Title

Model Forecasting of the number of Dengue Fever Diseases in Trang Thailand during the El Niño phenomenon.

Student name

Mr. Thanaphat Parichatnon

Mr. Thanakorn Phetbua

Teacher name

Mrs. Patchara Pongmanawut

Miss Apasri Chumchuen

School name

Princess Chulabhorn Science High School Trang

country

Thailand

date

5 march 2024

Abstract:

Over the past 50 years, the world has faced a changing climate. and is becoming more and more intense as a result of various phenomena that occur, such as the El Niño phenomenon, a phenomenon that has a huge impact on the public health of countries in Southeast Asia. Due to the increase in air temperature and a decrease in rainfall on the continent. This causes various diseases, such as an increase in the incidence of dengue fever. The researcher therefore studied the occurrence of the El Niño phenomenon and the incidence of dengue fever in Trang Province between 1998 and 2022. It was found that in the years that the El Niño phenomenon occurred, the rate The incidence of dengue fever in Trang Province has increased. When studying weather data, including rainfall, temperature, relative humidity, air pressure, and wind speed. Then analyzed the relationship with the number of dengue fever patients with the Pearson's Correlation Coefficient using the SPSS program. It was found that the average air pressure And the maximum air pressure value has a negative relationship with the number of dengue fever patients at a statistical significance of .05 and has created a model to predict the number of dengue fever patients using the Deep Learning model from ARIMAX (1,1 ,1) It was found that the number of dengue fever patients increased and decreased at a significant interval according to climate change, starting to decrease in 1998 to 1999, after which it increased from 1999 to 2012, and It will decrease with a very steep slope in 2012-2019 and finally increase again in 2019-2024. From data collection, the number of dengue fever cases corresponds to the occurrence of the El Niño situation. Therefore, it can be predicted that In 2025-2027, the number of dengue fever cases will decrease and will increase again in 2027-2028 due to a new wave of El Niño phenomenon

Keywords: *El Nino; Dengue fever (dengue fever); ARIMAX(1,1,1)*

Research questions:

1. Does the occurrence of the El Niño phenomenon increase the incidence of dengue fever in Trang Province?

2. What weather factors are related to the incidence of dengue fever in Trang Province?

3. Can the Deep Learning model from ARIMAX (1,1,1) accurately predict the number of dengue fever cases in Trang Province in the future? How?

Hypotheses:

1. The occurrence of the El Niño phenomenon causes the rate The incidence of dengue fever in Trang Province has increased.

2. Highest air pressure. and average air pressure It has the greatest impact on the incidence of dengue fever in Trang Province.

3. The Deep Learning model from ARIMAX (1,1,1) can accurately predict the number of dengue fever cases in Trang Province in the future.

Introduction and Review of Literature:

Over the past 50 years, the world has faced a changing climate. and is becoming more and more intense as a result of phenomena such as global warming, glass gas, acid rain, or even the El Niño phenomenon. This change in weather conditions affects various areas of the world in both ways. agriculture economic conditions Including medical and public health. The occurrence of various epidemics in certain areas may be affected differently by these phenomena. Although global climate change may cause different impacts in some areas, And the El Niño phenomenon is a phenomenon that has a huge negative impact on public health in countries in Southeast Asia. This is due to an increase in air temperature and a decrease in rainfall on the continent. This causes various diseases, such as the increase in the incidence of dengue fever in Trang Province over the past 15 years as a result of the increased abundance of Aedes aegypti and Aedes albopitus mosquitoes. (Ministry of Health, 2022)

The World Health Organization (WHO) states that an average of 390 million people worldwide are infected with the dengue virus that causes dengue fever each year. and has a death rate of 2.5% per year (World Health Organization (WHO), 2022) in Indonesia. Malaysia, including the southern region of Thailand, are the most directly affected areas. From past statistics, when the El Niño phenomenon occurred in that year, there would be an increased epidemic of dengue fever as a result. And research studies have found that the El Niño phenomenon has resulted in higher average air temperatures and decreased rainfall in Southeast Asia. Directly affecting the abundance of mosquitoes in Southeast Asia. The changing weather conditions cause mosquitoes to be more abundant due to their suitability for living and a shorter life cycle (Climate factors predict future mosquito activity, 2022), leading to a higher incidence of dengue fever in Trang Province has a significantly increased number of Aedes aegypti and Andes albopictus mosquitoes as vectors.

