

The background of the entire image is a dark blue field filled with a pattern of red dots. These dots are arranged in a way that they form a large, faint, stylized circular shape in the center, with the density of the dots increasing towards the center.

# HUST

**ĐẠI HỌC BÁCH KHOA HÀ NỘI**  
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

ONE LOVE. ONE FUTURE.





**ĐẠI HỌC  
BÁCH KHOA HÀ NỘI**  
HANOI UNIVERSITY  
OF SCIENCE AND TECHNOLOGY

# PHÂN TÍCH CÁC YẾU TỐ ẢNH HƯỞNG ĐẾN SẢN LƯỢNG LÚA BẰNG MÔ HÌNH HỒI QUY TUYẾN TÍNH BỘI

**Người thực hiện:**

Nguyễn Đức Hùng – 20237441

Nguyễn Khắc Chí – 20237422

Hoàng Quang Huy – 20237444

Nhóm: 20

**ONE LOVE. ONE FUTURE.**

1. Giới thiệu bài toán
2. So sánh giữa các mô hình
3. Kết luận và đánh giá



# Giới thiệu bài toán

- ▶ Lúa gạo là cây lương thực chủ lực, giữ vai trò quan trọng trong an ninh lương thực.
- ▶ Sản lượng lúa chịu tác động của nhiều yếu tố như diện tích, giống, phân bón, lao động,...
- ▶ Cần định lượng mức độ ảnh hưởng của các yếu tố này.
- ▶ Nghiên cứu sử dụng hồi quy tuyến tính bội.



Ảnh cánh đồng

- ▶ Xác định các yếu tố chính ảnh hưởng đến sản lượng lúa.
- ▶ Định lượng mức độ tác động của từng yếu tố thông qua mô hình hồi quy tuyến tính bội.
- ▶ Đánh giá mức độ phù hợp của mô hình với dữ liệu thực tế.
- ▶ Đề xuất biện pháp nhằm nâng cao hiệu quả sản xuất lúa.



- ▶ Bộ dữ liệu *RiceFarms* phản ánh tình hình sản xuất lúa tại Indonesia.
- ▶ Dữ liệu dạng bảng với tổng cộng 1.026 quan sát, mỗi quan sát tương ứng với một nông hộ trồng lúa.
- ▶ Bao gồm thông tin về đặc điểm nông hộ, điều kiện canh tác, các yếu tố đầu vào, lao động, giá cả và sản lượng lúa.
- ▶ Các nông hộ được khảo sát tại nhiều làng: Wargabinangun, Langan, Gunungwangi, Malausma, Sukaambit và Ciwangi tại tỉnh Tây Java, Indonesia.



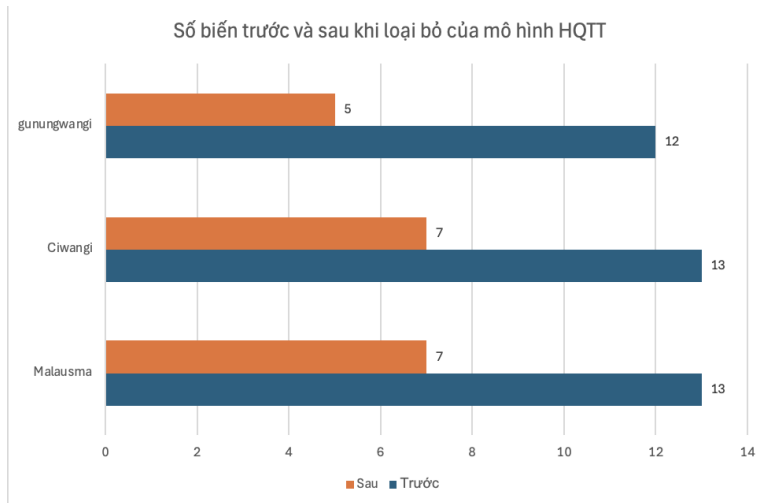
Tây Java, Indonesia



- ▶ Nghiên cứu tập trung phân tích dữ liệu sản xuất lúa của các nông hộ tại Indonesia.
- ▶ Nhóm lựa chọn dữ liệu của **3 vùng** đại diện để xây dựng mô hình hồi quy.
  - ▶ Chí: Ciwangi
  - ▶ Huy: Gunungwangi
  - ▶ Hùng: Malausma
- ▶ Việc lựa chọn nhằm phục vụ mục tiêu ước lượng, so sánh và đánh giá các mô hình hồi quy.
- ▶ Kết quả nghiên cứu giúp minh họa sự khác biệt trong mối quan hệ giữa các yếu tố đầu vào và sản lượng lúa giữa các vùng.

