Lapai Journal of Applied and Natural Sciences
LAJANS Vol 1(1): 44 - 57

RESEARCH PAPER

Spatial Analysis of West African Measles Data Set

[1]Bolarin G., [2]Bello O.A. and [3]Raji R.A.

[1] *Department of Mathematics, Federal University of Technology, Minna, Niger State, Nigeria*
[2] *Department of Statistics, Federal University of Technology, Minna, Niger State, Nigeria*
[3] *Department of Mathematics /Statistics, Osun State Polytechnic, Iree, Osun State, Nigeria*

ABSTRACT

The spread of infectious diseases has been tied to movement of people across borders. Statistical data have shown that there are a lot of factors that can make a disease to break out in a given community or country. In this work we analyzed measles data set of West African countries and we found that there is a synchrony in outbreak of measles in countries sharing borders. Also, from our analysis, we discovered that vaccination has great effect on the number of reported cases.

Keywords:

INTRODUCTION

Infectious diseases had proven to shed more light on the concept of 'spatiotemporal'. Spatial transmission of directly transmitted infectious diseases is ultimately tied to the movement of the hosts. A metapopulation consists of a group of spatially separated populations of the same species which interact at some level. The term metapopulation was coined by Richard Levins in 1970 to describe a model of population dynamics of insect pests in agricultural fields, but the idea has been most broadly applied to species in naturally or artificially fragmented habitats. It consists of a population of populations. The network of spatial spread (the disease's spatial coupling) may therefore, be expected to be related to the pattern of human movement or traffic within the host metapopulation.

The outbreak of Measles is noted to be seasonal and it can be synchronous across the host metapolulation. Movements across borders have reached a very complex degree, to the extent that monitoring has become unrealistic. Most countries in Africa have varying visa policies, but a lot of them have policies that are favourable to other African citizens. The harmonized international passport programme introduced by ECOWAS has lead to great trans-border movements by citizens of member states. This is also true of most regional settings all over the world e.g. Southern African Development Commission, EU, etc.

This practice has resulted in transportation of various goods, humans, diseases etc. continuously across borders. As we know, measles does not need a physical contact before it can be transmitted from one individual to another that is why it is being regarded as one of the most contagious diseases in the world [www.wikipedia.com]. With this, it is very easy for people to 'transport' the disease across borders. Rubeola and rubella viruses are spread through the respiratory route. This means they are contagious through coughing and sneezing. In fact, rubeola virus is one of the most contagious viruses known to man [www.wikipedia.com]. As a result, it can spread rapidly in a susceptible population. Infected people carry the virus in their respiratory tract before they get sick, so they can spread the disease without being aware of it. More so, most West African countries do not have strict health policies when it comes to immigration within the sub-region. Measles dynamics have sparked a long history of data analysis and modelling across a range of disciplines. Mathematical epidemiologists have drawn by the importance of the disease and have dissected most of the main features of measles transmission within local communities (McLean et al. 1988; Mollison et al. 1993; Xia et al. 2004).

Epidemiologists had attempted to study a lot of scenario when it comes to spatial dynamics of

Received 19 October, 2016
Accepted 22 November, 2016
Address Correspondence to:
g.bolarin@futminna.edu.ng

measles. Bjørnstad *et al.* **(**2002**)** and Dietz (1976), in their studies seasonality in transmission was considered, while Earn *et al.* (2000), Finkensta¨dt and Grenfell (2000), Finkensta¨dt *et al***. (**1998**),** Grenfell et al. (2002), McLean and Anderson (1988) considered the host demography (birth rate and vaccination). Xai *et al.* (2004) related measles dynamics to consumer-resource dynamics by studying the spatiotemporal distribution of measles.

There are basically two types of transmissions when it comes to measles; the first one is 'local' transmission among individuals within an enclaves i.e., cities, towns etc., and the second one is transmission between enclaves (inter-cities, across countries etc.). In (Grenfell et al. 1997; Earn *et al.* 1998; Swinton, 19998) a relationship was established between the regional dynamics of infectious diseases and the dynamics of ecological met populations. Xai *et al.* (2004) underscored the importance of infection process at local level and how it affects the spatiotemporal pattern of epidemics.

The network of spatial spread (the disease's spatial coupling) may therefore be expected to be related to movement within the host metapopulation. Spatial coupling was initially assumed to be a simple inverse function of distance (Okubo, 1980). Again it has been noted that emigration and immigration rates depend on population size (Hanski, 1998). Also distance between enclaves/countries can be a simple but effective assumption in spatial analysis (Erlander, 1990).