Therefore, the researcher has an idea to study the El Niño phenomenon and the incidence of dengue fever in Trang Province. Analyze the relationship between weather data such as rainfall, temperature, relative humidity, wind speed, air pressure and the rate of dengue fever cases using the Pearson's Correlation

Coefficient using the SPSS program and using the moving average model. Integrated Autoregression Model (ARIMAX) for predictive analysis of the number of dengue fever cases in Trang Province.

Materials – Equipment:

1. Digital thermometer
2. Digital hygrometer
3. Rain Gauge
4. Anemometer
5. SPSS program
6. Google Colab program

Method of operation:

1) Set the study point. This research was carried out to survey the weather at 1 study point, Meteorological Station, Trang Province, coordinates 6.69640° N, 101.38297° E.

Picture 1: Meteorological Station Study Point, Trang Province.

2) Collection and collection of weather data. Go to the area to collect weather data including temperature, humidity, rainfall, air pressure and wind speed at 1 study point using methods consistent with GLOBE principles during the months of January - December 2023 as follows:

2.1 Air temperature measurement Using a digital thermometer By requiring the thermometer to read the temperature every noon.

2.2 Measuring the relative humidity in the air. Using a digital hygrometer which reads the relative humidity value directly from the device.

2.3 Measuring the amount of rain using a Rain gauge. Collects the amount of rain once a day (every 24 hours) at noon (Solar noon).

2.4 Fill in the data in the Data Entry.

3) Creating a Deep Learning model from ARIMAX (1,1,1)

3.1 Creating a patient forecast equation using ARIMAX (AutoRegressive Integrated Moving Average with exogenous inputs) requires time series data. and use exogenous variables to help make predictions. In this case, the exogenous variables are weather data. Before we can use ARIMAX, we need to check the consistency of the data. (stationarity) and must adjust the data regularly if necessary It is also necessary to specify the parameters for the ARIMAX model, namely the AR parameter (p), the MA parameter (q), and the degrees of differencing (d). As well as considering the appropriate exogenous variables, we begin by checking the uniformity of Number of patients information And then proceed to build the ARIMAX model:

3.1.1 Check the consistency of the number of patients data We will use the Augmented Dickey-Fuller (ADF) test to test the consistency of the number of patients data: The results from the Augmented Dickey-Fuller (ADF) test for patient count data is as follows: • Test Statistic: -2.686707 • p-value: 0.076367 where the p-value is greater than 0.05, which indicates that we cannot reject the null hypothesis that the data has a unit root and is not uniform. can Therefore, patient population data may require differencing before being used in ARIMAX modeling.

3.1.2 ARIMAX Modeling

Due to the small size of our data, this analysis is only a demonstration of the feasibility of using ARIMAX in Forecasting does not divide data into training and test sets. Instead, it will demonstrate modeling with all available data. We will choose to use external variables that have a significant statistical relationship with the number of patients. The previous analysis is Pressure(Mean). For this demonstration, we will use a simple selection of p, d, q parameters such as p=d=q=1 for simplicity. Perform ARIMAX model creation: The generated ARIMAX model has the following results:

• Coefficients for Pressure(Mean) is -92.2622 with a p-value of 0.058, indicating that this variable tends to have a statistically significant relationship with the number of patients. But it is outside the specified threshold (0.05).

• Parameters AR and MA have p-values less than 0.05, indicating a statistically significant relationship with the number of patients. Note:

• There is a warning about the stability of the coefficients. and the risk of using unstable variance values.

• The limited amount of data (only 9 observations) may affect the reliability of the model.

