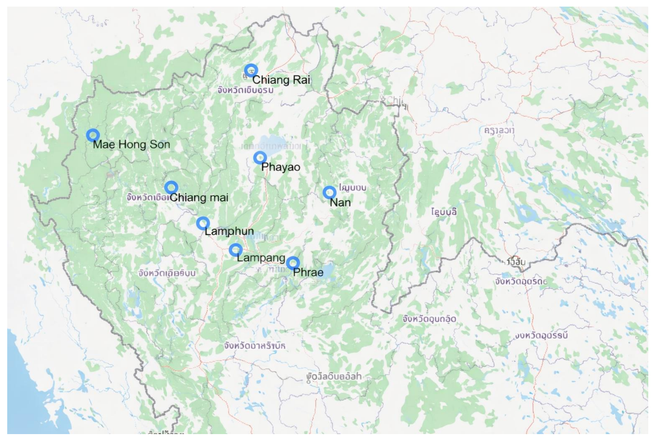
Modeling Rainfall and Temperature Variations in the Upper Northern Region of Thailand

Niphitpon Leunsai Advisor: Assoc. Prof. Dr. Manad Khamkong

FACULTY OF SCIENCE
CHIANG MAI UNIVERSITY

INTRODUCTION

Rainfall and temperature are important factors in water resource management, agriculture, and disaster prevention, especially in the upper northern region of Thailand, which is a major headwater area for several river basins in the country and a region that depends heavily on seasonal rainfall. However, climate variability and climate change have altered rainfall patterns in terms of timing, intensity, and consistency, affecting water security and agricultural productivity. Therefore, the researcher is interested in studying monthly rainfall forecasting in the upper northern region by comparing time-series models, namely SARIMA, SARIMAX, and ETS (Holt–Winters). Maximum and minimum temperatures are also included as external variables in the SARIMAX model to examine the effects of climatic factors on rainfall. The findings of this study can help identify the most appropriate forecasting model and support more effective planning for water resource management, disaster preparedness, and agricultural decision-making.



Objectives

1. To examine the trend and seasonality of monthly rainfall data using statistical analysis.
2. To identify the best forecasting model for monthly rainfall by comparing SARIMA, SARIMAX, and ETS (Holt-Winters), using sMAPE, MAE, and RMSE as the performance evaluation metrics.
3. To propose guidelines for water resource management and agricultural planning based on the forecasting models.

Scope of the Study

Study rainfall and temperature data in the upper northern region of Thailand.
Use data from 2017 to 2025 (B.E. 2560–2568).
Compare three forecasting models: SARIMA, SARIMAX, and ETS (Holt–Winters Exponential Smoothing).
Evaluate model performance based on forecasting results on the test set using sMAPE, MAE, and RMSE.

Materials

Time Series

SARIMA (Seasonal Autoregressive Integrated Moving Average)
Use the equation

$$\phi_p(B)\,\Phi_P(B^s)\,(1-B)^d(1-B^s)^D x_t = \delta + \theta_q(B)\,\Theta_Q(B^s)\,\varepsilon_t$$

SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Variables)
Use the equation

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D y_t = \delta + \boldsymbol{\beta}^{\top}\mathbf{X}_t + \theta_q(B)\Theta_Q(B^s)\varepsilon_t$$

ETS/Holt-Winters (Additive Holt–Winters)
Use the equation

Level $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$
Trend $b_t = \beta(\ell_t - \ell_{t-1}) + (1-\beta)b_{t-1}$
Seasonal (Additive) $s_t = \gamma(y_t - \ell_t) + (1-\gamma)s_{t-m}$
สมการพยากรณ์ล่วงหน้า hช่วง $\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$
โดย $k = \left\lfloor \frac{h-1}{m} \right\rfloor$