As we all know, epidemiology study involves comparing, planning, implementing, evaluating, and optimizing various detection, prevention, therapy, and control programs. Epidemiological modelling can contribute to the design and analysis of epidemiological surveys, suggest important data that should be collected, identify trends, make general forecasts, and estimate the uncertainty in forecasts. So, in this work an attempt is made to analyze trends in measles data (like synchronicity, prevalence, etc) for West African countries. Infectious diseases dynamics in a single individual is fairly simple and well understood, so connecting models to data-set of infectious diseases is an interesting study area for epidemiologists and this is alluded to by Grenfell et al. (2002). The scope of this study is to perform a comprehensive analysis of measles data-set of West African subregion.

MATERIALS AND METHOD

Data Collection

The data-set used in this study was obtained from the WHO website [www.who.int]. The data captured number of reported cases of measles' outbreak for each country in the subregion from 1980 to 2015. The vaccination data captures the percentage of vaccinated individuals that are less than 1 year old in the population. Some of the data were missing (there were no data for some years in some of the countries under consideration) and so a code was written to take care of the missing data.

Detrending

Trend in a time series is a slow, gradual change in some property of the series over the whole interval under investigation.

Trend is sometimes loosely defined as a long term change in the mean, but can also refer to change in other statistical properties.

Detrending is the statistical or mathematical operation of removing trend from the series. Detrending is often applied to remove a feature thought to distort or obscure the relationships of interest.

Detrending is also sometimes used as a preprocessing step to prepare time series for analysis by methods that assume stationarity. Anderson [1], describes differencing as a way to remove nonstationarity from time series in general.

RESULTS AND INTERPRETATION

At this point, there is a need for a pictorial representation of the data, so the number of reported cases was plotted (frequency) against the year to see the trend in the data and how each country is faring from year to year as campaign towards eradication of measles is being intensified.

In addition to the graphical representation of reported cases, a graph of the prevalence is also presented.

Prevalence is the measure of how commonly a disease or condition occurs in a population. Prevalence measures how much of some disease or condition there is in a population at a particular point in time. It is calculated by dividing the number of persons with the disease or condition at a particular point by the number of individuals examined.

The graphs of the incidence rate/number of reported cases for various counties are presented in Figures 1-6 and the prevalence profiles are provide in Figures 7-10.

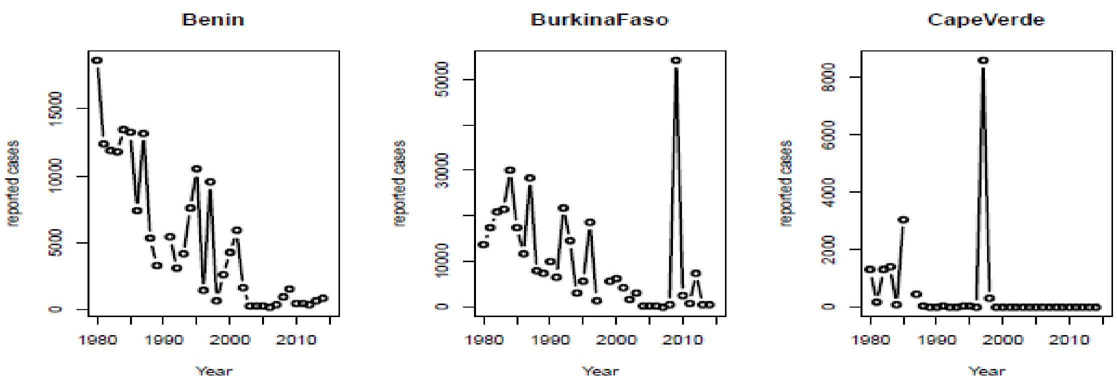

Figure 1: Number of measles reported cases in Benin Republic, Burkina Faso and Cape Verde between 1980 and 2015. (The break in the lines is as a result of missing data from the data-set*)*

