

Spatial analysis

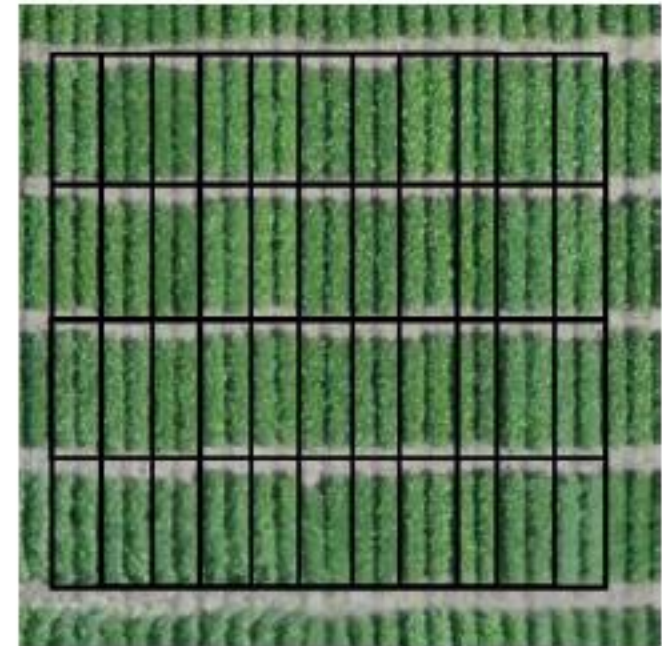
AX 04/17/2017

<http://alenxav.wix.com/home>

Why

- Need for accurately control field heteroscedasticity in heterogeneous environments
- Better use of information

Space



When

- Some traits are more sensitive than others (genetic architecture)
- Small plots
- Few replications or non-replicated trials
- Unknown sources of soil variance (fertility, fungus or nematodes)
- Unobvious way of blocking

How

- Spatial covariate
 - Spatial covariance
 - Trait as % nearest check
 - NN as intercept (FW)
-
- Local control does not override the other two principles of exp. design
 - Randomization
 - Replication

Case of study: Lado et al. 2013

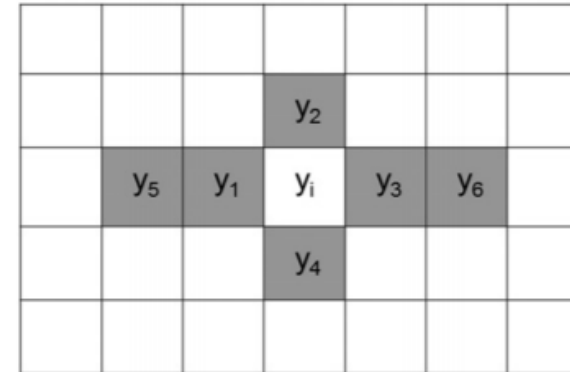
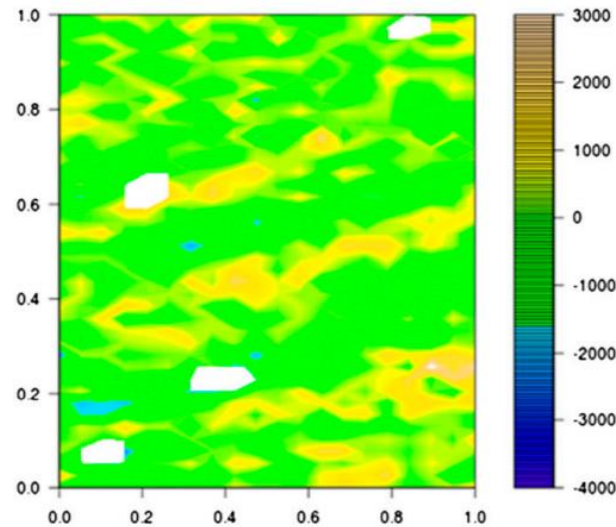


Figure 1 Diagram to calculate the covariable x_i . y_i is the phenotypic value in the plot. The neighboring plots are indicated with gray color.