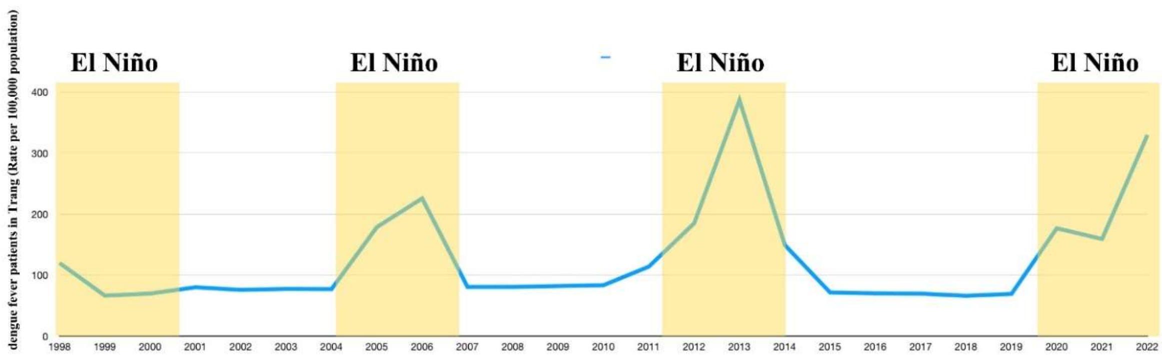
In conclusion, the ARIMAX model may be useful in predicting the number of patients. Considered from the average air pressure However, due to limitations on the size of the data and risks associated with model accuracy, Use of this model should be carried out with caution.

4) Analysis of research results

4.1 Pearson's Correlation Coefficient using the SPSS program.

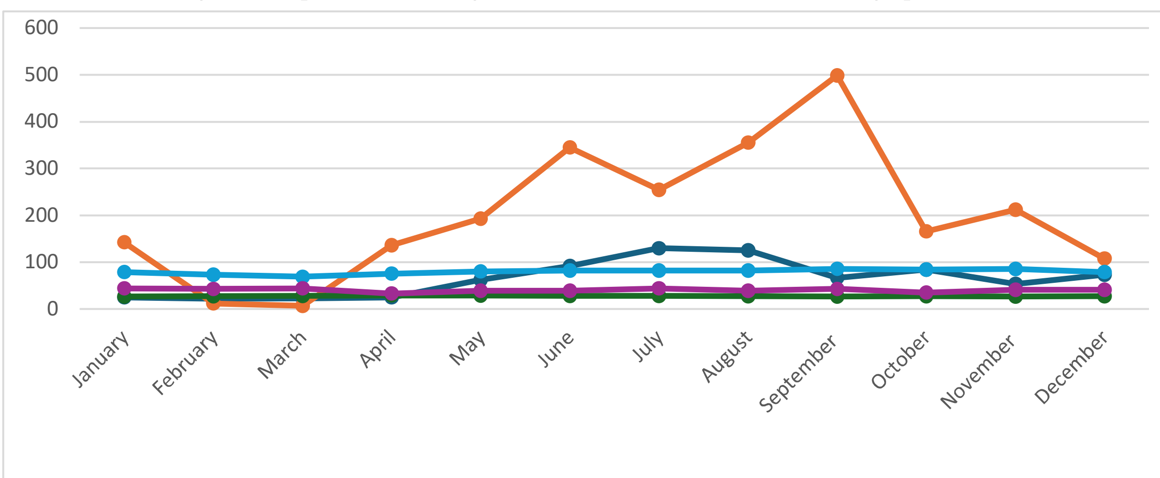
Result:

1. Results of the study of the occurrence of the El Niño phenomenon and the incidence of dengue fever. In Trang Province It is as in Graph



From the results of the study it was found that The occurrence of the El Niño phenomenon affects the number of dengue fever patients.

2. The results of weather studies include temperature, humidity, rainfall and wind speed. with the number of dengue fever patients of Trang Province in 2023 is as shown in the graph.



The dark blue line graph shows the number of patiens(people).

The orange line graph shows rainfall(millimeter) .

The green line graph shows temperature(degree celsius).

The blue line graph shows humidity(%).

The purple line graph shows wind speed (km/h).