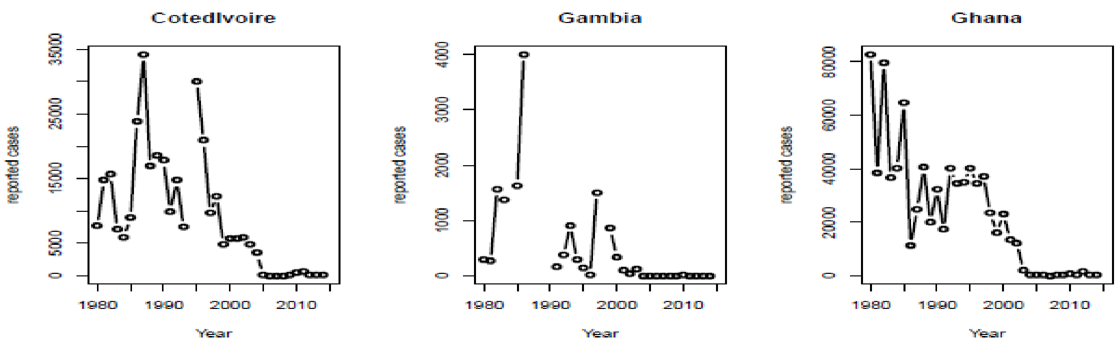

Figure 2: Number of measles reported cases in Cote d'Ivoire, Gambia and Ghana between 1980 and 2015. (The break in the lines for Cote d'Ivoire is as a result of missing data due to war while the relative low and missing data recorded for Gambia is due largely to political instability like coup.

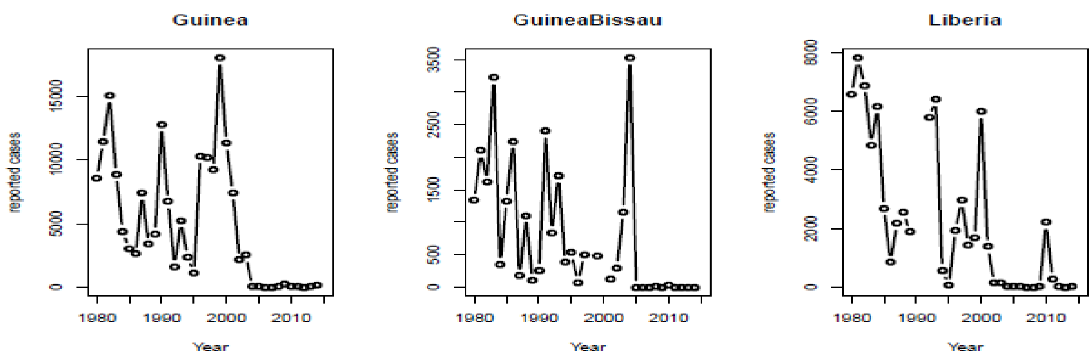

Figure 3: Number of measles reported cases in Guinea, Guinea-Bissau and Liberia between 1980 and 2015. (The missing data record in Guinea-Bissau was due to Civil war triggered by army uprising. Also the missing data recorded in 1990 and 1991 in Liberia was due to Civil wars)

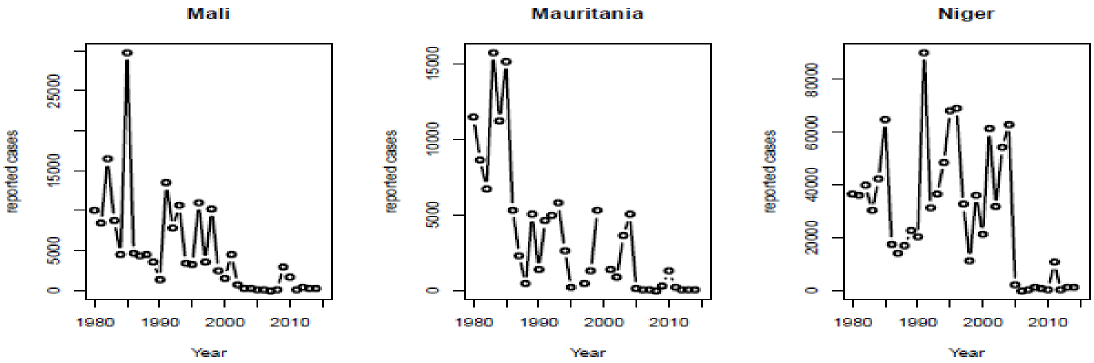

Figure 4: Number of measles reported cases in Mali, Mauritania and Niger between 1980 and 2015. (Niger is known for high rate of measles outbreak)

