THE USE OF APPROPRIATE STATISTICAL TECHNIQUES IN ADDRESSING FOOD PRODUCTION CHALLENGES OF SMALLHOLDER FARMERS

Prof Peter Njuho

14th July 2022

Department of Statistics

Dedication

To my late parents:

- Godfrey Njuho Kioi
- Leah Wanjiru Njuho

Appreciation

Family

- My wife: Wanjiku Mungai
- My sons: Njuho Mungai

  Munyeria Mungai
- Daughter-in-law: Khanyisile Njuho

Colleagues

- Department of Statistics

Introduction

- Smallholder farmers encounter multifaceted challenges in food production
- Answers in food security lie in the hands of scientists
- Multi-disciplinary approaches
- Culture and dynamics in food production at farm level

The 1 acre plot

Motivation

- Growing up
- Daily household food
- Harvesting whatever was mature
- Questioned how and why these miseries
- Can smallholder farmers feed the world?
- Where are we now and how can we address these challenges?

Motivation - Questions

- Why have things remained the same despite the trainings & investment?
- Are our scientists recommending appropriate technologies?
- How do we increase food production at smallholder farms?
- If they are, why then do we still have these challenges?

Motivation - Questions

- Could assumptions made while developing these technologies misplaced?
- Do we really understand the domain of operation at smallholder farm level?
- How could we empower each farmer to make informed decision on adopting technology that works?

What do we know or Challenges?

- Research concerns:
  - improvement of an existing technology
  - formulating and nurturing an innovation
  - comparing technologies under different environment
- Conventional principals involve:
  - identification of the population of interest
  - deciding on the sampling producer to apply
  - implementing the plan, collecting data
  - analysing data and interpreting the findings

What do we know or Challenges?

- Statistician role:
  - designing of the experiment
  - analyzing the data, and
  - writing the statistical section
- A common understanding required
  - to provide solutions to the challenges faced by the smallholder farmers
  - our culture is resistant to new food
  - want to eat what is common and produced elsewhere

What do we know or Challenges?

- encourage the production of traditional foods
- are resilient to the effects of climate change

- widen the range of Africa’s food baskets
  - increase nutrient bioavailability
  - drastically cut food insecurity
  - address malnutrition on the continent
  - integrate African foods both from plants and animals into our different food systems

What do we know or Challenges?

- Smallholder farms differ in
  - climate, soil type, management, farmer's knowledge, preferences, and access to new information
- Overcoming these limitations
  - evaluating the new technologies against the farmer's technology
  - apply proper statistical procedures

Scenario -1

- Smallholder farmer
- On-farm animal feed experiment
- Compare new feed (**A**) against farmers' traditional feed (**B**)
- Sale of milk is the farmer's income
- Two cows available Cow 1 – receives A

  Cow 2 –receives **B**
- Increase in milk production from cow 1
- Farmer abandon feed (**B**) and
- Feed both animals with (**A**)

A compromised experiment

Scenario - 2

- East coast fever (ECF) a tropical disease
  - Test a new management strategy to control ECF
  - Strategy involves **vaccination** and **dipping** in acaricide
  - Treatment combinations were:
    - vaccination and dipping (VD)
    - no vaccination and dipping (NVD)
    - vaccination and no dipping (VND)
    - no vaccination and no dipping (NVND)

Scientific control (NVND) - Impossible

Lesson from scenarios

- Understand the social and economic factors
- Farmer - researcher perception differ
- Make farmer understand long-term benefits
- No amount of data manipulation will result to meaningful solutions to such messy trial

Research Vs Difficulties in Statistics

RESEARCH PROCESS	STATISTICS DEFICIENCES
Problem Identification	Survey Design & Interpretation
Hypothesis Formulation	Relationship to Expt'l Design & prior Tests
Research Approaches	Incorrect Approaches to Mult-Loc & Repeat Meas
Data Compilation & Analysis	Data Quality Control & Model Assumptions
Synthesis & Interpretation	Insufficient Use of Design & Treatment Structures

