

The use of R as part of a large-scale information management and decision system

Joris DE WOLF, Koen BRUYNSEELS, Rindert PEERBOLTE and Willem BROEKAERT
CropDesign, Belgium

Why did *CropDesign* choose **R** as the statistical environment for its *high-throughput plant evaluation platform TraitMill* ?

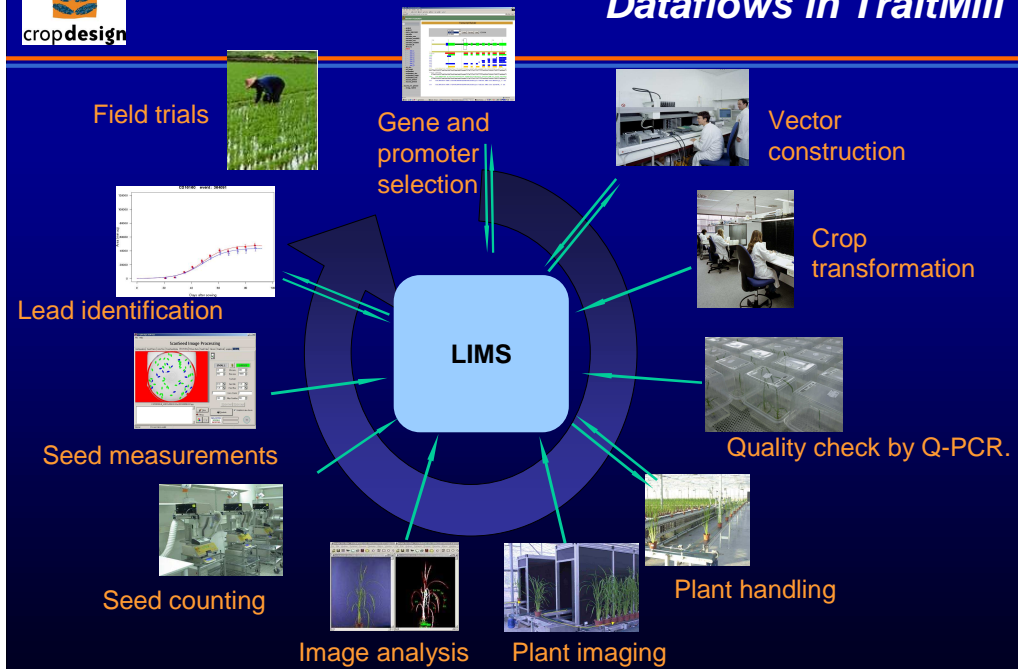
We give our experience since 2001 with R as part of an industrial-scale data management and decision support system.