When the data was analyzed for the relationship between weather data and the rate of dengue fever cases with the Pearson's Correlation Coefficient using the SPSS program, there was a value According to Table

Correlations

		MaxRain	MeanRain	MinRain	MaxPress	MeanPress	MinPress	MaxWin	MeanWind	MinWind	MaxHumidity	MeanHumidity	MinHumidity	MaxTemp	MeanTemp	MinTemp
Patients	Pearson Correlation	-.354	-.261	.022	-0.66	-.714*	.016	.510	.533	.543	-.111	.120	-.264	.537	.533	.533
	Sig. (2-tailed)	.349	.497	.955	.043	.031	.967	.160	.140	.131	.777	.758	.493	.136	.140	.140
	N	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9

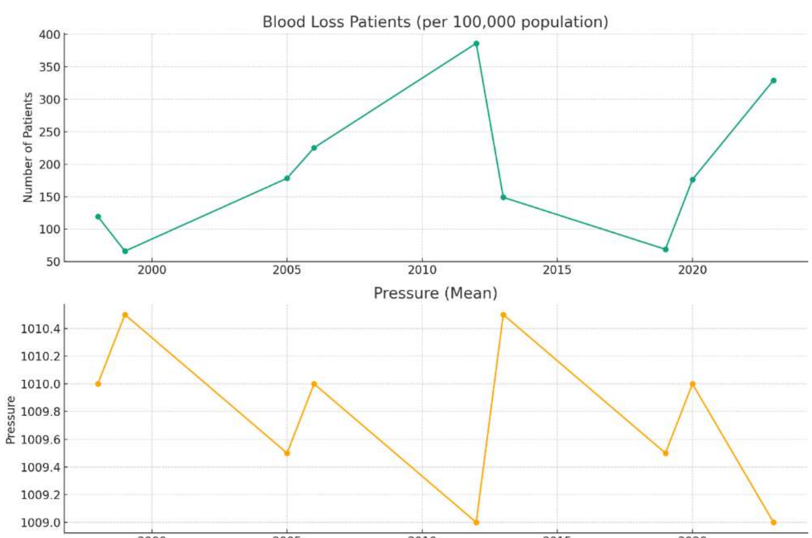
**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

from the results of the study it was found that Weather information includes average air pressure. And the highest air pressure has a negative relationship with the rate of dengue fever cases at the significance level of .05, that is, if the average air pressure is high, the rate of dengue fever cases is less.

3. Results of the study of creating equations to predict the number of patients. Using ARIMAX (AutoRegressive Integrated Moving Average with exogenous inputs)

3.1 Checking the consistency of the data We will start by checking the consistency of the time series data for number of patients and Pressure (Mean), and will make adjustments if necessary.



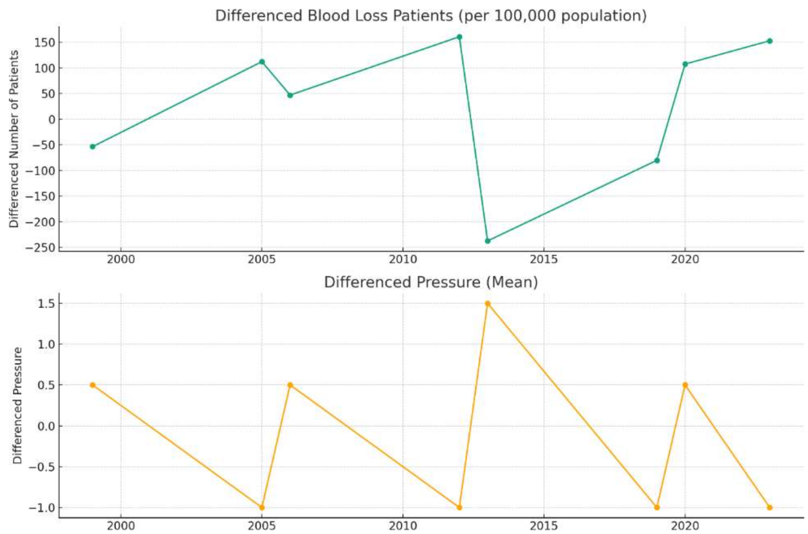
From the data plot and the Augmented Dickey-Fuller (ADF) test results: for the number of patients

the ADF statistic is -2.6867 and the p-value is 0.0764, which indicates that the data cannot reject the hypothesis of unit root. At the 95% confidence level, but close. This indicates that the data may need to be transformed to make it more consistent.