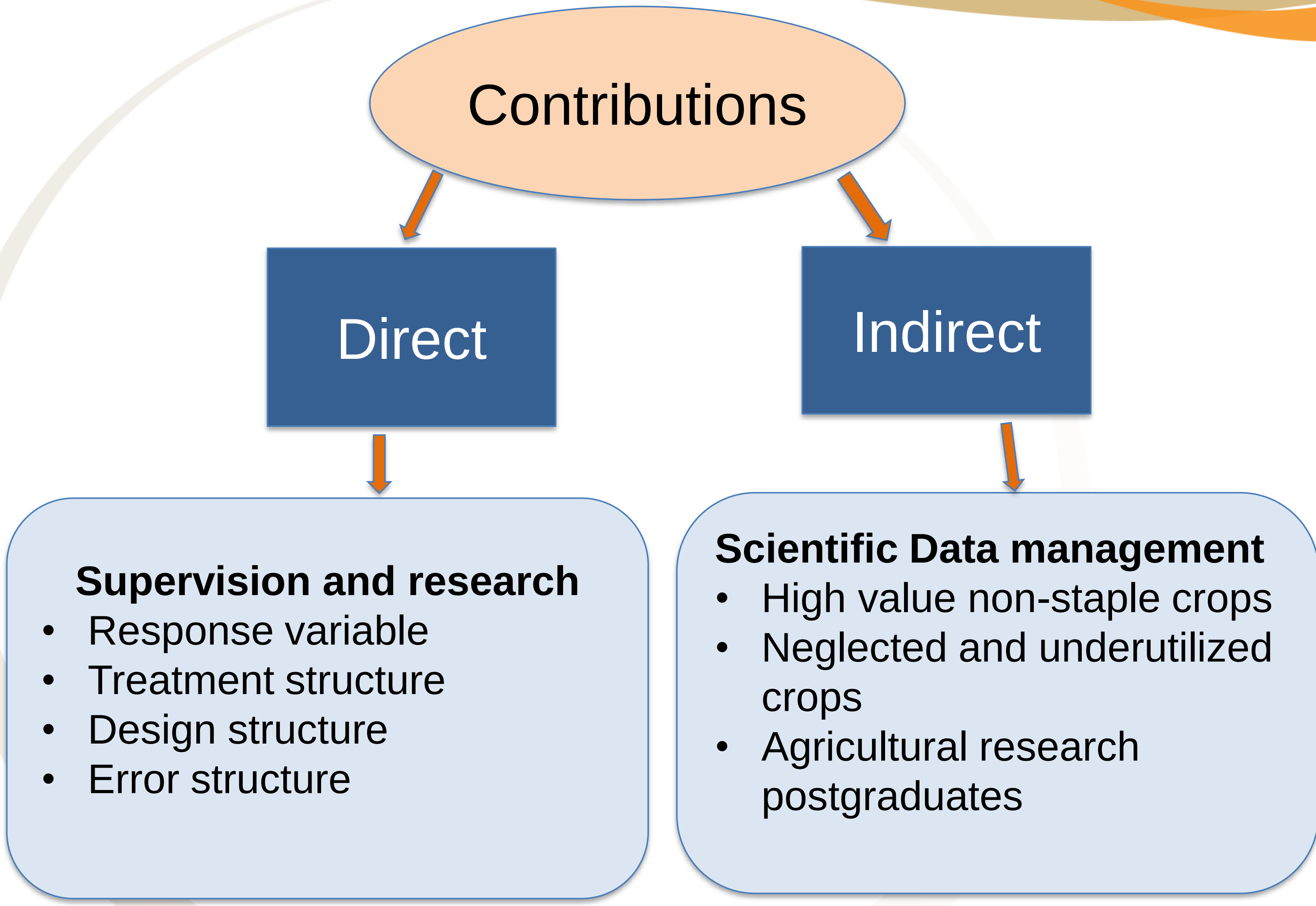
Questions to answer

- Any justification for recommending **New technology** to all farmers?
  - based on the assumption of a **Common Control**
- Are we making the correct decision at **Farm level?**
  - to accept a new technology

Possible Answers

- All farms have the **same** practice - Common Control
- **Each** farm has it’s control
- Common control for a **group** of farms
- Groups:
  - Group **A** - Farmers who would adopt the new variety
  - Group **B** - Farmers who would stick to their own variety
  - Group **C** - Farmers who would find it difficult to decide

