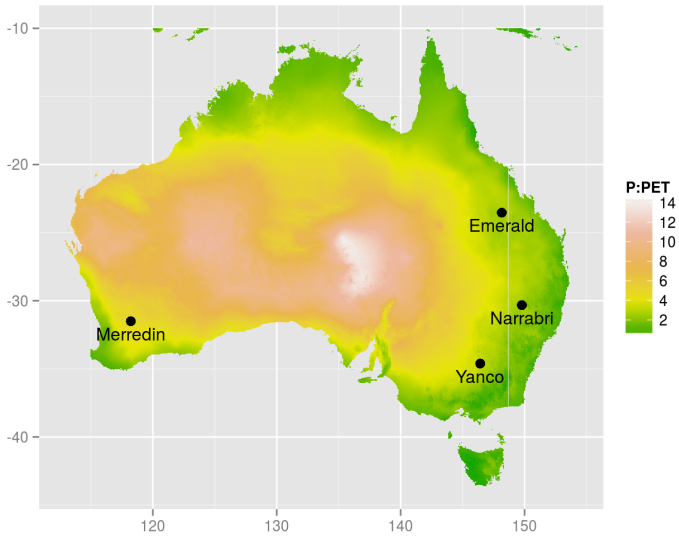
Genotypes (Parameters) $n = 9100$

- Morris method [4] (*sensitivity*)
- 6 levels by factor
- 100 random rep. of OAT designs



Environments $n = 9000$

- 2 CO_2 levels
- 3 sowing dates
- 3 N fertilization amount
- 4 locations
- 125 years of climatic data



Simulation and sensitivity index computation

Simulation and output variables

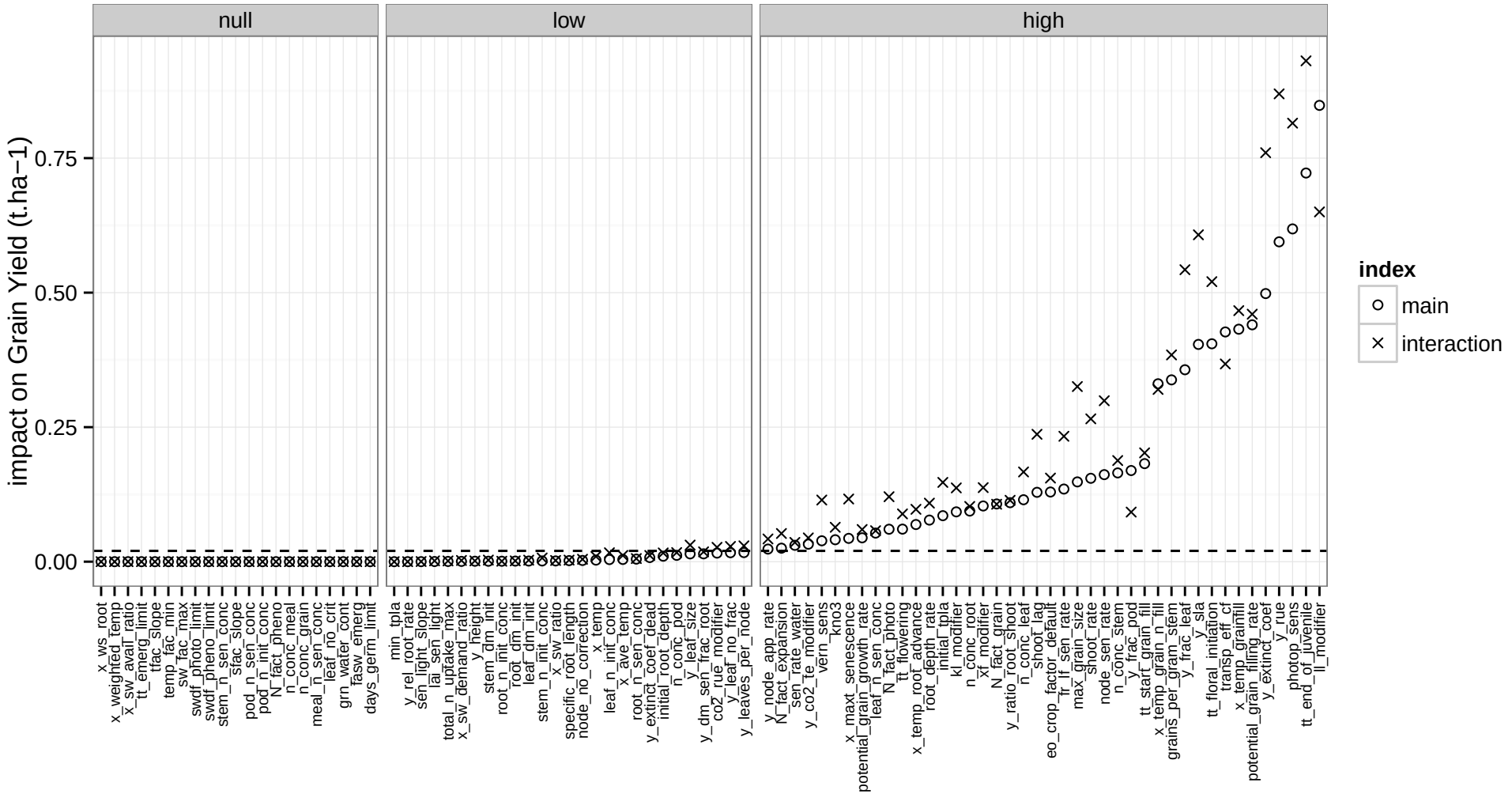
- 8 output variables
- distributed computing in CSIRO (81.9×10^6 simulations)
- R packages *ncdf4*, *dplyr*

