

# 整合氣候因子與時序模型之水稻產量預測

## Integration of Climate Factors and Time Series Models for Rice Yield Prediction

組別:A4\_1

研究生:林宇哲、徐竣緯、陳彥君

指導教授:林真如 博士



### 摘要

水稻產量預測對於糧食安全與農業決策具有重要意義。然而，氣候因子變化複雜且與產量之間存在高度非線性關係，使得傳統迴歸與統計模型在預測上常受限於精確度不足。

本研究引用一套「雙層式隨機混合預測框架 (Two-Stage Stochastic Hybrid Forecasting Framework)」，整合氣候因子與時序模型進行水稻產量預測。研究首先蒐集並整合2014至2023年之氣候相關因子(如降雨量、氣溫、日照時數)與歷史產量資料，比較極端梯度提升樹(XGBoost)、自迴歸整合移動平均模型(ARIMA)與隨機森林(RF)之效能，整合2014-2023年台灣稻米產量與氣象資料，建立多模型預測系統。資料來源包括農委會農糧署的總產量及種植面積，與中央氣象署的氣象觀測數據。以單位產量(公斤/公頃)為依變數，消除面積變動影響，準確反映氣候對生產力的實際效應。

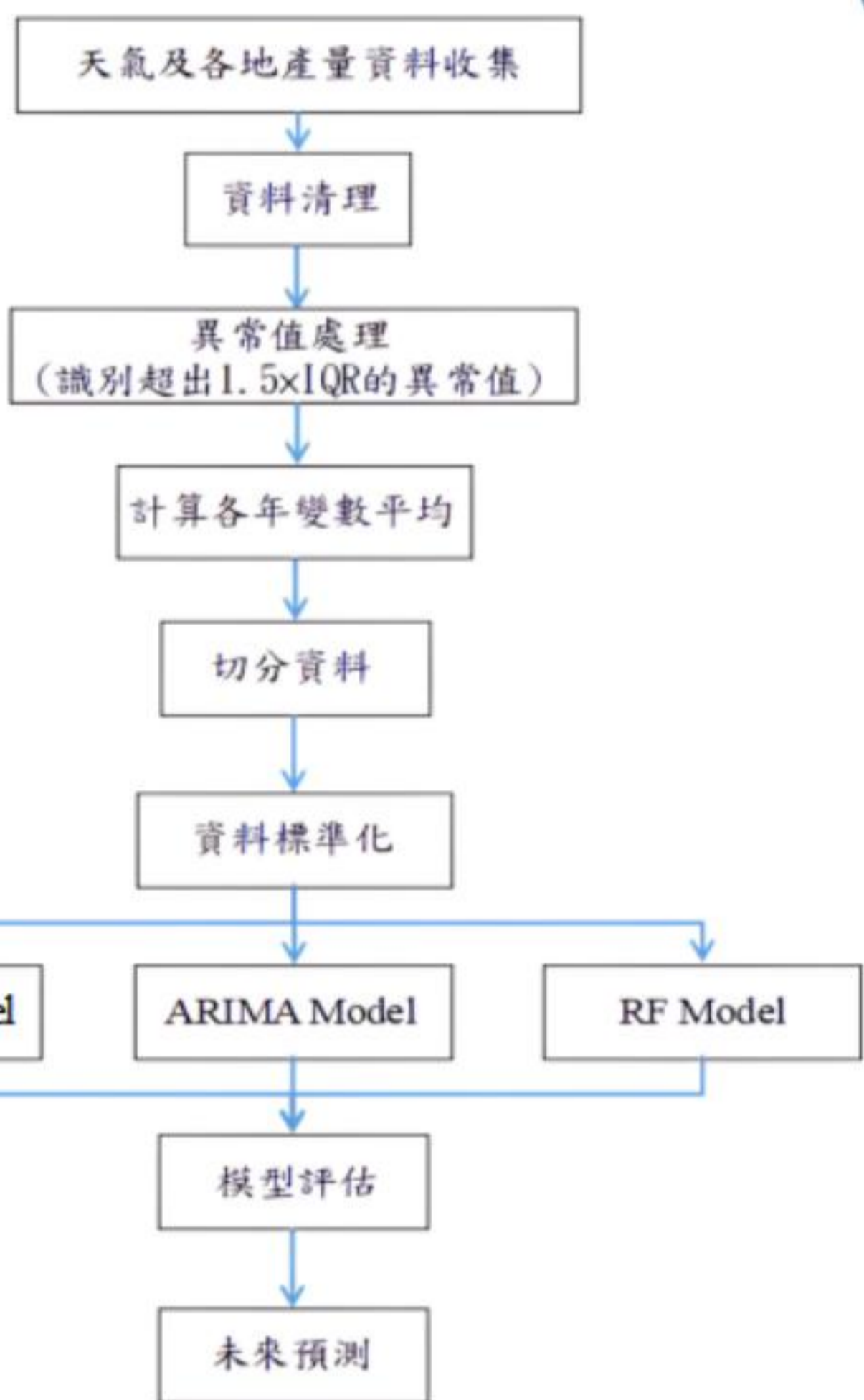
在未來預測方面，本研究針對2024至2028年導入隨機過程模擬生成氣候情境，並輸入模型進行推估。本研究結果驗證了混合預測框架的可行性，為未來智慧農業決策與糧食供應預警提供重要參考。