Contribution to solving smallholder farmer challenges


Contributions
Direct
Indirect
Supervision and research
• Response variable
• Treatment structure
• Design structure
• Error structure
Scientific Data management
• High value non-staple crops
• Neglected and underutilized crops
• Agricultural research postgraduates

DIRECT: Linear Mixed Model

$$y = X\beta + Zu + \varepsilon$$

Response Variable = Treatment Structure + Design Structure + Error Structure

Response Variable	Treatment Structure	Design Structure	Error Structure
2 PhDs & 1 MSc	2 PhDs	1 PhD & 2 MScs	1 PhD & 1 MSc
(A)	(B)	(C)	(D)

Work on Response Variable (A)

2PhD and 1 MSc produced

Theses:

- Statistical Techniques for Combining Parameter Estimates: Case of Food Production in Sub-Saharan Africa
- Discrete Regression Models Using Quasi-Likelihood Estimation Method
- Understanding patterns of aggregation in count data

Issues Addressed

- Understanding aggregation, seasonality and other unique patterns associated with parasitological data
- Understanding the occurrence of certain species causing diseases in grazed animals
- Low production affect the livelihood of the smallholder farmers

Issues Addressed

- Quasi-score test for testing over dispersed data
- Agricultural production estimates combined to a single stable estimate
- Contribute to the formulation policies on food insecurity

Research output -A

Publications in peer-reviewed journals

- A Comprehensive Approach for Integrating Meta-analysis into Structural Equation Modelling: Case of Food Production. International Journal of Agricultural and Statistical Sciences. Vol.17(2), pp 479-492
- Use of linear mixed effects model in meta-analysis for studies with multiple outcomes. International Journal of Agricultural and Statistical Sciences. Vol. 16 (1), pp11-21
- Improved Structural Equation Models using Factor Analysis. Pakistan Journal of Statistics and Operation Research, Vol. XIV(4), pp 995-1012
- Statistical models for helminth faecal egg counts in sheep and goats. Small Ruminant Research Journal. Vol. 170, pp 26-30

Work on Treatment Structure (B)

2PhD Theses

- Using Mixed Models to Analyze Data from On-Farm Trials
- Analysis of Linear Mixed-Models with an Extension to Three or More Factors Each Having Both Fixed and Random Levels

Issues Addressed

- Performance of one or more improved technologies in comparison to the farmer's practice
- Unpack the assumption-smallholder farmers have a common control
- Same experimental design applicable to all
- Comparing a treatment to different controls
- Different experimental design on each farm