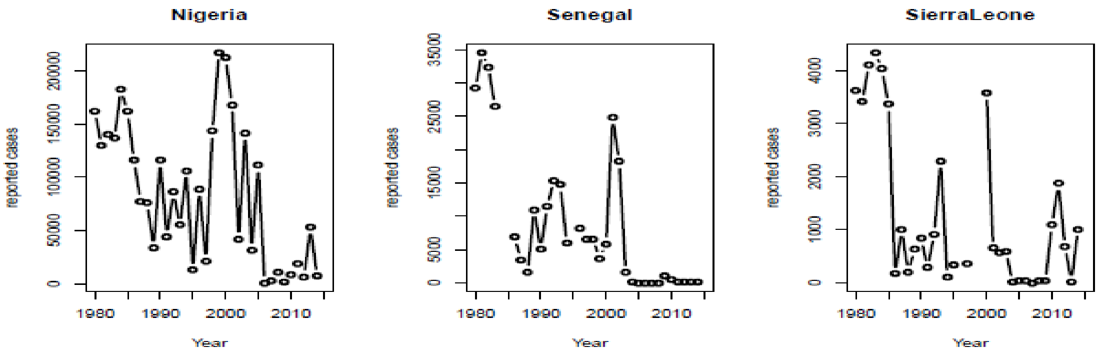

Figure 5: Number of measles reported cases in Nigeria, Senegal and Sierra Leone between 1980 and 2015. (Nigeria remain the leading country when it comes to measles outbreak, this is due to the large population of Nigeria; the low cases reported in Senegal was due to war between government and two separatists and the missing data in Sierra Leone was due to extension of Liberia's war into Sierra Leone, which lead to brutal civil war for 10 years.)

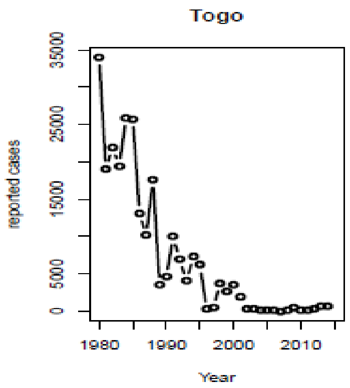

Figure 6: Number of measles reported cases in Benin Republic, Burkina Faso and Cape Verde between 1980 and 2015. (The break in the lines is as a result of missing data from the data-set)

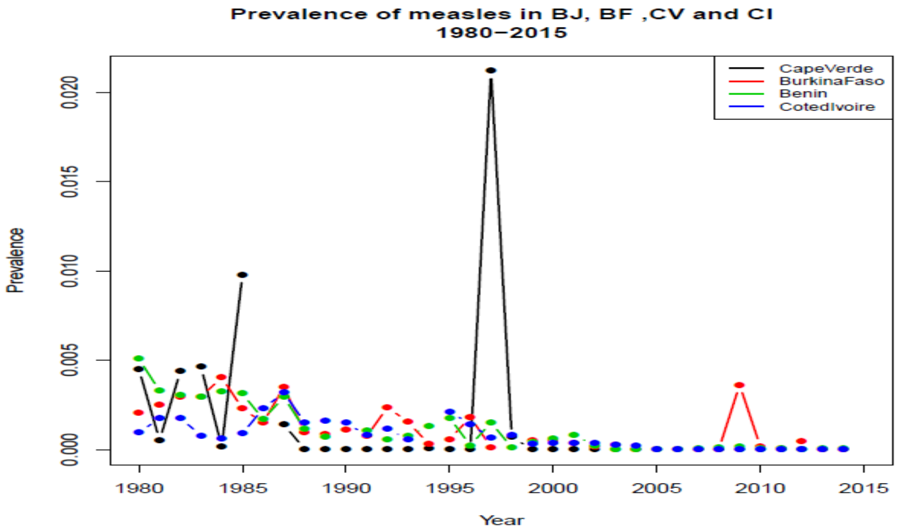

Figure 7: Prevalence of measles in Benin, Burkina Faso, Cape Verde and Cote d'Ivoire with Cape Verde reporting highest level of prevalence at some point in the time series.