### 變數選擇

選取五個核心氣象因子作為自變數

1. 平均氣溫 (影響光合與呼吸作用)
2. 相對濕度 (影響蒸散與病蟲害)
3. 累積降雨量 (提供生長水分)
4. 日照時數 (影響光合效率)
5. 雨日數 (反映降雨分布)

### 研究流程與方法



### 實驗結果與分析

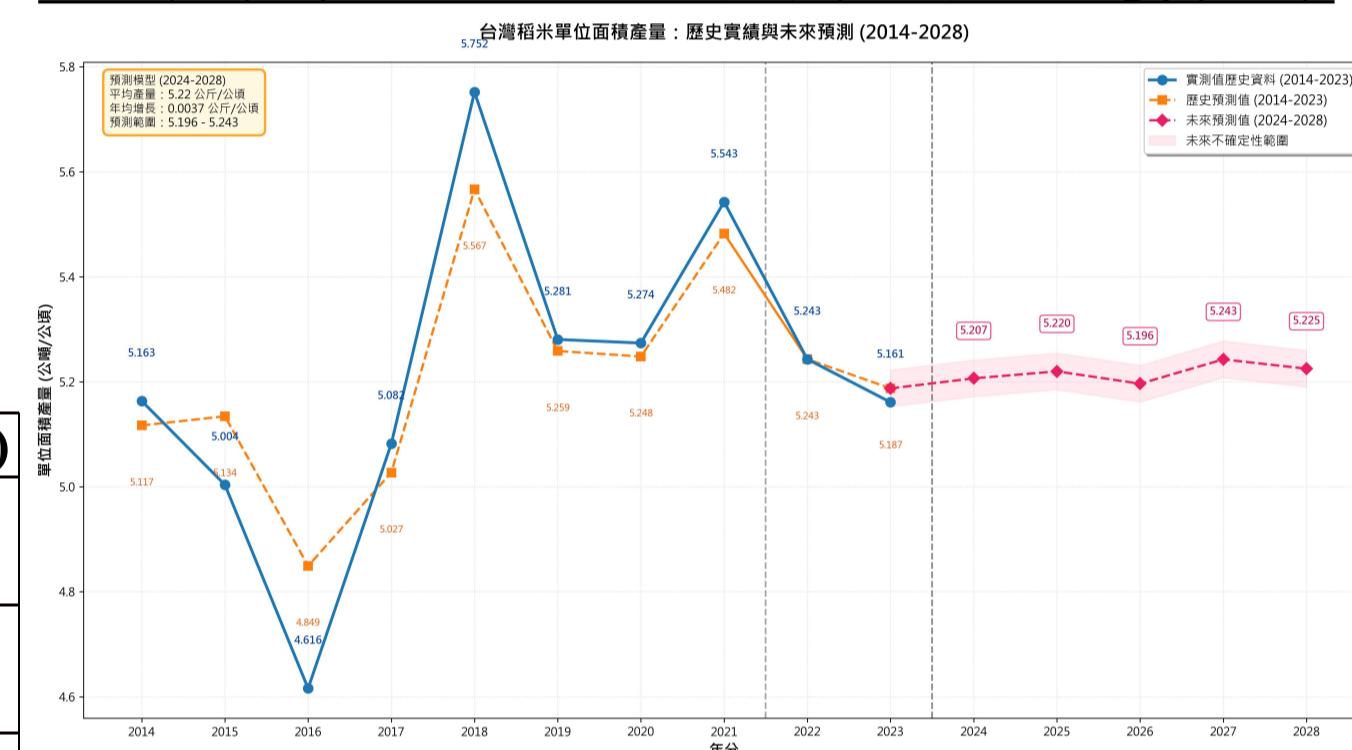
本研究整合2014-2023年台灣稻米產量與氣象資料，建立多模型預測系統。資料來源包括農委會農糧署的總產量及種植面積，與中央氣象署的氣象觀測數據。以單位產量(公斤/公頃)為依變數，消除面積變動影響，準確反映氣候對生產力的實際效應。