Mixed Model Analysis

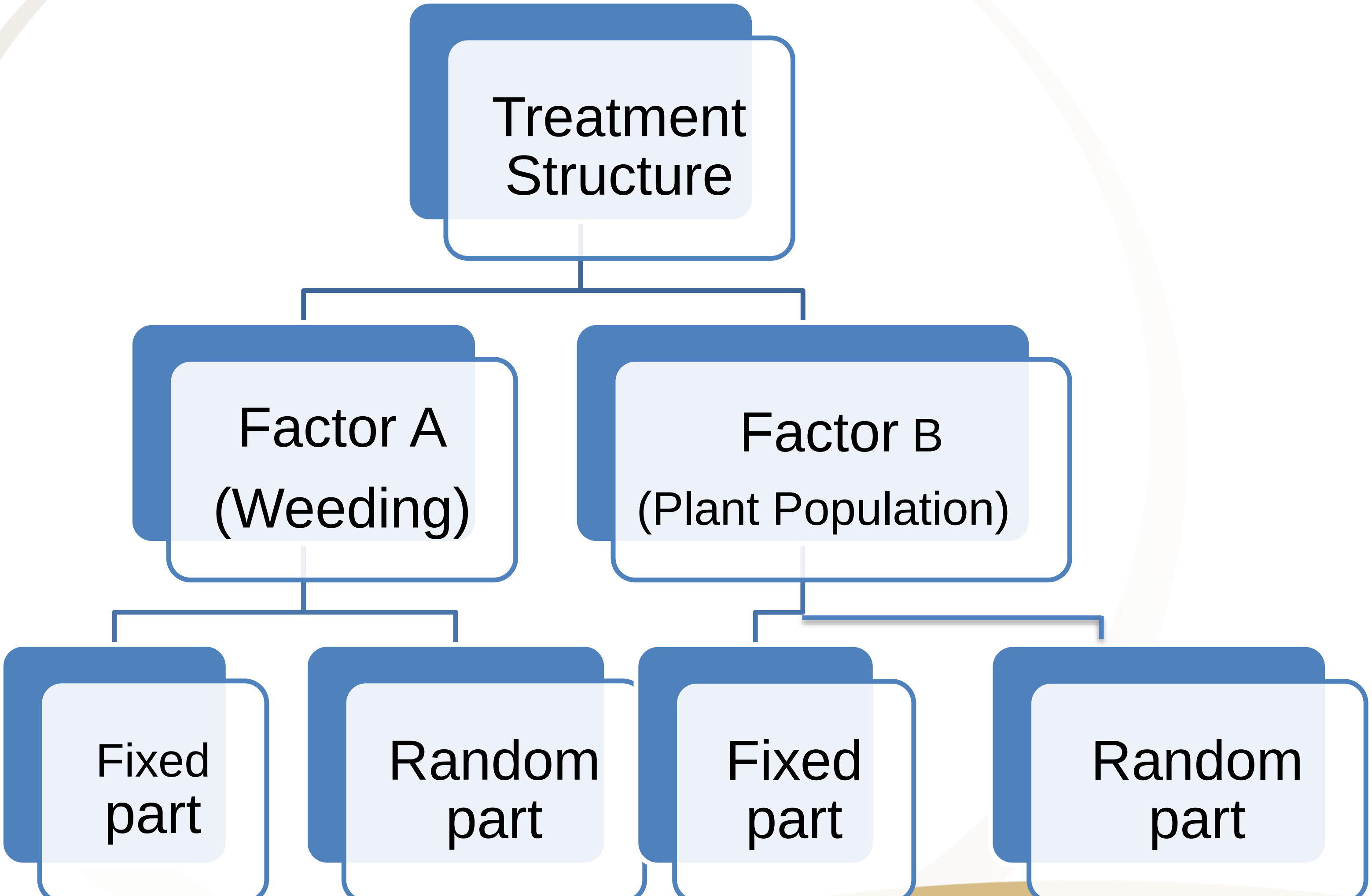
- Two components

Y = Fixed Part + Random Part

Random Part = Blocking + **Error**

- Must state which factors are fixed and which are random

Two-way Treatment Structure



Types of Models

- Fixed-fixed (FF)
- Fixed-random (FR)
- Random-fixed (RF)
- Random-random (RR)

Models for each of the combinations

		Levels of Factor B	
		$1,2,\ldots,f_b$	$f_b+1, f_b+2,\ldots, f_b+r_b=b$
Levels of Factor A	$1,2,\ldots, f_a$	Fixed-Fixed $y_{FFijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\varepsilon_{FFijk}$	Fixed-Random $y_{FRijk}=\phi_i+b_{FRj}+c_{FRij}+\varepsilon_{FRijk}$
	$f_a+1, f_a+2,\ldots, f_a$	Random-Fixed $y_{RFijk}=a_{RFi}+\omega_j+c_{RFij}+\varepsilon_{RFijk}$	Random-Fixed $y_{RRijk}=\mu_{RR}+a_{RRi}+b_{RRj}+c_{RRij}+\varepsilon_{RRijk}$

Example 1

Smallholder farm experiment

Two Factors**:** Weeding methods & Plant Populations

- Fixed Part:
  - Weeding methods (FW1 & FW2)
  - Plant populations (FP1, FP2 & FP3)
- Random part:
  - Weeding methods (RW1, RW2 & RW3)
  - Plant populations (RP1 & RP2)