TraitMill: Greenhouse



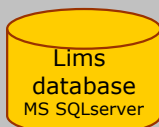
TraitMill: Plant imaging





Operations

Measuring equipment
Image analysis software
Transponder readers
Interfaces



JDBC

Java Lims server

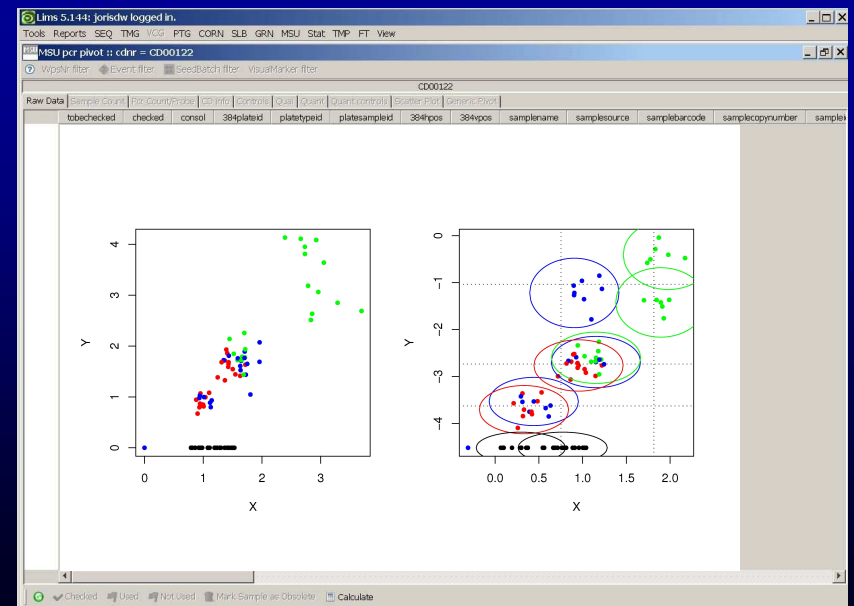
JRclient

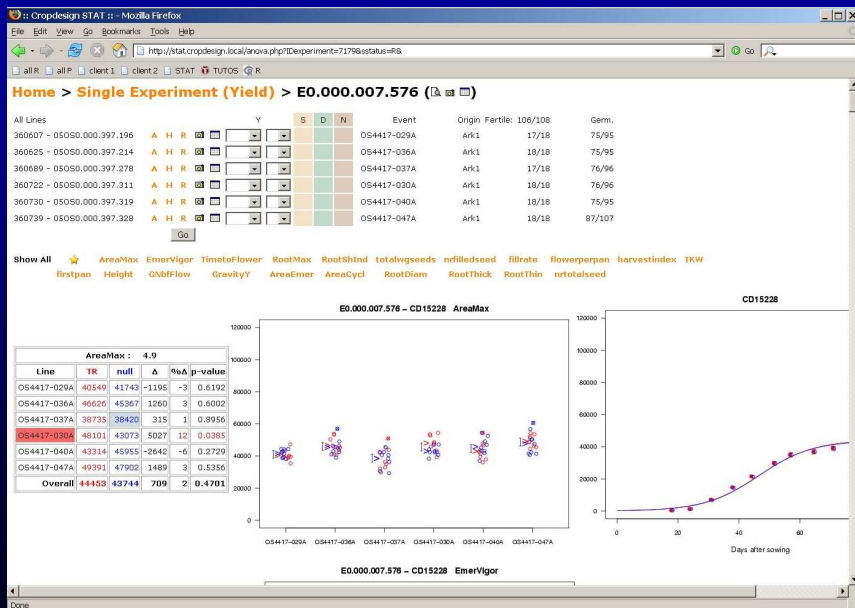
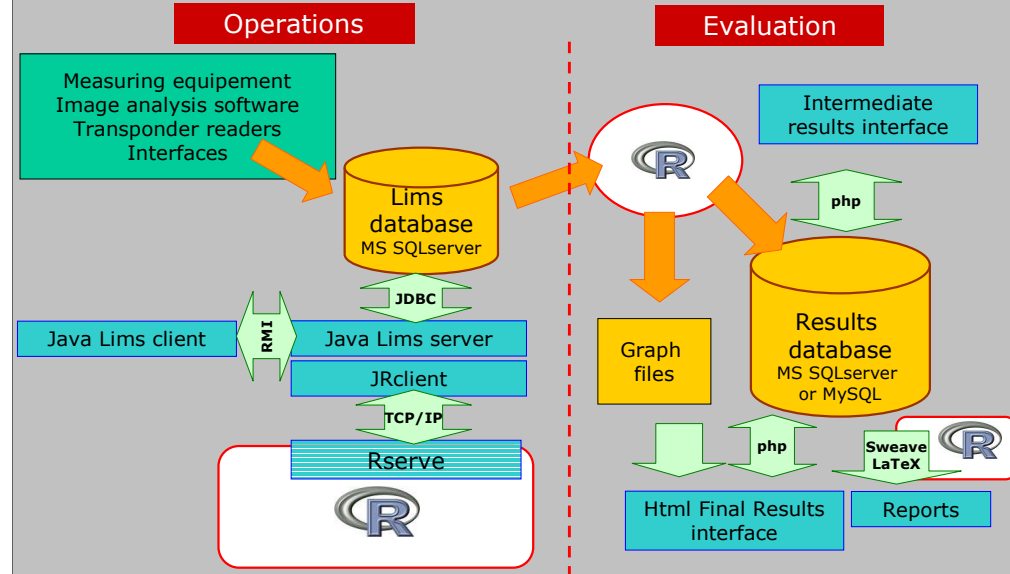
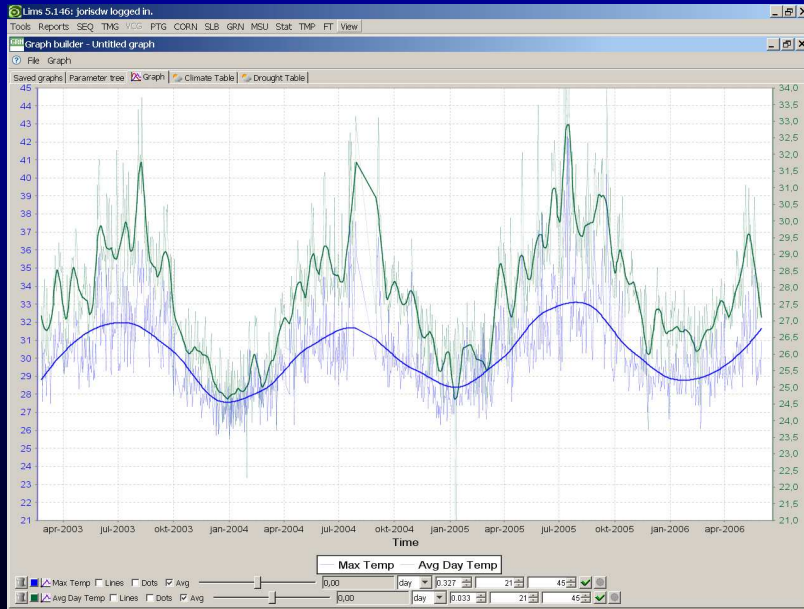
TCP/IP