Sensitivity indexes

- main (μ^*): estimation of the linear effects of inputs
- interaction (σ): estimation of the non-linear/interaction effects

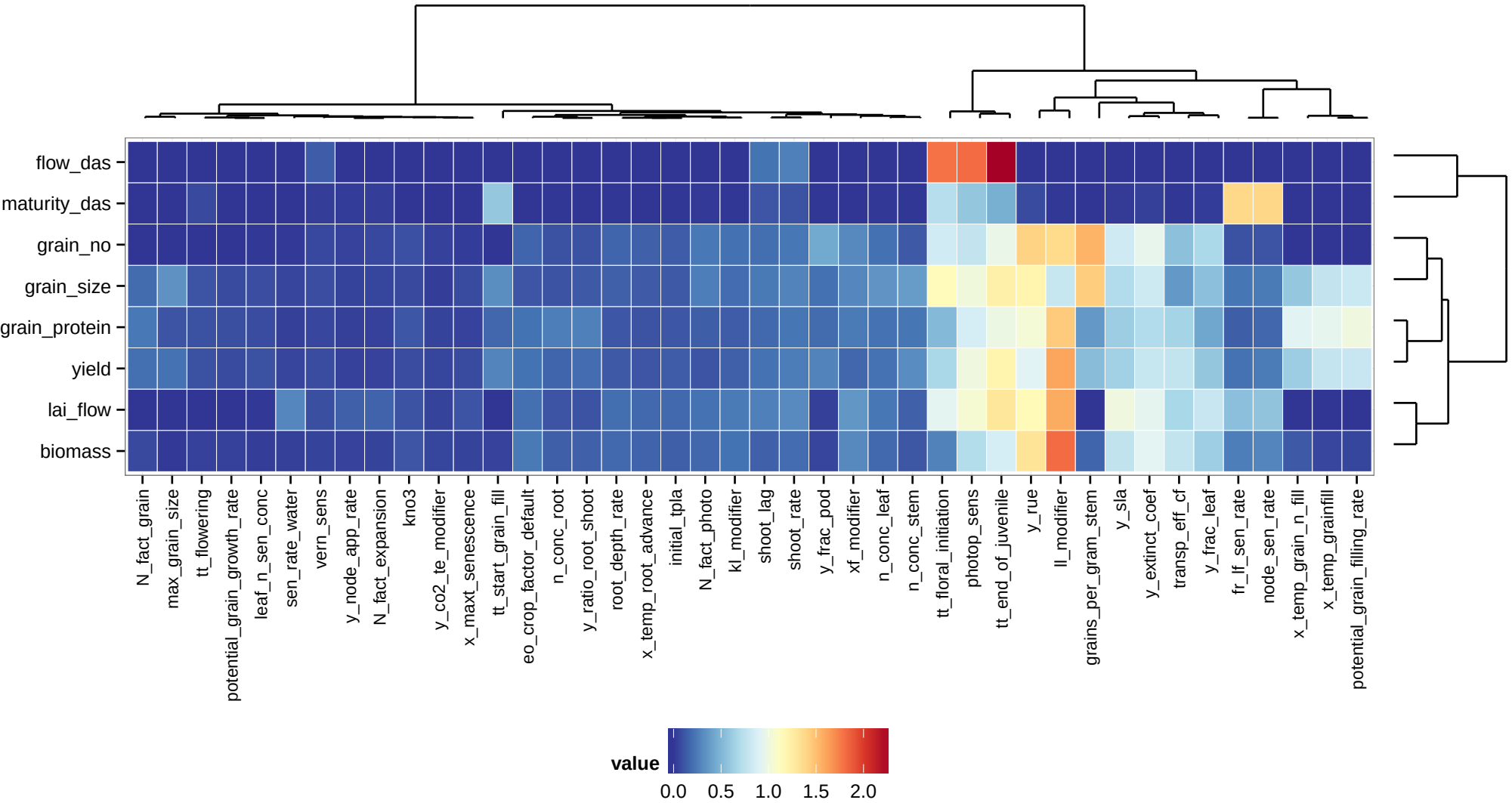
Screening for impactful parameters

About a half of the parameters are not or weakly impacting yield in control conditions



