Week-04-L-01

Agricultural Statistics in Practice

Regression Path Analysis

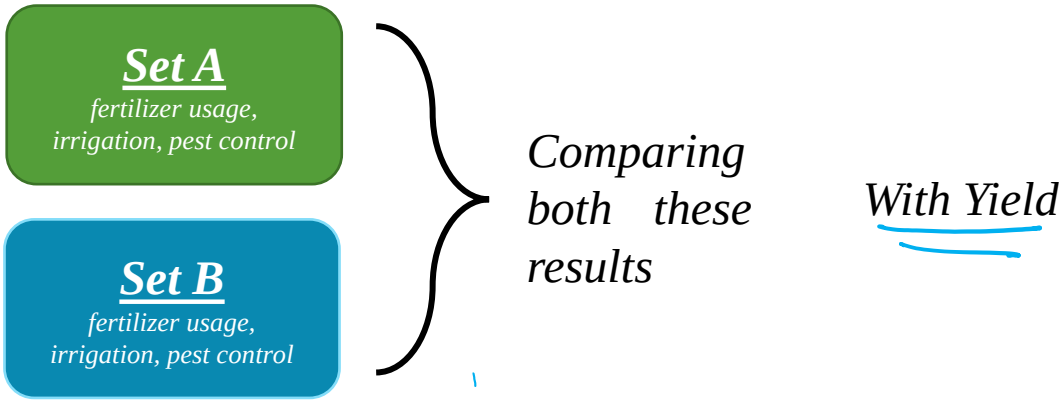
Introduction

Prof. J. Ramkumar

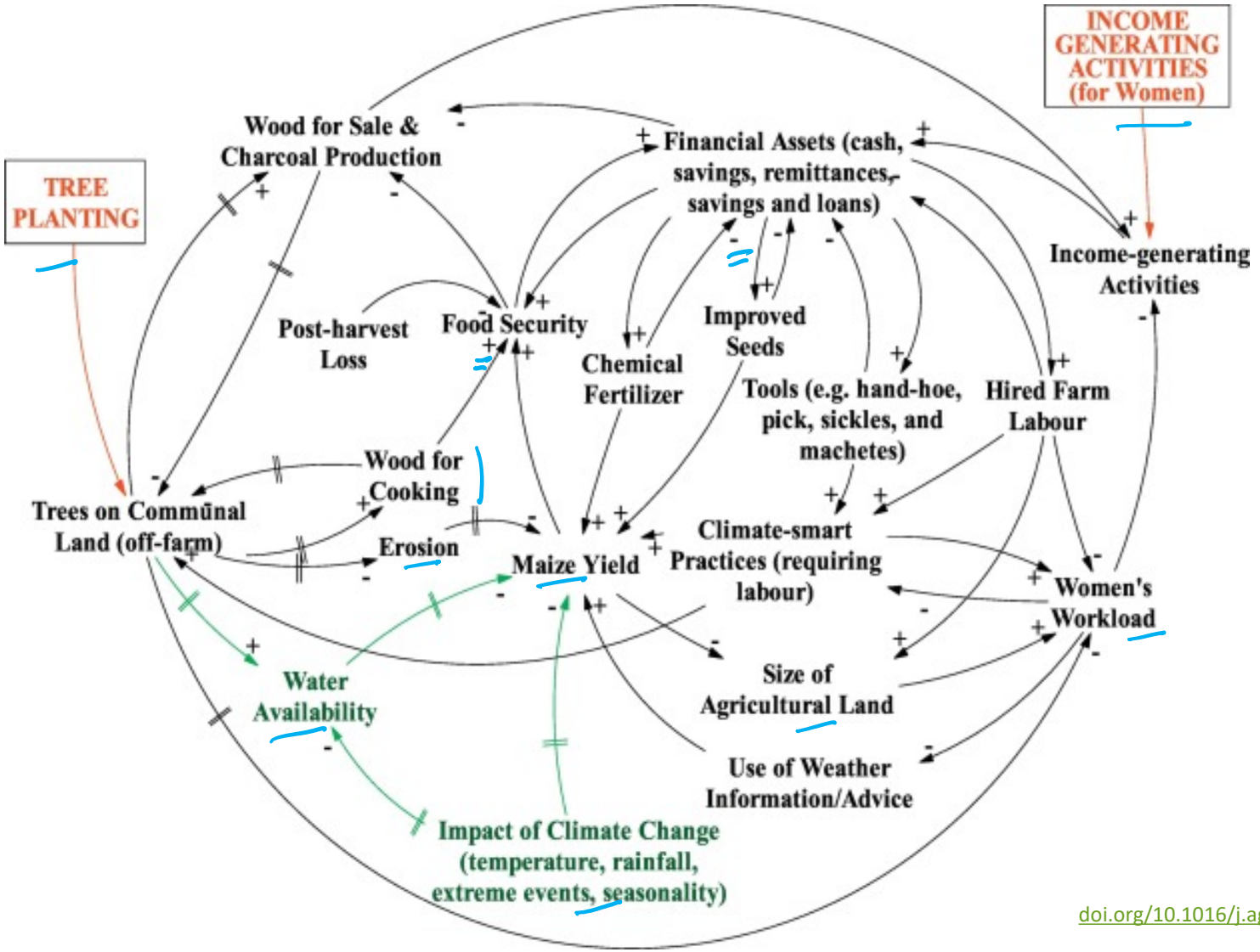
Dept. of ME & Design
Indian Institute of Technology Kanpur

A case study

- A researcher wants to investigate the relationship between different agricultural practices (e.g., fertilizer usage, irrigation, pest control) and crop yield.
- They also want to examine how these practices ***mediate*** the impact of environmental factors (e.g., temperature, rainfall) on crop yield.
- Analyze the direct and indirect effects of these variables on crop yield.