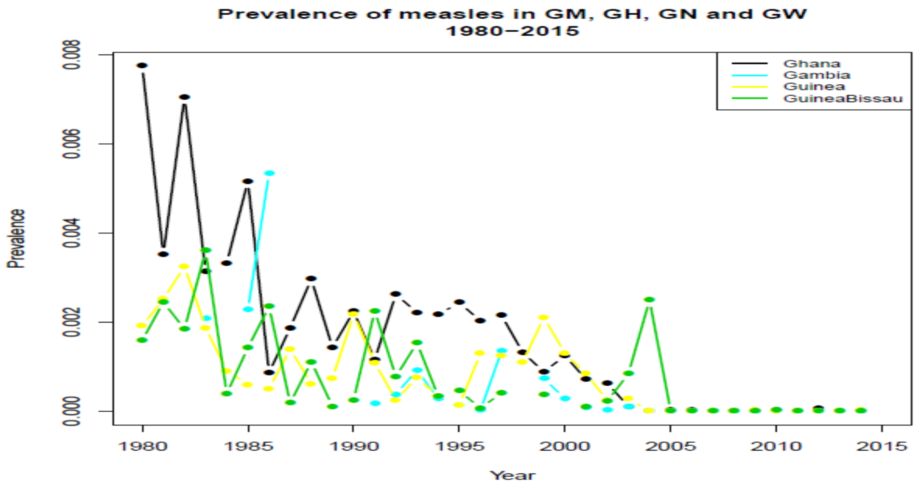

Figure 8: Prevalence of measles in Gambia, Ghana, Guinea and Guinea-Bissau with Ghana reporting highest level of prevalence at some point in the time series.

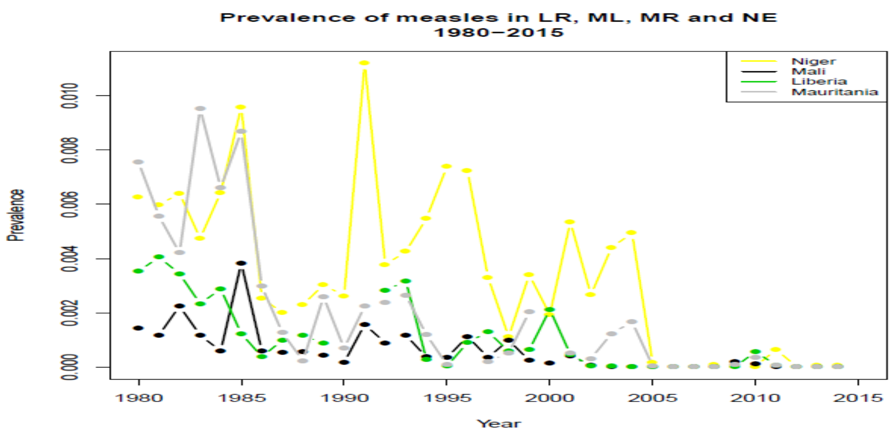

Figure 9: Prevalence of measles in Liberia, Mali, Mauritania and Niger with Niger reporting highest level of prevalence at some point in the time-series and also with the highest reported cases in this segment.

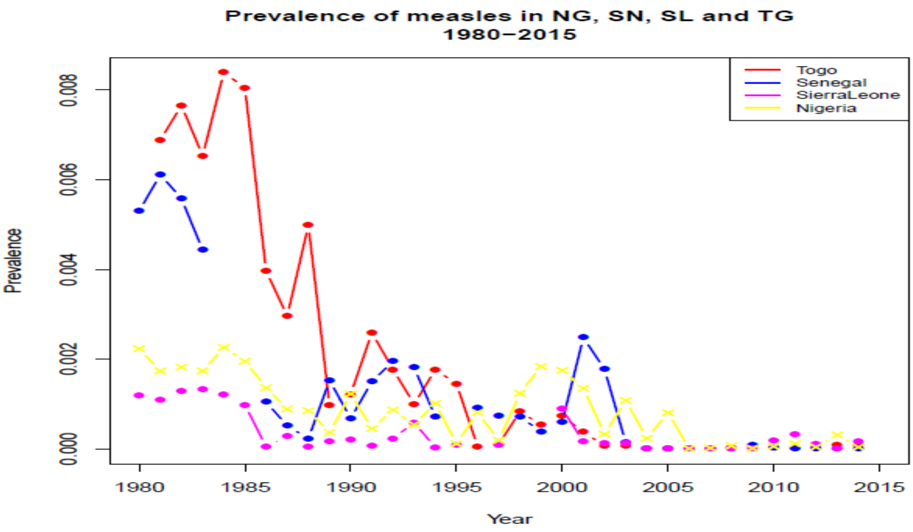

Figure 10: Prevalence of measles in Nigeria, Senegal, Sierra Leone and Togo, though Nigeria has the highest reported cases, Togo has the highest prevalence rate at a time.

