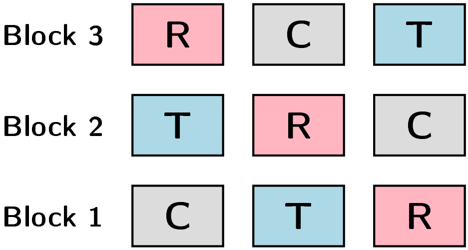
Geomatic Techniques to Support Phytosanitary Products Tests whithin the EPPO Standard Framework

Samuele Bumbaca

University of Turin

August 28, 2025

The Traditional Approach to Agricultural Trials



C Control
T Tested Product
R Reference Product

ANOVA Model:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

Where:

- y_{ij} = response
- μ = overall mean
- α_i = treatment effect
- β_j = block effect
- ε_{ij} = random error

Note:

This is the additive model. Modern approaches may include interaction terms: $\alpha_i \times \beta_j$

Key Assumptions of Traditional ANOVA

Statistical Assumptions:

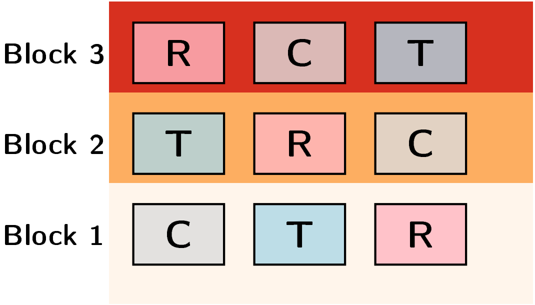
- **Randomization**: Treatments randomly assigned within blocks
- **Replication**: Each treatment appears in each block
- **Independence**: Observations are independent given the design
- **Homoscedasticity** : Equal variances across treatments
- **Normality**: Residuals follow normal distribution

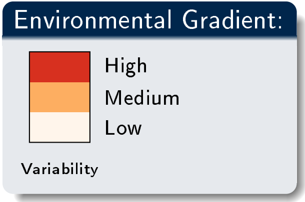
Consequences of Assumption Violations:

- **Invalid conclusions of parametric tests**: Need for non-parametric tests leading to reduced statistical power

Based on R. A. Fisher, Statistical Methods for Research Workers, in S. Kotz & N. L. Johnson (eds.), Breakthroughs in Statistics: Methodology and Distribution, pp. 66–70, Springer, New York, 1992.

The Right Blocking: Capturing Environmental Variability



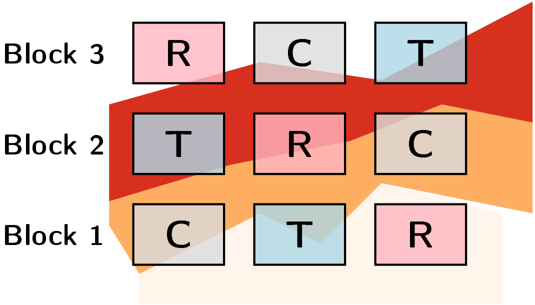


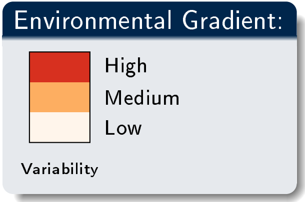
C Control
T Tested Product
R Reference Product

Success of Blocking Strategy:

- **Within-block homogeneity**: Treatments compared under similar conditions
- **Between-block heterogeneity**: Environmental gradient captured by block effects

The Wrong Blocking: Assumption Violation





C Control
T Tested Product
R Reference Product

Heteroscedasticity Assumption Violation Problem:

- **Blocks fail to capture environmental variability**: Treatments compared under different conditions
- **Invalid parametric test**: Residual variance differs across treatments

Current Limitations in Statistics for Agricultural Trials

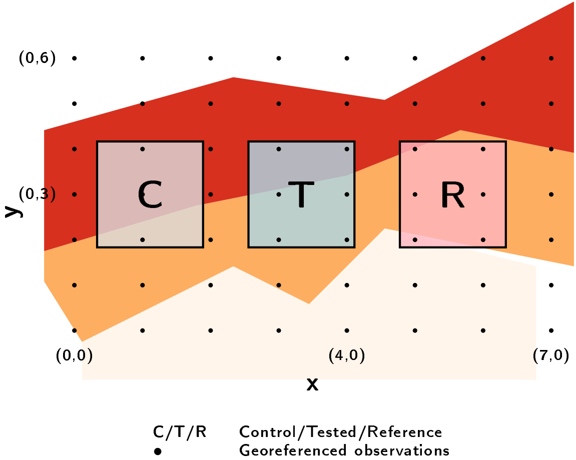
Traditional Approach Issues:

- **Human-dependent blocking**: Environmental variability assessment relies on experimenter experience
- **A priori identification**: Must identify variance sources BEFORE data collection

The Challenge:

How can we capture environmental variability mathematically rather than through human judgment?

Geostatistical Approach: Spatial Linear Mixed Models



Spatial LMM:

$$y(s_i) = \mu + \alpha_j + f(s_i) + \varepsilon_i$$

Where:

- $y(s_i)$ = response at s_i
- μ = overall mean
- α_j = treatment effect
- $f(s_i)$ = spatial random field
- ε_i = error
- $s_i = (x_i, y_i)$ = coordinates

Benefits:

- **No blocking**: Spatial correlation captures variability
- **Post-hoc**: No a priori variance identification
- **Homoscedasticity**: Assumption satisfied in more cases in respect blocking

Statistical Methods Comparison: Introduction

Comparison Objective:

Evaluate the performance of **traditional RCBD** versus **spatial geostatistical methods** (SpATS) in capturing environmental variability and estimating treatment effects.

Synthetic Dataset:

- **54 observations**(6×9 grid)
- **3 treatments**: Control (0 t/ha), Reference (0.5 t/ha), Test (1.0 t/ha)
- **3 blocks**(18 plots each)
- **Environmental zones**: Low (-1.5 t/ha), Medium (0 t/ha), High (+1.5 t/ha)

Tested Models:

1. **RCBD Model**: Linear Mixed Model with random block effects

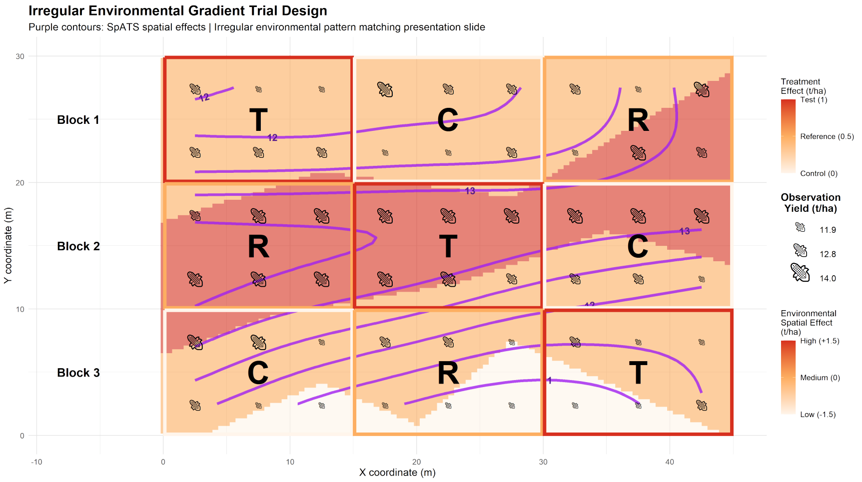
$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

2. **SpATS Model**: Spatial model with PSANOVA splines

$$y(s) = \mu + \alpha_i + f(s) + \varepsilon(s)$$

Where: α_i = treatment effects, β_j = block effects, $f(s)$ = spatial smooth

Statistical Methods Comparison: The Field Trial Design



Statistical Methods Comparison: Results

Model Performance (Mean Absolute Errors tonn/ha):

Model	Treat. Error	Env. Error
RCBD Model	0.13	0.62
SpATS Spatial	0.03	0.45

Treatment Effect Estimation (tonn/ha):

Treatment	True	RCBD	SpATS
Control	0.00	0.00	0.00
Reference	0.50	0.40	0.45
Test	1.0	0.69	0.94

Key Findings:

- **Both models satisfied assumptions**
- **SpATS outperformed RCBD**:
  - 3.8× better treatment effect estimation
  - 1.4× better environmental effect estimation
- **RCBD underestimated** by 20-31%
- **SpATS <6% error**

Implications:

Even when traditional RCBD meets statistical assumptions, **spatial modeling provides superior accuracy** in treatment effect estimation by properly accounting for environmental spatial variability.