Rserve



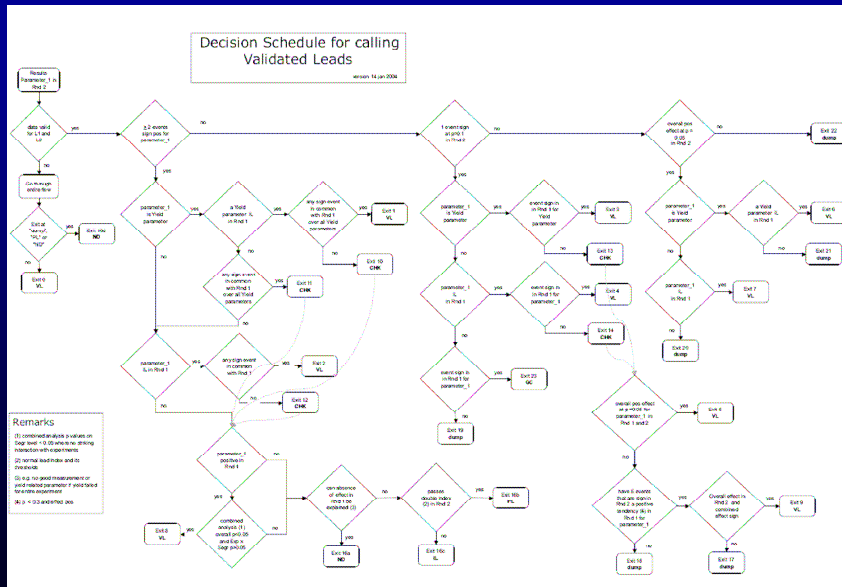
Java Lims client





- General linear models
- Mixed linear models
- Generalised linear models
- Survival analysis
- Clustering
- Time series
- Graphical visualisation

- base
- stats(4)
- graphics
- MASS
- nlme
- splines
- survival
- mvtnorm
- ellipse
- gtools
- cluster
- sweave
- ...



- Database connection
- Mimimize/optimize interactivity (400 dataset/week)
- Data preparation
- Data consistency checking (missing & odd values)
- Robustness of the methods
- Error capture
- Fall-back systems
- Standardize output
- Graphical visualization

- ☺ Easy to integrate into the architecture of databases and interfaces
- ☺ Various options of using the R engine
- ☺ State-of-the-art statistical techniques
- ☺ High quality and versatile graphical facilities
- ☺ Programming language
- ☺ Free, no problems with licenses
- ☺ Available for Linux

- ☹ No interface for MS SQL server

- ☺ Relatively easy programming language
- ☺ Algorithms and output can be easily stripped down and adapted to the needs
- ☺ Quick assistance through a helpful user community
- ☺ Fast and stable

- ☹ Biometricians need programming skills
- ☹ Input and output of functions not standardized
- ☹ Frequent updates not always compatible with old scripts
- ☹ Difficult to find help on old versions

R is a very valuable resource for this sort of application.

After 5 year of experience, we still think we made the right choice.