首先利用歷史數據進行模型訓練與驗證，透過分析預測值與實際值的誤差來評估模型準確度，確保模型具備足夠的泛化能力。隨後，針對未來產量的推估，本研究引入一套「雙層式隨機混合預測框架 (Two-Stage Stochastic Hybrid Forecasting Framework)」。

此框架分為兩個核心階段：第一階段針對氣候因子與種植面積等獨立變數，導入「動態模型競爭機制」與「隨機過程模擬」以生成未來情境；

第二階段則將上述模擬生成的特徵數據，輸入至已訓練完畢的三種模型(XGBoost、ARIMA、RF)，進行單位面積產量的非線性預測建模。首先利用歷史數據進行模型訓練與驗證，透過分析預測值與實際值的誤差來評估

### RF未來單位產量預測圖(包含歷史資訊)

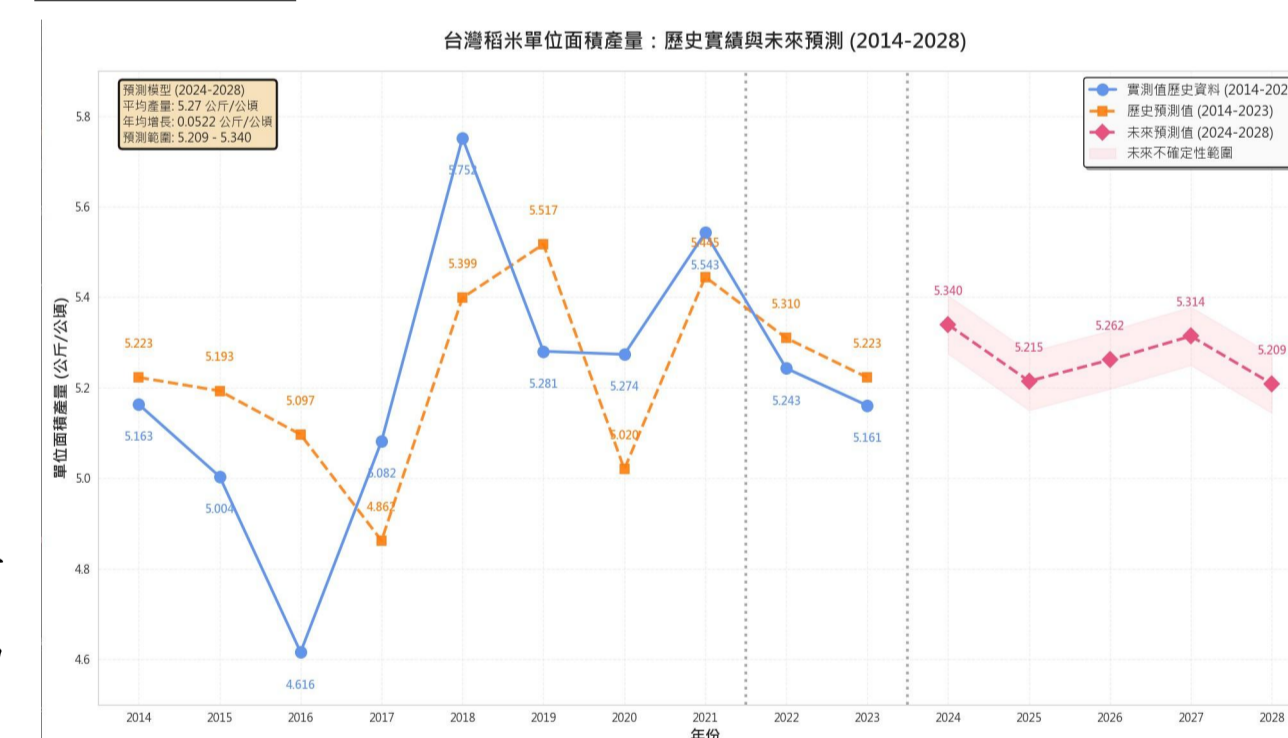


### 預測模型綜合評估指標

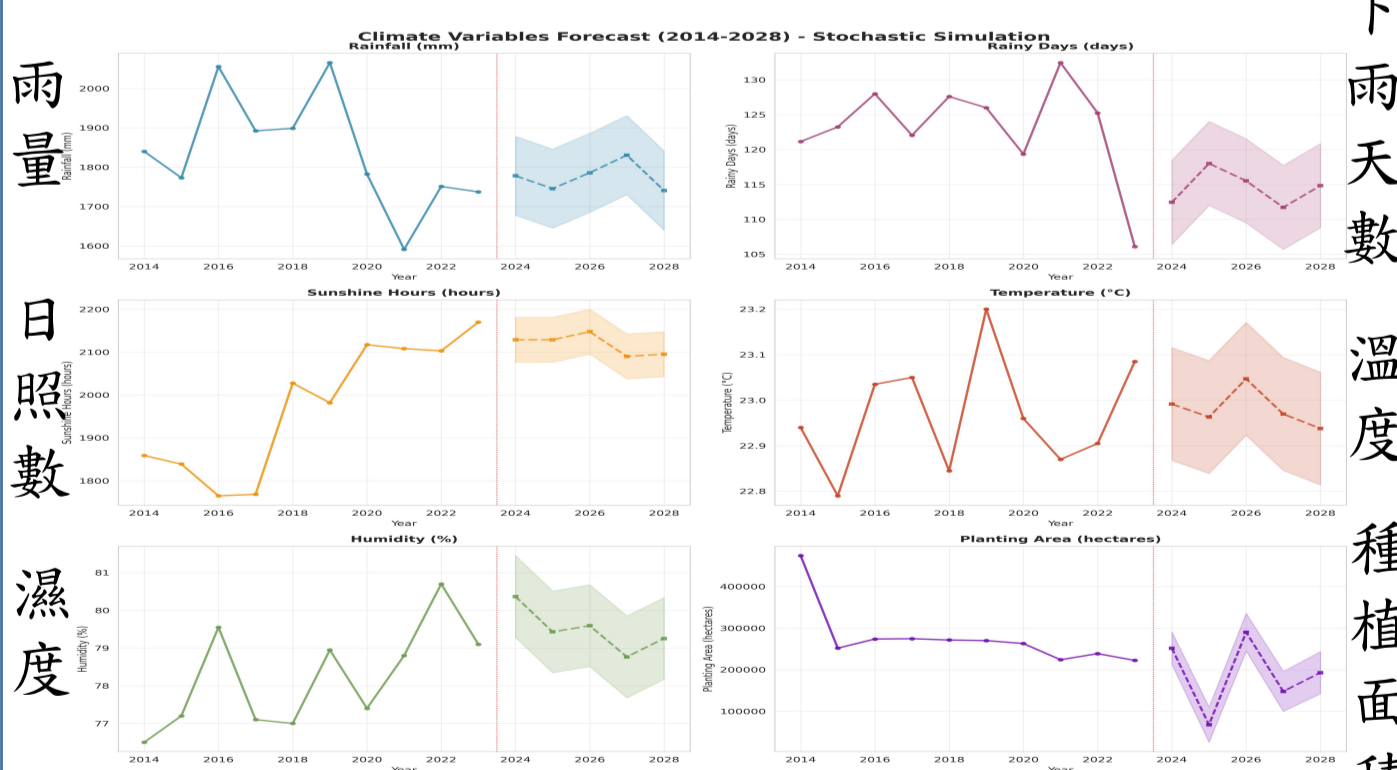
	MSE	RMSE	MAE	MAPE(%)
XGboost(訓練)	0.003	0.058	0.029	0.563
XGboost(測試)	0.014	0.117	0.117	2.218
ARIMA(訓練)	0.07	0.2677	0.2363	4.61
ARIMA(測試)	0.01	0.06	0.06	1.24
Random Forest(訓練)	0.01	0.119	0.083	1.63
Random Forest(測試)	0.01	0.094	0.080	1.56