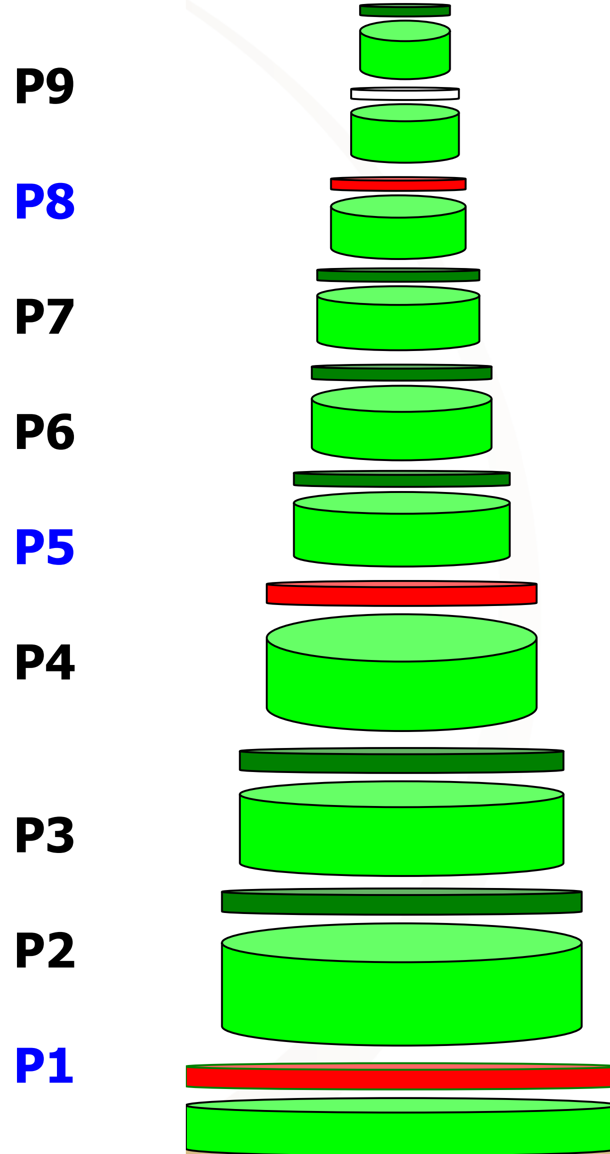
Example 2

Pulp yield estimation per tree

Two factors – Height and Diameter

- Height
  - Fixed (FP1, FP5 & FP8) levels
  - Random (RP2, RP3, RP4, RP6, RP7 & RP9) levels
- Diameter
  - Fixed (FD1 & FD4) levels
  - Random (RD2, RD3 & RD5) levels

Example 2



Example 2
Disc divided into sections

Five Diameters

- 2 Fixed levels
- 3 Random levels

Example 2

- **3 fixed** levels test for the height effect
- **6 random** levels estimate height variability
- Effect of diameter and variability also assessed
- Test: Interaction between height and diameter
- Estimate: Interaction variance component

Research output - B

Publications in peer-reviewed journals

- Analysis of Linear Models with One Factor Having Both Fixed and Random Levels. Journal of Communications in Statistics-Theory and Methods. Vol.34(9) Pp 1979-1989
- Analysis of Linear Models with Two Factors Having Both Fixed and Random Levels. Journal of Communications in Statistics Theory and Methods, 38(14), Pp 2348-2365
- Construction of a linear mixed model with each factor having both fixed and random levels: A case of split-split plot structure in a RCBD. International Journal of Agricultural and Statistical Sciences. Vol.17(2), pp 501-518
- Repeated-Measures Analysis in the Context of Heteroscedastic Error Terms with Factors Having Both Fixed and Random Levels. Stats 2022, 5(2), 458-476
- Mixed models approach to on-farm trials: An alternative to meta-analysis for comparing one treatment to possibly different controls. Proceedings of the 1995, Kansas State University Conference on Applied Statistics in Agriculture. Pp 196-213

Work on Design Structure (C)

1 PhD and 2 MSc Theses

- Using spatial modelling techniques to improve data analysis from agricultural fields trials
- Assessment of Variability in On-Farm Trials: A Uganda Case
- Analysis and Efficiency of Systematic Designs in Intercropping Experiments

Issues Addressed

- Smallholder farms characterized by
  - unique constraints, high variability within and between farms, missing observations, and differing farming knowledge
- Strife to allow farmers' conditions to remain the same while the experiment on the farm progress
- Simple & effective designs needed
  - control the inherent farm variability