# So sánh giữa các mô hình

# So sánh số lượng biến trước và sau khi lựa chọn mô hình



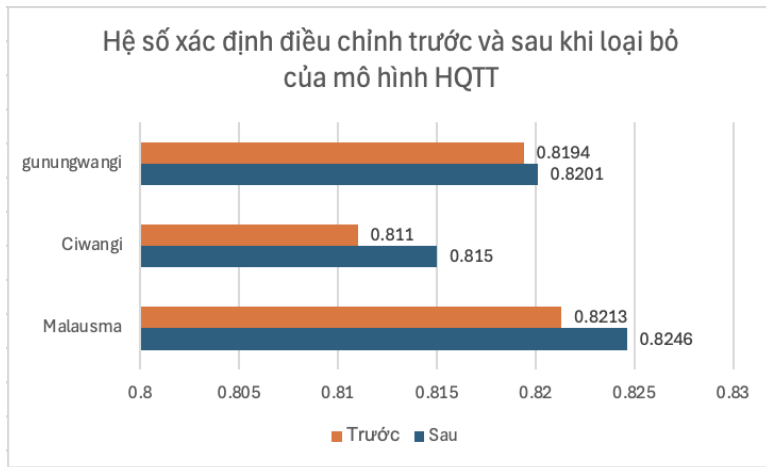
# So sánh 3 mô hình hồi quy

<b>Biến</b>	<b>Mô hình Ciwangi</b>	<b>Mô hình Malausma</b>	<b>Mô hình Gunungwang</b>
Hàng số	513,54	-409,7	-607,8
Diện tích canh tác (ha)	1492,88	1605,57	1668,44
Lượng giống sử dụng (kg)	-	11,79	-
Lượng phân lân (kg)	6,98	-	-
Giá phân lân (Rupiah/kg)	-	-	8,56
Lượng phân urê (kg)	-	1,88	2,21
Giá phân urê (Rupiah/kg)	-	6	-
Chi phí thuốc BVTV (Rupiah)	0,03	-	-
Lao động gia đình (giờ)	-1,36	-	0,76
Lao động thuê (giờ)	-	-	0,43
Tổng lao động (giờ)	1,14	0,4	-
Tiền công lao động (Rupiah/giờ)	7,45	5,64	-
Giá lúa (Rupiah/kg)	-10	-5,17	-

# Kiểm định ý nghĩa của hồi quy

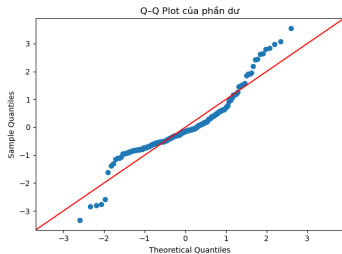
- ▶ Kiểm định F được sử dụng để đánh giá ý nghĩa thống kê tổng thể của mô hình hồi quy.
- ▶ Giả thuyết kiểm định:
  - ▶  $H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$  (mô hình không có ý nghĩa).
  - ▶  $H_1: \exists \beta_j \neq 0$  với ít nhất một  $j = 1, 2, \dots, k$ .
- ▶ Kết quả kiểm định cho thấy p-value của cả ba mô hình tại ba khu vực đều bằng 0.000.
- ▶ Vì p-value  $< 0,05$ , bác bỏ giả thuyết  $H_0$ .
- ▶ Kết luận: Mô hình hồi quy có ý nghĩa thống kê.