Mean Absolute Error (MAE)
Use the equation

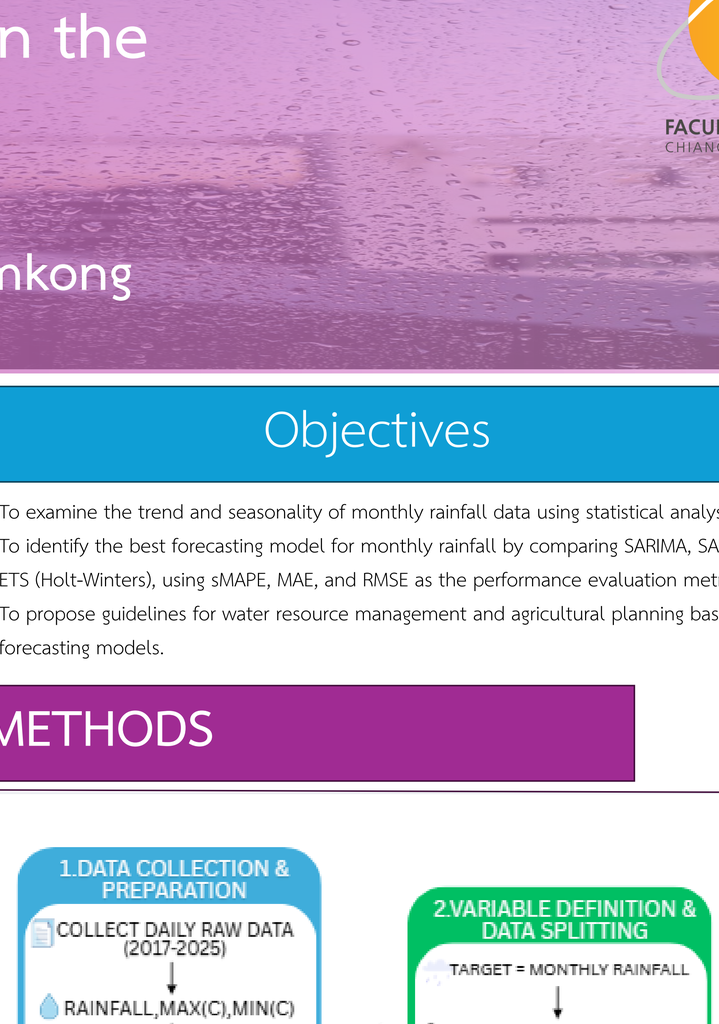
$$\mathbf{MAE} = \frac{1}{n}\sum_{t=1}^{n} |\, y_t - \hat{y}_t \,|$$

Root Mean Squared Error (RMSE)
Use the equation

$$\boldsymbol{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}$$

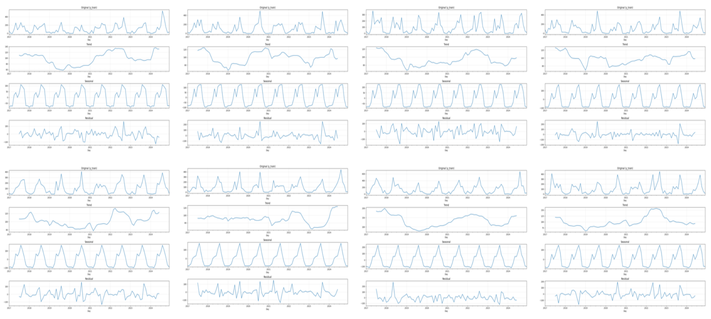
sMAPE (Symmetric Mean Absolute Percentage Error)
Use the equation

$$\mathrm{sMAPE} = \frac{100}{n}\sum_{t=1}^{n}\frac{2\,|\,y_t - \hat{y}_t\,|}{|\,y_t\,| + |\,\hat{y}_t\,|}$$



Results

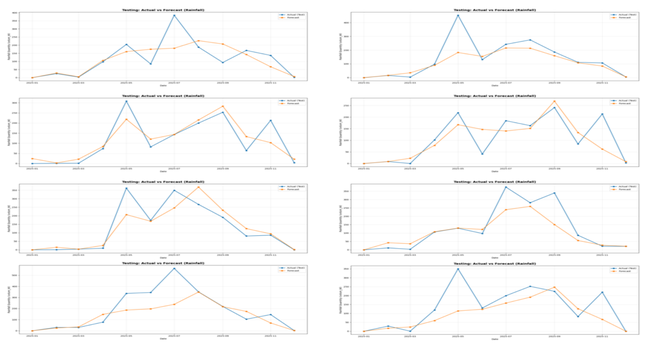
Trends and seasons



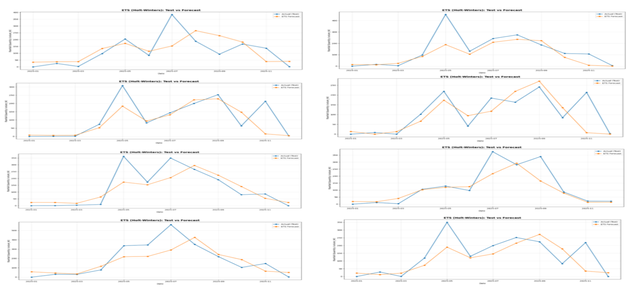
SARIMA (Seasonal Autoregressive Integrated Moving Average)



SARIMAX (Seasonal Autoregressive Integrated Moving Average with Exogenous Variables)



Holt-Winters



(From the figure, from left to right, the provinces are Phayao, Phrae, Lamphun, Lampang, Mae Hong Son, Nan, Chiang Rai, and Chiang Mai, respectively.)