Issues Addressed

- Incomplete block designs
  - Allows a farmer decide on the new interventions
- Improved design then model to account for the spatial variation

On-farm mathematical model

$$y = \mu + f + b(f) + \tau + f\tau + \epsilon$$

Response

Overall mean

Farm effect

Block within farm effect

Treatment effect

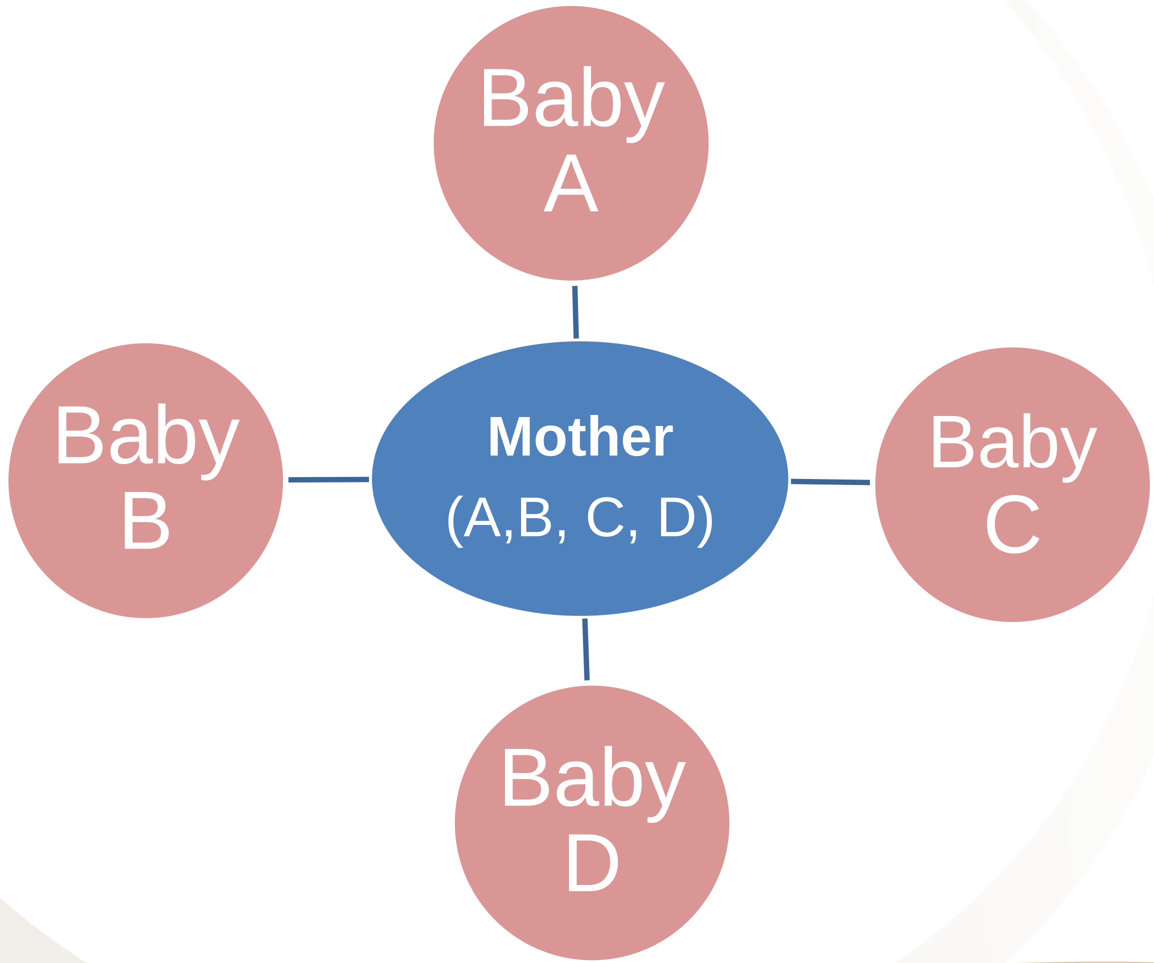
Interaction effect

Random error effect

Selection of experimental material

- Monitor the trend per block level
  - Soil characteristics non-random, fertility trend, spatial autocorrelation, and periodicity
- A strategy for modelling spatial variability
- Allow for investigation - many technologies per farm