For Pressure (Mean), the ADF statistic is -2.2684 and the p-value is 0.1824, indicating that the data cannot reject the root hypothesis at any level. Confidence 95%

3.2 Data conversion

This is because the two datasets do not show consistency. We will be converting data to improve this feature. Possible conversions are using diff. (Differencing) to reduce the effects of trends and fluctuations that change over time. We will perform a one time difference and then run the ADF test again

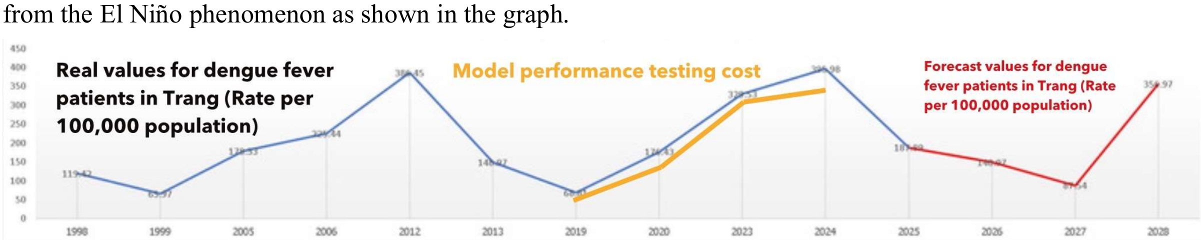


after one time difference:

• For the number of patients after the difference, the ADF statistic is -1.759 and the p-value is 0.401, which is still not The hypothesis of unit roots can be rejected at the 95% confidence level, showing that the data is still inconsistent even after divergence.

• For Pressure(Mean) after divergence, the ADF statistic is -8.914 and the p-value is very small (1.08×10—141.08×10—14), which can reject the hypothesis of having unit roots and show that the data is consistent after differences.

The graph shows the study of data from Model ARIMAX that shows the number of dengue fever cases in waves from 1998-2024 and predicts the number of dengue fever cases that will occur in 2025-2028 from the El Niño phenomenon as shown in the graph.



The blue line graph shows the number of people. Suffering from dengue fever Information from the Trang Provincial Public Health Office

The red line graph is the forecast for the number of dengue patients from the program from 2025 to 2028.

The yellow line graph is an estimate of the accuracy of the ARIMAX model to compare with the actual value in the period 2019-2023.

Summary and discussion:

The results of the study of the occurrence of the El Niño phenomenon and the incidence rate of dengue fever in Trang Province between 1998 and 2022 found that in the years that the El Niño phenomenon occurred, the incidence rate Dengue fever in Trang province is increasing. When studying weather data, including rainfall, temperature, relative humidity, air pressure, and wind speed. Then analyzed the relationship with the number of dengue fever patients with the Pearson's Correlation Coefficient using the SPSS program. It was found that the average air pressure And the maximum air pressure value has a negative relationship with the number of dengue fever patients at a statistical significance of .05 and has created a model to predict the number of dengue fever patients using the Deep Learning model from ARIMAX (1,1 ,1) It was found that the number of dengue fever patients increased and decreased at a significant interval according to climate change, starting to decrease in 1998 to 1999, after which it increased from 1999 to 2012, and It will decrease with a very steep slope in 2012-2019 and finally increase again in 2019-2024. From data collection, the number of dengue fever cases corresponds to the occurrence of the El Niño situation. Therefore, it can be predicted that In 2025-2027, the number of dengue fever cases will decrease and will increase again in 2027-2028 due to a new wave of El Niño phenomenon.

Bibliography/Citations:

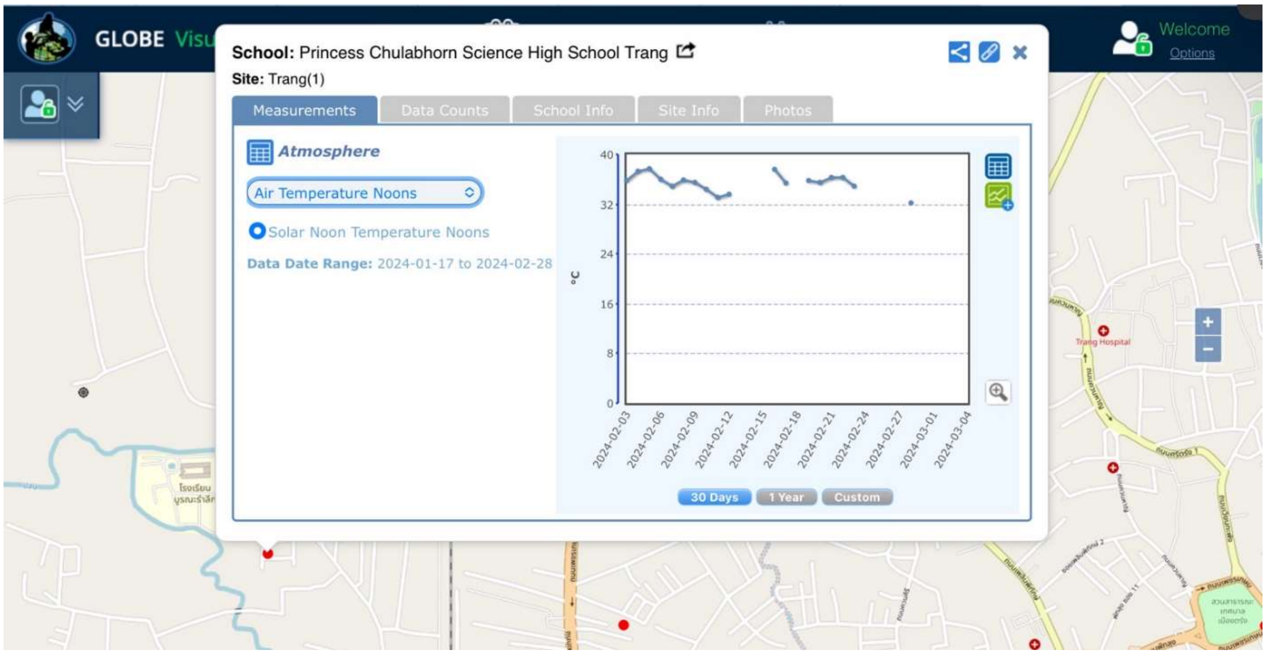
Ministry of Health (Myanmar). (2022). Coordination Meetings in 2022 for Preventing Dengue Outbreaks and Monitoring Monkeypox Epidemic 2022. From https://www.mohs.gov.mm/page/18212.

Ministry of Health (Myanmar). (2020) National Strategic Plan for Dengue Prevention and Control 2016-2020. Nay Pyi Daw: Vector Borne Disease Control Programme 2016. From http://mohs.gov.mm/ su/nzodH0.

World Health Organization (WHO). (2024). Dengue in the South-East Asia 2023. From https://www.who.int/southeastasia/health-topics/dengue- and-severe-dengue.

Submission of environmental measurement data in the GLOBE project database.

Weather data in GLOBE DATA ENTRY.



Badge Descriptions/Justifications:

I am a Data Scientist:

In this research, the researcher used skills and knowledge in statistics. Computer Science and Machine Learning to analyze huge amounts of data Find a relationship and build a model to predict the future. In this case, use weather data from the Trang Meteorological Department and data on dengue fever patients in Trang Province. From the Department of Epidemiology of Trang, going back 25 years (1998 - 2023) to develop a prediction model for dengue fever patients. Creating an ARIMAX model to predict future patient numbers This result can help public health agencies in allocating resources and planning disease prevention.

I make an Impact:

This research is a model developed that can predict the number of dengue fever cases in Trang Province 5 years in advance. This information helps people. Formulate public health policies to efficiently allocate resources and prepare to deal with disease outbreaks. Helps reduce the number of patients and the severity of the disease. My research helps public health agencies to Predict the number of dengue fever patients in the future. Allocate resources efficiently Plan disease prevention and control Reduce the number of patients and deaths from dengue fever.

I am Collaborator:

In this research, all researchers are involved in collecting weather data. and dividing the work duties into each section, including weather data analysis modeling In addition, we have coordinated with the officials of the Meteorological Department, Trang Province, to request historical weather information. Public health officials in Trang Province provide information on the number of dengue fever patients.