Table sMAPE, MAE and RMSE

province	Mode	sMAPE(%)	MAE	RMSE
Phayao	SARIMA	81.3348	63.4366	93.3057
	SARIMAX	61.3752	52.3619	77.4150
	ETS (Holt-Winters)	83.1128	64.7520	88.8859
Phrae	SARIMA	67.5382	51.1319	88.0368
	SARIMAX	46.9099	43.9627	83.2922
	ETS (Holt-Winters)	64.6904	48.0926	84.0368
Lamphun	SARIMA	90.1956	48.5704	76.1183
	SARIMAX	72.6511	35.6395	49.1323
	ETS (Holt-Winters)	74.2806	43.1529	72.8062
Lampang	SARIMA	85.6397	51.9661	74.1126
	SARIMAX	71.1633	45.1826	61.6179
	ETS (Holt-Winters)	79.5063	49.6000	69.7675
Mae Hong Son	SARIMA	84.3561	50.5668	73.9832
	SARIMAX	68.5825	42.1848	64.3819
	ETS (Holt-Winters)	83.5020	49.2885	78.1529
Nan	SARIMA	64.3619	39.9405	72.133
	SARIMAX	58.7352	39.7320	65.5790
	ETS (Holt-Winters)	55.8234	38.9662	69.5836
Chiang Rai	SARIMA	76.1346	81.4339	123.9908
	SARIMAX	51.6113	81.4339	123.9908
	ETS (Holt-Winters)	65.8028	78.1777	103.9248
Chiang Mai	SARIMA	120.7714	82.0915	110.0286
	SARIMAX	69.1976	59.9899	88.7669
	ETS (Holt-Winters)	89.4730	59.8178	81.1892

discussion

Trend and seasonality were identified from the results of the Runs Test, the Kruskal–Wallis test, and the Seasonal Decomposition plots. The trend component showed periods of increase and decrease, while the seasonal component displayed a repeating pattern every year, indicating clear seasonality in the data.
According to the evaluation results, SARIMAX performed best in almost all provinces when primarily considering sMAPE. The provinces where SARIMAX gave the best results were Phayao, Phrae, Lamphun, Lampang, Mae Hong Son, Chiang Rai, and Chiang Mai. However, for Nan Province, Holt–Winters performed better. Therefore, SARIMAX should be selected as the first choice, Holt–Winters as the second choice, and SARIMA as the last choice in terms of sMAPE performance.
This suggests that including external variables, such as maximum temperature (Max °C) and minimum temperature (Min °C), helps improve forecasting accuracy in many provinces. Nevertheless, there are still some provinces where Holt–Winters provides better results.

Conclude

Time-series rainfall data for all provinces show clear seasonality, and an appropriate model for forecasting monthly rainfall in the upper northern region must also take into account the high variability of the data. The comparison results indicate that SARIMAX is the primary model that provides the best overall forecasting accuracy, while Holt–Winters can be a good alternative in some provinces where the data pattern is more suitable for smoothing methods.
The findings of this study can be applied to support water resource management planning. For example, in months when high rainfall is expected, drainage plans should be prepared in advance, whereas in months when low rainfall is forecast, water consumption should be reduced. Such applications are part of preparedness measures for both droughts and floods. In the agricultural sector, understanding clear seasonal patterns or anticipating whether rainfall will be above or below normal can help reduce crop damage. For instance, whether rainfall arrives in season or out of season, planting dates can be adjusted to align with the forecast in order to prevent losses in agricultural production.

REFERENCES

- Samsulaiman, N. H., & Ismail, N. (2023). Rainfall modelling and forecasting using Seasonal Autoregressive Integrated Moving Average (SARIMA) and exponential smoothing in Pahang and Terengganu, Malaysia. Proceedings of Science and Mathematics, 18, 1-10
- Lee, S., & Kim, S. U. (2016). Comparison between change point detection methods with synthetic rainfall data and application in South Korea. KSCE Journal of Civil Engineering, 20(4), 1558-1571. https://doi.org/10.1007/s12205-014-0228-5