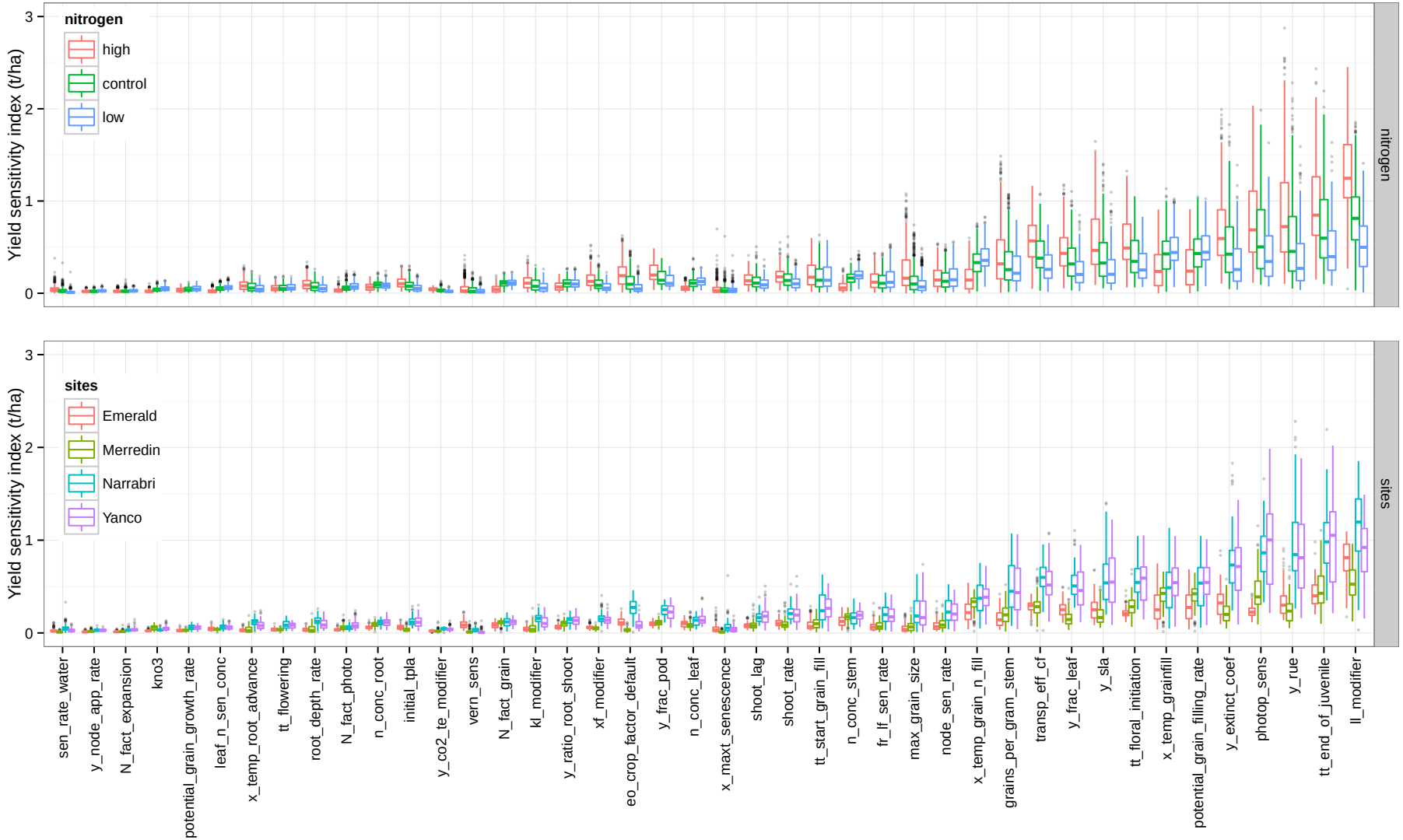
Screening for impactful parameters

Parameters mainly impact specific output variables in control conditions



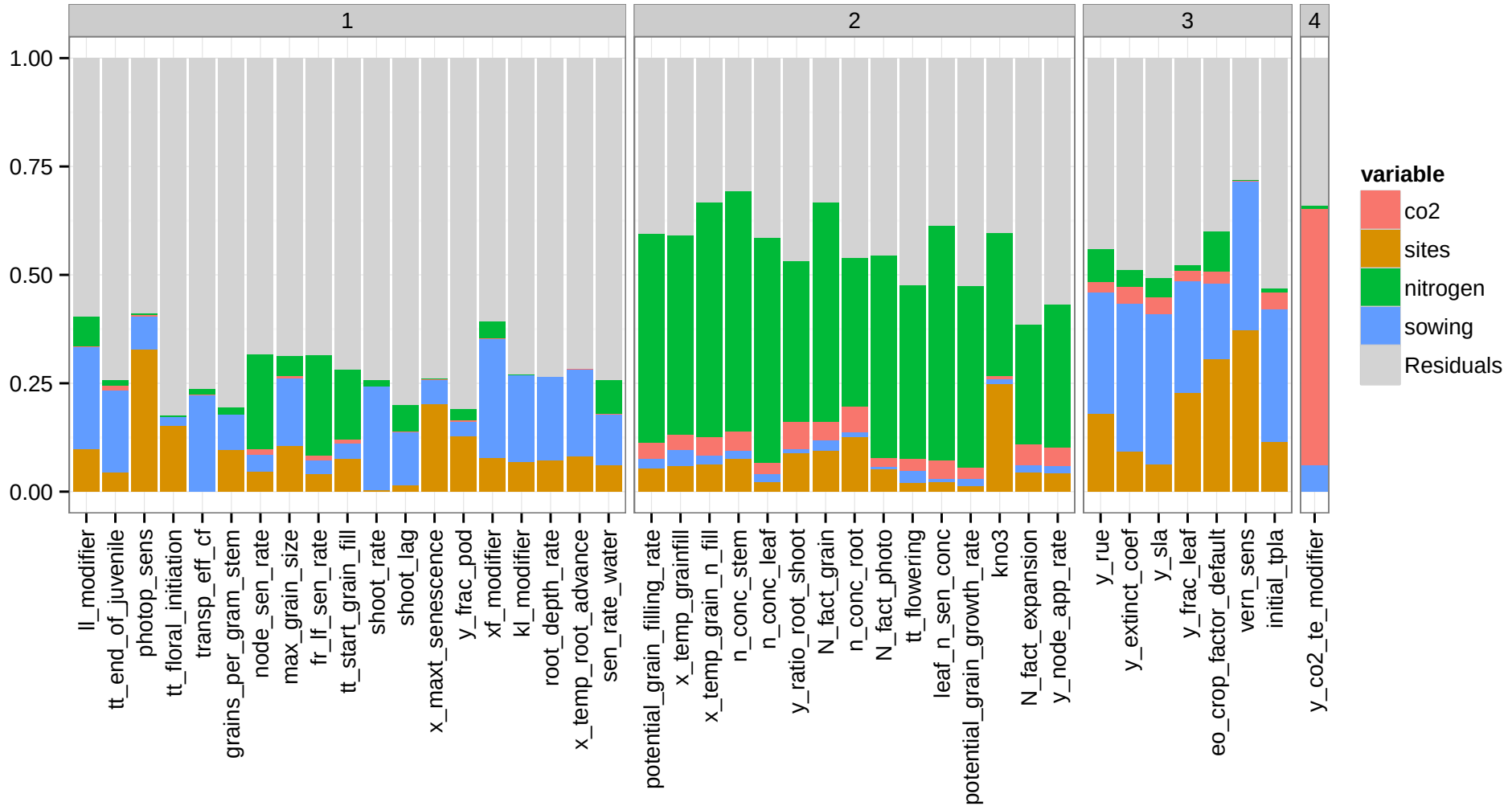
Searching for candidate parameters

The impact of traits is strongly affected by environment and management



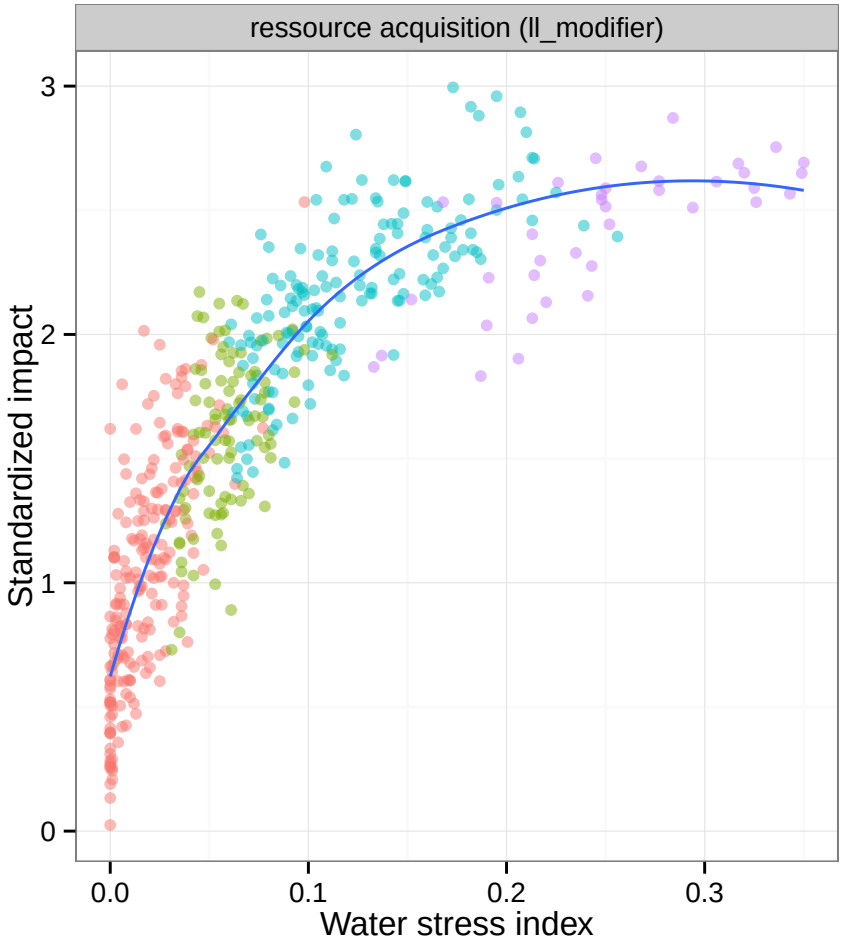
Searching for candidate parameters

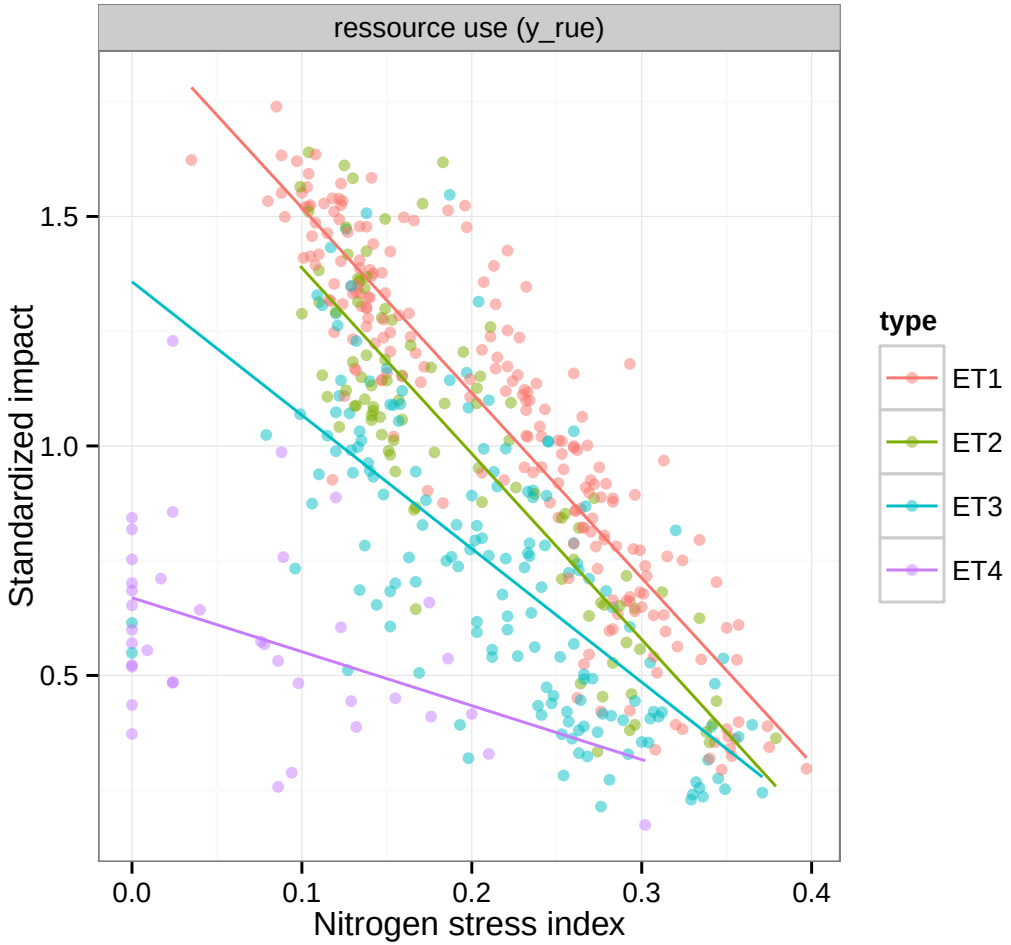
A subset of the most impactful parameters is less dependant to environmental effect



Searching for candidate parameters

The impact of parameters is linked to ressource availability





Conclusions

Numerical exploration