# So sánh mức độ phù hợp của mô hình

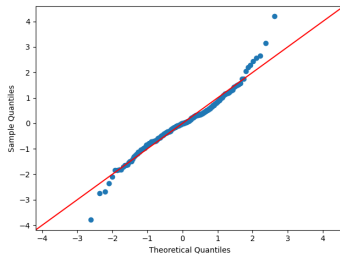


# Kết luận và đánh giá

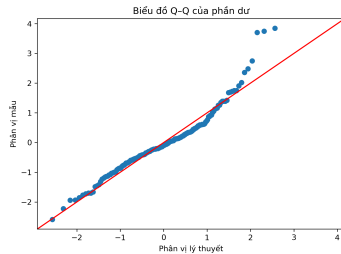
# Kiểm tra phân phối chuẩn của phần dư



Ciwangi



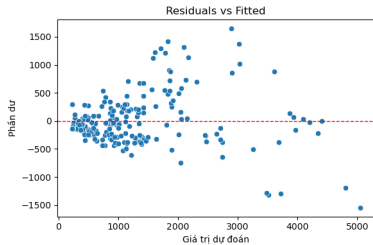
Gunungwangi



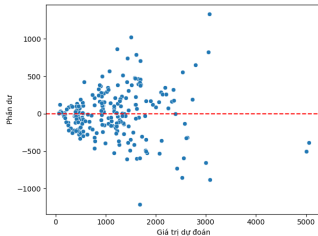
Malausma



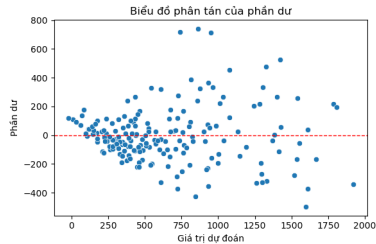
# Biểu đồ phân tán của phần dư



Ciwangi



Gunungwangi



Malausma

- ▶ Ba mô hình được xây dựng **riêng cho từng vùng**.
- ▶ Tất cả các mô hình đều có  $R^2_{\text{adj}} > 0,81$ :
  - ▶ Mô hình giải thích tốt biến thiên của sản lượng lúa.
- ▶ Biến diện tích canh tác xuất hiện ở trong cả ba mô hình và có hệ số hồi quy lớn.
  - ▶ phản ánh đúng bản chất của nông nghiệp, diện tích lớn thì sản lượng lúa sẽ lớn.
- ▶ Sự khác nhau về tập biến phản ánh **đặc điểm sản xuất khác nhau giữa các vùng**.

# Phân tích mô hình hồi quy vùng Ciwangi

- ▶ Phân lân (6,98) và chi phí thuốc BVTV (0,03)
  - ▶ Cần bón phân lân khi đất thiếu lân, khi đất chua, đất kiềm hoặc đất nghèo dinh dưỡng.
  - ▶ Môi trường có sâu bệnh.
- ▶ Lao động gia đình (-1,36)
  - ▶ Gợi ý hiệu quả lao động gia đình thấp hoặc dư thừa lao động.
- ▶ Tiền công lao động (7,45)
  - ▶ Tăng năng suất lao động dẫn tới sản lượng lúa tăng
- ▶ Giá lúa (-10)
  - ▶ Khi giá lúa tăng, người tiêu dùng có nhu cầu tìm sản phẩm thay thế.

$$\begin{aligned}\widehat{\text{Sản lượng lúa}}_{\text{Ciwangi}} = & 513,54 + 1492,88 \cdot \text{Diện tích canh tác} \\ & + 6,98 \cdot \text{Lượng phân lân} + 0,03 \cdot \text{Chi phí thuốc BVTV} \\ & - 1,36 \cdot \text{Lao động gia đình} + 1,14 \cdot \text{Tổng lao động} \\ & + 7,45 \cdot \text{Tiền công lao động} - 10 \cdot \text{Giá lúa}\end{aligned}$$

# Phân tích mô hình hồi quy vùng Malausma

- ▶ Lượng phân urê (1.88)
  - ▶ Thiếu đạm nito, là đất chua phèn hoặc đất trung tính.
- ▶ Tiền công lao động (5.64)
  - ▶ Tăng năng suất lao động dẫn tới sản lượng lúa tăng
- ▶ Giá lúa (-5.17)
  - ▶ Khi giá lúa tăng, người tiêu dùng có nhu cầu tìm sản phẩm thay thế.

$$\begin{aligned}\widehat{\text{Sản lượng lúa}}_{\text{Malausma}} = & -409,70 + 1605,57 \cdot \text{Diện tích canh tác} \\ & + 11,79 \cdot \text{Lượng giống sử dụng} + 1,88 \cdot \text{Lượng phân urê} \\ & + 6 \cdot \text{Giá phân urê} + 0,4 \cdot \text{Tổng lao động} \\ & - 5,17 \cdot \text{Giá lúa}\end{aligned}$$

- ▶ Lượng phân urê (2.21)
  - ▶ Thiếu đạm nito, là đất chua phèn hoặc đất trung tính.
- ▶ Lao động gia đình (0.76) và lao động thuê (0.43)
  - ▶ Các nông hộ còn thiếu nhân lực

$$\widehat{\text{Sản lượng lúa}}_{\text{Gunungwangi}} = -607,8 + 1668,44 \cdot \text{Diện tích canh tác} \\ + 8,56 \cdot \text{Giá phân lân} + 2,21 \cdot \text{Lượng phân urê} \\ + 0,76 \cdot \text{Lao động gia đình} + 0,43 \cdot \text{Lao động thuê}$$

- ▶ Ciwangi:
  - ▶ Mô hình chi tiết, phản ánh nhiều yếu tố sản xuất.
  - ▶ Tuy nhiên cấu trúc phức tạp, khó diễn giải nhanh.
- ▶ Malausma:
  - ▶ Mô hình có độ phù hợp thống kê cao nhất.
  - ▶ Thể hiện rõ vai trò của các đầu vào sản xuất.
- ▶ Gunungwangi:
  - ▶ Mô hình gọn, ổn định, dễ trình bày.
  - ▶ Phù hợp cho so sánh và kết luận tổng quát.

**Cảm ơn mọi người  
đã lắng nghe!**