The Missing Link: Spatial Coordinates

Geostatistical Methods Advantages:

- ✓ **Mathematical modeling** of environmental variability
- ✓ **Post-hoc analysis** - no need for prior knowledge of the environment variables and of their distribution
- ✓ **Superior performance** in handling spatial heterogeneity
- ✓ **EPPO recognized** approach (PP1/152(4) - Design and analysis of efficacy evaluation trials)

Current Barrier:

- × **Requires spatially referenced observations**
- × **Traditional manual assessments lack coordinates**
- × **Implementation gap** in practical field trials

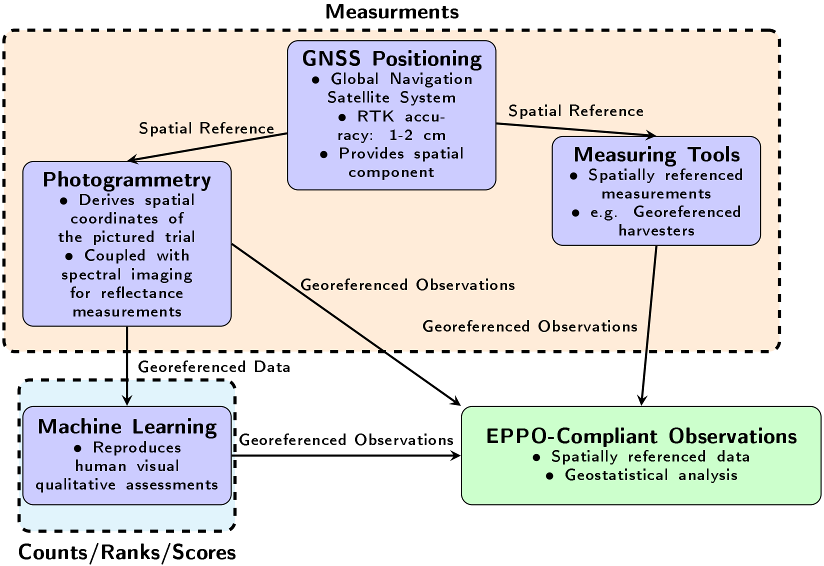
Central Research Question

Can geomatics technologies provide spatially referenced observations that enable geostatistical analysis within EPPO-compliant Plant Protection Product trials?

Specific Objectives:

1. Establish which geomatics technologies can be used to collect spatially referenced observations
2. Demonstrate the feasibility of collect spatially referenced observations in compliant with EPPO standards
3. Validate performance against traditional methods
4. Provide practical implementation guidelines

Geomatic Technologies: Workflow for Spatially Referenced Observations



Georeferencing EPPO Standard Assessments

Table: EPPO's types of variables

Type of Variable	Measurement	Ranking	Scoring
Binary			X
Nominal			X
Ordinal		X	X
Discrete	X		
Continuous limited	X		
Continuous not limited	X		

Summary from EPPO PP 1/152: Design and analysis of efficacy evaluation trials

Current State of Georeferencing in Agricultural Trials:

Measurements that can be tool-collected can be also easily georeferenced (e.g., yield harvesters). Whereas observation depends on experimenters' visual assessments (all not measurement variables, except visual counts), ML is required to spatial integration.

EPPO ML integration

EPPO PP 1/333(1): Digital Technologies in PPP Trials

ML integrated assessments must meet the same quality standards as manual assessments and require validation through comparison with manual assessments (golden sample).

Validation Benchmarks[a]

[a] Based on EPPO PP 1/333(1): Use of digital technologies in efficacy and selectivity trials

- **Continuous/Discrete**: $R^2 > 0.85$ (1:1 relationship)
- **Ordinal/Nominal**: Cohen's $\kappa > 0.7$
- **Binary**: Accuracy > 0.85

Table: EPPO's types of variables

	Type of Variable	Measurement	Ranking	Scoring
	Binary			X
	Nominal			X
	Ordinal		X	X
$\rightarrow$	Discrete	X		
	Continuous limited	X		
	Continuous not limited	X		