- low-impact parameters are targets for code refactoring
- methods to integrate G×E interactions in breeding approaches

UseR !

- numerical design, decoupling, index computation: *sensitivity* [5]
- data maniplation and visualization: *reshape2* [6], *dplyr* [7], *ggplot2* [8]
- reproducible code (markdown, *knitr* [9], github repository)

References

[1] Graeme Hammer, Mark Cooper, Francois Tardieu, Stephen Welch, Bruce Walsh, Fred van Eeuwijk, Scott Chapman, and Dean Podlich. Models for navigating biological complexity in breeding improved crop plants. *Trends in Plant Science*, 11(12):587–593, December 2006.

[2] Xinyou Yin, Paul C. Struik, and Martin J. Kropff. Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Science*, 9(9):426–432, September 2004.

[3] B. A. Keating, P. S. Carberry, G. L. Hammer, M. E. Probert, M. J. Robertson, D. Holzworth, N. I. Huth, J. N. G. Hargreaves, H. Meinke, and Z. Hochman. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy*, 18(3-4):267–288, 2003.

[4] Max D Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2): 161–174, 1991.

[5] Gilles Pujol, Bertrand Iooss, and Alexandre Janon. *sensitivity: Sensitivity Analysis*, 2014. URL `http://CRAN.R-project.org/package=sensitivity`. R package version 1.8-2.

[6] Hadley Wickham. Reshaping data with the reshape package. *Journal of Statistical Software*, 21(12):1–20, 2007. URL `http://www.jstatsoft.org/v21/i12/`.

[7] Hadley Wickham and Romain Francois. *dplyr: a grammar of data manipulation*, 2014. URL `http://CRAN.R-project.org/package=dplyr`. R package version 0.2.

[8] Hadley Wickham. *ggplot2: elegant graphics for data analysis*. Springer New York, 2009. ISBN 978-0-387-98140-6. URL `http://had.co.nz/ggplot2/book`.

[9] Yihui Xie. knitr: A comprehensive tool for reproducible research in R. In Victoria Stodden, Friedrich Leisch, and Roger D. Peng, editors, *Implementing Reproducible Computational Research*. Chapman and Hall/CRC, 2013. URL `http://www.crcpress.com/product/isbn/9781466561595`. ISBN 978-1466561595.