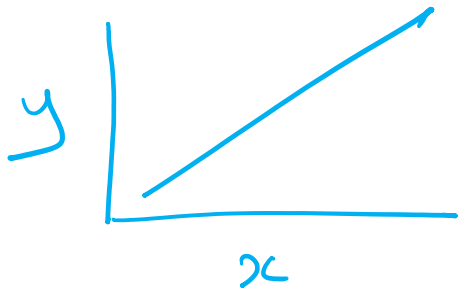


A Model



Source:
doi.org/10.1016/j.agsy.2018.12.008

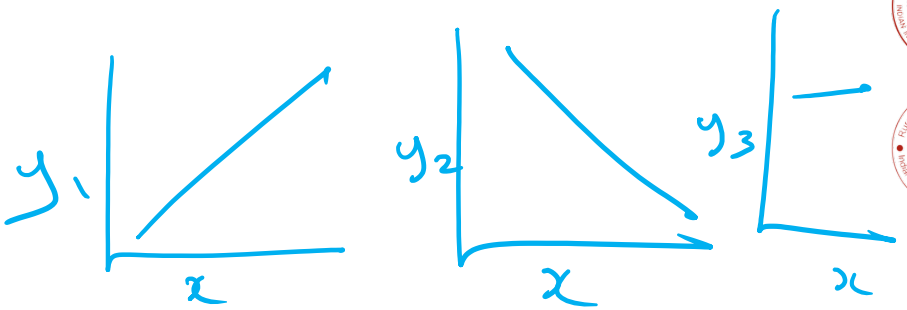
Terminologies



- ***Independent variable:*** The variable(s) that are believed to have a direct or indirect influence on the dependent variable (e.g., agricultural practices, environmental factors).
- ***Dependent variable:*** The variable of interest that is being predicted or explained (e.g., crop yield).
- ***Mediating variable:*** An intermediate variable that links the independent and dependent variables, helping to explain the relationship between them (e.g., crop growth, nutrient uptake).
- ***Direct effect:*** The direct relationship between an independent variable and the dependent variable.
- ***Indirect effect:*** The relationship between an independent variable and the dependent variable that is mediated by one or more mediating variables.

$x_1, x_2, x_3 \ldots x_n \Rightarrow y_n$

Regression Analysis



- Used to observe relationships among multiple variables by assessing direct and indirect effects.

Assumptions

- Sample is representative of pop at large & error in measuring independent var is 0
- Deviations from the model have an expected value of zero.
- The variance of the residuals e_i (actual y value − predicted y value)is constant across observations (homoscedasticity).
- The residuals e_i are uncorrelated with one another. Mathematically, the variance–covariance matrix of the errors is diagonal.

$y = mx_1 + m_1x_2 + m_2x_3$

10 mangoes → 11 mangoes.

General Model & Usability in agriculture

$$Y_i = f(X_i, \beta) + e_i$$

dependent variable.

independent variable

Unknown parameter which is Scalar or Vector

error term

- Farmers can ***prioritize*** the most ***influential practices*** and focus their ***resources*** on ***optimizing*** those factors that have the ***greatest*** impact on crop ***productivity***.
- Enables ***resource allocation***, such as determining the optimal levels of fertilizer, irrigation, or pest control practices for ***maximizing yield*** under specific environmental conditions.

Classification of Regression Analysis

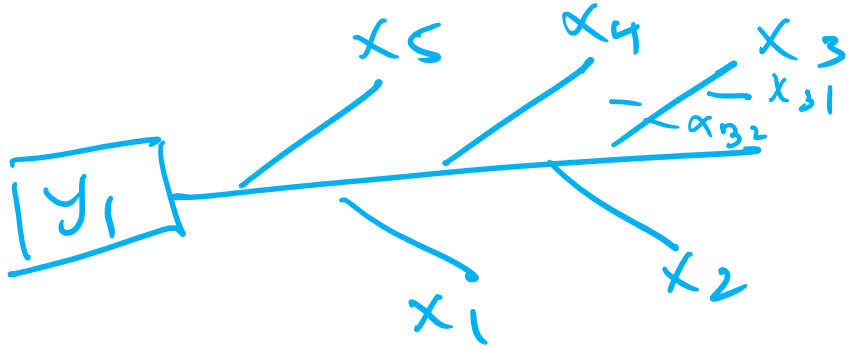
- ***Simple Linear Regression Analysis:***
  - Involves one independent variable, one dependent variable, and potentially one mediating variable.
  - $X_1 = f(X_2, u)\ X_2 = f(X_1, v)$ where "u" and "v" are the respective random components
- ***Multiple Linear Regression Analysis:***
  - Involves multiple independent variables, one dependent variable, and potentially one or more mediating variables.
  - $X_1 = f(X_2, X_3, X_4, X_5 \ldots\ldots X_k, u)\ X_2, X_3, X_4, X_5 \ldots\ldots X_k$ are the variables

$$y = mx + c$$

$$y = mx + m_1x_1 + m_2x_2 + m_3x_3 \ldots + x_1x_2 + x_2x_3 + x_3x_1 + \ldots + c$$

Does Regression Indicate Cause and Effect Relationship?

- Allows us to ***estimate the expected change*** in the dependent variable while holding other variables constant.
- However, it ***does not determine the causality*** between the independent variable and the observed change in the dependent variable.
- Regression analysis ***establishes the linear relationship*** among variables rather than examining cause and effect.
- To assess causality, Granger's causality test can be employed to investigate the relationship between variables.



Thank You


ideas to products
IMAGINEERING
LAB | IIT KANPUR
Rural Technology Action Group
Indian Institute of Technology Kanpur
MedTech
IIT KANPUR