Mother-baby design



Mother-baby design

- Investigation of many technologies on the farm
- Farmer’s chance to try any interventions
- Allow for non-replication on the farm
- Block size depends on farmer’s ability to handle
- **Mother** accommodate more technologies
- **Baby** select any technology to try
- **On station** – Mother & **babies** smallholder farms

Research output -C

Publications in peer-reviewed journals

- Smoothing fertility trends in agricultural field experiments. Statistics: A Journal of Theoretical and Applied Statistics. Vol. 43(3), Pp 275-289
- Monitoring field variability using confidence interval for coefficient of variation. Journal of Communications in Statistics Theory and Methods. Vol. 37(6), Pp 831-846
- Improvement on Papadakis covariate to account for spatial variation. Journal of Agricultural, Biological, and Environmental Statistics
- Optimum plot block dimensions and effect of plot shape on significance tests in potato. Journal of the Ethiopian Statistical Association, Vol. XII, Pp 39-59.
- The Efficiency of Incomplete Block Designs in On-Farm Trials

Work on Error Structure-D

1 PhD and 1 MSc Theses

- Using spatial modelling techniques to improve data analysis from agricultural fields trials
- Assessment of Variability in On-Farm Trials: A Uganda Case
- Analysis and Efficiency of Systematic Designs in Intercropping Experiments

Issues Addressed

- Smallholder farmers’ production systems
- Involve constraints that require different solutions
- Characterize different sources of variability
  - High variability leads to inefficient estimates
- Farming systems characterized by:
  - Crops grown & combinations, animals kept, educational & farmer’s wealth, cultural practices & socio-economic status

Issues Addressed

- Contributors of direct or indirect variation:
  - Agronomic, animal husbandry and socio-economic
- A stochastic relationship

$$y = f(G, E, M, S) + Error$$

G - Plant genotype effect

E - Biophysical environmental effect

M - Past and present management effect

S - Socioeconomic factors effect

Error-Random effects

Indirect Contribution

Indirect contribution

Training agricultural research scientists under:

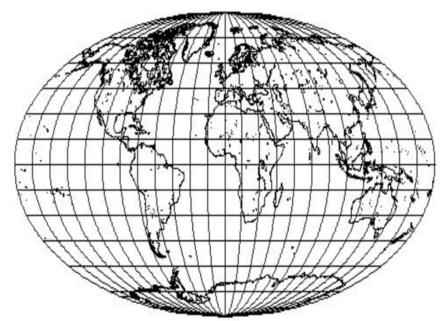
- RUFORUM - Scientific data managements
- AWARD- Science writing and proposal writing skills
- ASARECA- Research on high value non-staple crops
- IFS - Research on neglected and underutilized crops

Professional contributions

- International Biometric Society (IBS)
  - Board member of directors
  - Educational representative
  - Chair of the Education committee
  - Chair of the Young Statistician Showcase section

Professional contributions

- Sub-Saharan Africa Network (SUSAN)
  - Coordinator
- South Africa Statistical Association (SASA)
  - Executive member
- IBS Group South Africa
  - National Secretary



Take home message

- Critical review on postgraduate curriculum
- Enhance farmer training centres
- Sensitize young students to pursue career in agricultural & statistics
- IT research and innovation on smallholder farms
- Design of experiments which apply to smallholder
- Act now to reverse the decline in food production

Take home message

- Creative scientists are the engines that run innovative technology generation and adoption
  - How can we best **ignite** this creativity?
- Statisticians are **servants** of science
  - Collaborate with them
- Ensure research conducted use gender **lens** to help in maintaining harmony amongst smallholder farmers

Take home message

- Smallholder farmers are most affected by climate change

- Let us all join hands in alleviating the miseries of smallholder farmers in **all** ways

Acknowledgement

KALRO
Kenya Agricultural & Livestock Research Organization

RUFORUM
Capacity Building in Agriculture
RUFORUM C10 Implementation

KANSAS STATE UNIVERSITY

AWARD
African Women in Agricultural Research and Development

university of south africa

Thank you