Table 3 Accuracy of predictions for each trial in 2011 using random training sets with 100 independent randomizations

			IB	RC	RCB_MVNG	MVNG
SR_FI	GY	RR	0.298 ± 0.117	0.296 ± 0.119	0.319 ± 0.114	0.319 ± 0.113
		GAUSS	0.312 ± 0.117	0.310 ± 0.120	0.325 ± 0.117	0.326 ± 0.116
SR_MWS	GY	RR	0.236 ± 0.141	0.275 ± 0.147	0.231 ± 0.127	0.347 ± 0.134
		GAUSS	0.231 ± 0.144	0.273 ± 0.150	0.260 ± 0.128	0.370 ± 0.132

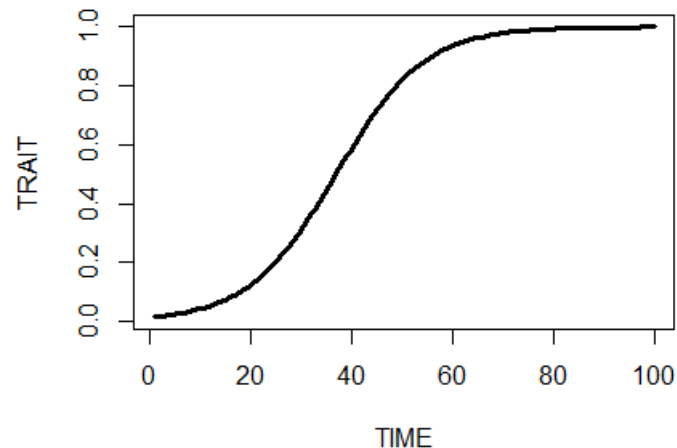
IB, incomplete blocks, field design; RC, row by column model; RCB_MVNG, random complete block model with moving means as covariable; MVNG, linear regression model with moving means as covariable; SR_FI, Santa Rosa under full irrigation; GY, grain yield; RR, Ridge regression kernel; GAUSS, Gaussian kernel; TKW, thousand kernel weight; DH, days to heading; NKS, number of kernels per spike; SR_MWS, Santa Rosa under mild water stress.

Getting a spatial covariate in R

- Two functions implemented in the NAM package
 - `NAM::NNsrc()`
 - `NAM::NNcov()`
- **NNsrc**
 - Input: data.frame with spatial information: Block, Row, Column
 - Output: “map” object indicating who is neighbors with who
- **NNcov**
 - Input: map, output from NNsrc, and the phenotype to create covariate
 - Output: covariate informing the phenotype neighbor average

Logistic curve (temporal neighbors)

1. Logit transform: $\psi = \log\left(\frac{y}{1-y}\right)$
 2. Fit the regression: $\psi = \beta_0 + t\beta_1 + e$
 3. Find fitted values: $\hat{\psi} = \beta_0 + t\beta_1$
 4. Transform back: $\hat{y} = \frac{\exp(\hat{\psi})}{1 + \exp(\hat{\psi})}$
- **Example:** Function that fits and predicts



Spatial analysis

```
# t = numeric vector, Days after planting
# x = numeric vector, Coverage
# p = numeric vector, Prediction
logistic = function(t, x, p){
```

FITTING THE LOGISTIC CURVE

```
y = log(x/(1-x)) # Logit transform
b1 = cov(t,y)/var(t) # Compute coefficient (beta1)
b0 = mean(y - b1*t) # Compute the intercept (beta0)
yhat = b0+b1*t # Fit the observed days
z = exp(yhat)/(1+exp(yhat)) # Transform back
```

PREDICTION ON NEW VALUES

```
pred = b0+b1*p # predict new values
pred = exp(pred)/(1+exp(pred)) # covert predictions
coef = c('b0'=b0,'b1'=b1)
final = list( fit = z, coef = coef, pred = pred )
return(final)}
```

USE R RUN THE EXAMPLE BELOW

```
t = c(28,35,42,50,56,65,73) # Time of observations
y = c(0.25,0.47,0.61,0.77,0.93,0.97,0.98) # Phenotype
p = 1:100 # Predict all values between day 0 and 100
Object = logistic(t, y, p) # Fit logit with function above
plot( Object$pred ) # Check the results
```