- ARIMA模型在測試集表現最佳
- RF具有最佳泛化能力
- XGBoost在訓練集達到最高精度

### ARIMA未來單位產量預測圖(包含歷史資訊)

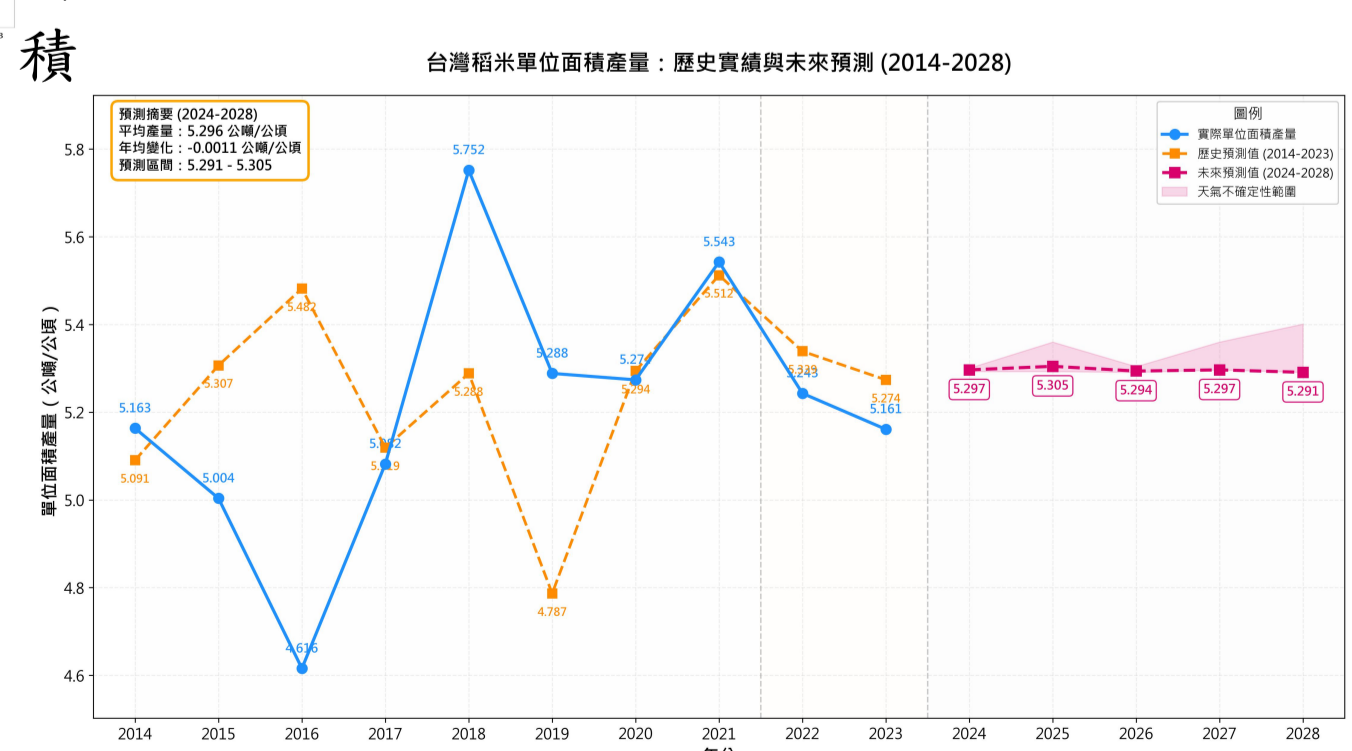


### 未來氣候預測圖



下雨天數  
溫度  
種植面積

### XGBoost未來單位產量預測圖(包含歷史資訊)



第一階段利用隨機過程生成具備波動特性的未來氣候情境；第二階段將此情境輸入模型進行非線性建模預測結果顯示，綜合而言，三種模型對未來五年的預測共識均值落在 5.2 至 5.3 公噸/公頃之間，預期未來五年不會出現結構性的產量異常上升或崩跌。

### 結論

儘管本研究成功量化了極端氣候(如日照不足)可能導致的產量下行風險邊界。成果驗證了混合預測框架的可行性，但仍受限於資料時間跨度僅 10 年的極小樣本特性，限制了深度學習模型的應用潛力。

此外，為求全台總體預測而進行的空間資料聚合，亦犧牲了縣市層級的氣候異質性資訊。未來研究建議可朝向整合衛星遙測數據作為外生變數，以彌補地面觀測站點的不足；或導入階層式預測方法，同步產出全台與各縣市的預測值。進而為未來智慧農業之風險管理與糧食供應預警提供了具體的科學依據。

### 實務應用隨機森林

訓練-測試一致性最佳，MAPE差距接近零，泛化能力優秀，預測最穩健可靠。適合需要穩定預測的實務場景。

### 短期預測ARIMA

測試集表現最優(MAPE=1.24%)，泛化能力最強。適合短期預測、小樣本情境，且計算效率高。

### 大數據XGBoost

訓練配適最強(MAPE=0.56%)，揭示日照為關鍵因子。小樣本易過擬合，但資料擴充後潛力最大。