

# Quantifying the Status of Galamsey In Ghana With Time Series Classification

Kalong Boniface  
Fugah Seletay Mitchell

May 1, 2022



**University Of Energy and Natural Resource, Sunyani**  
**Department of Mathematics and Statistics**

# 1 Introduction

The purpose of this paper is to establish an understanding of time series analysis on remotely sensed data. Which will introduced us to the fundamentals of time series modeling, including decomposition, autocorrelation and modeling historical changes in Galamsey Operation in Ghana, the Cause,Dangers and it's Environmental Impact.

Galamsey("gather them and sell")(OwusuNimo2018), is the term given by local Ghanaian for illegal small-scale gold mining in Ghana(DavidYawDanquah2019). The major cause of Galamsey is unemployment among the youth in Ghana(Gracia2018). Young university graduates rarely find work, and when they do it hardly sustains them. The result is that these youth go the extra mile to earn a living for themselves and their family. Another factor is that lack of job security.

On November 13, 2009 a collapse occurred in an illegal, privately owned mine in Dompoase, in the Ashanti Region of Ghana. At least 18 workers were killed, including 13 women, who worked as porters for the miners. Officials described the disaster as the worst mine collapse in Ghanaian history(News2009).

Illegal mining damages the land and water supply(Ansah2017). In March 2017, the Minister of Lands and Natural Resources, **Mr. John Peter Amewu**, gave the galamsey operators/illegal miners a three-week ultimatum to stop their activities or be prepared to face the law(Allotey2017). The activities by galamseymen have depleted Ghana's forest cover and they have caused water pollution, due to the crude and unregulated nature of the mining process(Gyekye2021).

Under current Ghanaian constitution, it is illegal to operate as galamseymen. That is to dig on land granted to mining companies as concessions or licenses and any other land in search for gold. In some cases, Galamseymen are the first to discover and work extensive gold deposits before mining companies find out and take over. Galamseymen are the main indicator of the presence of gold in free metallic dust form or they process oxide or sulfide gold ore using liquid mercury.

Between 20,000 to 50,000, including thousands from China are believed to be engaged in Galamsey in Ghana. But according to the Information Minister 200,000 and nearly 3 million people, recently are now into Galamsey operation and rely on it for their livelihoods(Burrows2017). Their operations are mostly in the southern part of Ghana where it is believe to have substantial reserves of gold deposits, usually within the area of large mining companies(Barenblitt2021). As a group, they are economically disadvantaged. Galamsey settlements are usually poorer than neighboring agricultural villages. They have high rates of accidents and are exposed to mercury poisoning from their crude processing methods. Many women are among the workers, acting mostly as porters for the miners.

## 2 Problem Statement

The Footprint of Galamsey is Spreading at a much faster rate.

## 3 Objectives

### 3.1 Overview

The purpose is to establish an understanding in time series analysis on remotely sensed data. We will be introduced to the fundamentals of time series modeling, including decomposition, autocorrelation and modeling historical changes.

- Perform time series analysis on satellite derived vegetation indices
- Gain familiarity with R Markdown, Reticulate and the R Spatial Ecosystem
- Process satellite imagery using the Google Earth Engine API
- Create a Statistical interactive dashboard

## 4 Data collection and Methodology

### 4.1 Data

As Galamsey is considered an illegal activity, their operations are hidden to the eyes of the authorities. So locating them is quite tricky, but with satellite imagery, it is now possible to locate their operations and put an end to it. One of the features of Google Earth Engine is the ability to access years of satellite imagery without needing to download, organize, store and process this information. For instance, within the Satellite image collection, now it is possible to access imagery back to the '90s, allowing us to look at areas of interest on the map to visualize and quantify how much things have changed over time. With Earth Engine, Google maintains the data and offers its computing power for processing. Users can now access hundreds of time series images and analyze changes across decades using *GIS and R* or another programming language to analyze these datasets.

## 4.2 Method

Time series data is the collection of observations made sequentially at different points in time. Because data points in time series are collected at adjacent time periods there is potential for correlation between observations. we propose some new tools to allow machine learning classifiers to cope with time series data. *We first argue that, time-series classification problems can be solved by detecting and combining local properties or patterns in time series.* Then, a technique is proposed to find patterns which are useful for classification. These patterns are combined to build interpretable classification rules. First, we will pull Sentinel 2 to select NDVI and EVI data from Google Earth Engine, applying a quality filter to mask poor quality pixels.

Instead of performing our analysis on the imagery itself, we will be summarizing the mean NDVI and EVI value, this will allow the analysis to take less time while producing a visually appealing and informative map.

Some cells may not contain NDVI and EVI for a given month, to correct this, we will apply smoothing method using an ARIMA function.

Once NA values are remove, we will decompose the time series to remove seasonality and fit a linear model to the normalized data.

Once we have extracted the linear trend, we will then make a move to classifier our data on the map and map it.

## 5 Significance Of the Study

## 6 Limitations

Time series modeling aims to build an explanatory model of the data without over fitting the problem set, to use as simple a model as possible while accounting for as much of the data as possible. When breaking down time series data into component parts, remote sensing data has additional limitations that make this more challenging. It is almost inevitable that you will not get this same level of precision from remote sensing data. Additionally, atmospheric conditions can skew the visual results, where the hue of the vegetation changes drastically from image to image due to atmospheric conditions such as (fog, ground moisture, cloud cover).