Loess

LOESS denotes a method that is also known as locally weighted polynomial regression. At each point in the data-set (comprising m points), a low-degree polynomial is fitted to a subset of the data, with explanatory variable values near the point whose response is being estimated. The polynomial is fitted using weighted least squares, giving more weight to points near the point which response is being estimated and less weight to points further away. The value of the regression function for the point is then obtained by evaluating the local polynomial using the explanatory variable values for that data point. The LOESS fit is complete after regression function values have been computed for each of the m data points. Many of the details of this method, such as the degree of the polynomial model and the weights are flexible.

The subsets of data used for each weighted least squares fit in LOESS are determined by a nearest neighbour algorithm. A user-specified input to the procedure called the "bandwidth" or "smoothen parameter" determines how much of the data is used to fit each local polynomial. The smoothen parameter, α , is a number between $(\lambda+1)/m$ and 1, with λ denoting the degree of the local polynomial. The value of α is the proportion of data used in each fit. The subset of data used in each weighted least squares fit comprises the $m\alpha$ (rounded to the next largest integer) points whose explanatory variable values are closest to the point at which the response is being estimated.

α is called the smoothen parameter because it controls the flexibility of the LOESS regression function.

The local polynomials fitted to each subset of the data are almost always of first or second degree; that is, either locally linear (in the straight line sense) or locally quadratic. Using a zero degree polynomial turns LOESS into a weighted moving average (www.wikipedia.com).

The use of the weights is based on the idea that points near each other in the explanatory variable space are more likely to be related to each other in a simple way than points that are further apart. Following this logic, points that are likely to follow the local model influence the local model parameter estimates the most. Points that are less likely to actually conform to the local model have less influence on the local model parameter estimates.

The traditional weight function used for LOESS is the tri-cube weight function which is given as:

$$w(x)=(1-|x|^3)^3 I[|x|<1].$$

In this work the LOESS fit per country is presented. Figures 11a-11e show the LEOSS fits.

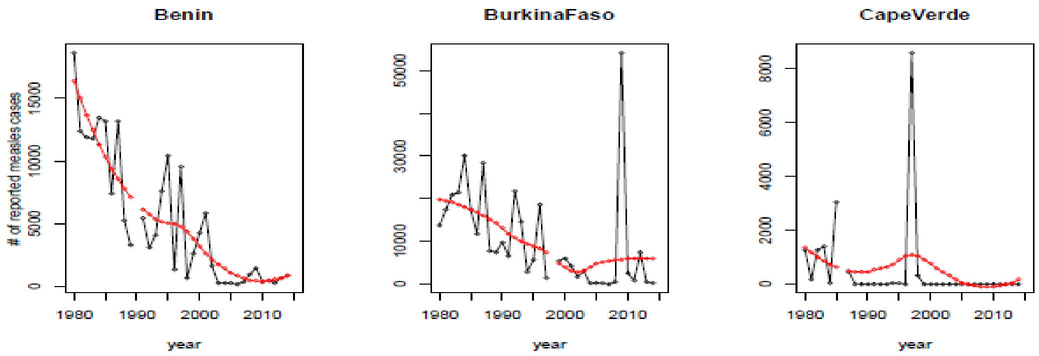

Figure 11a: Loess Fit for Benin, Burkina Faso and Cape Verde.

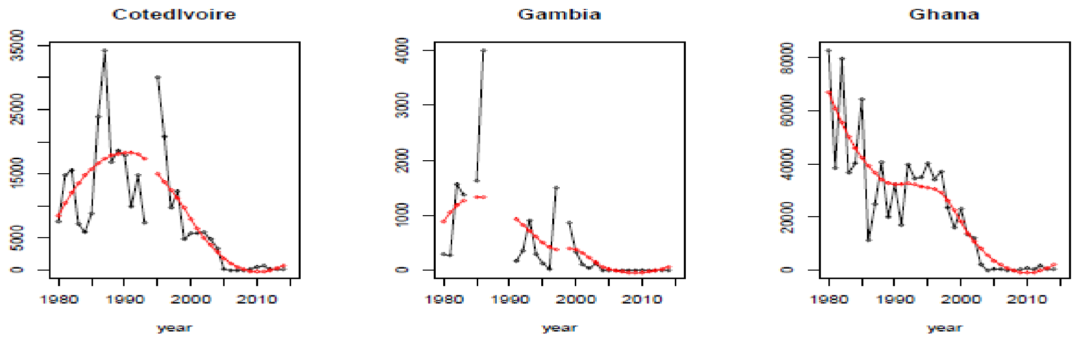

Figure 11b: Loess Fit for Cote d'Ivoire, Gambia and Ghana.

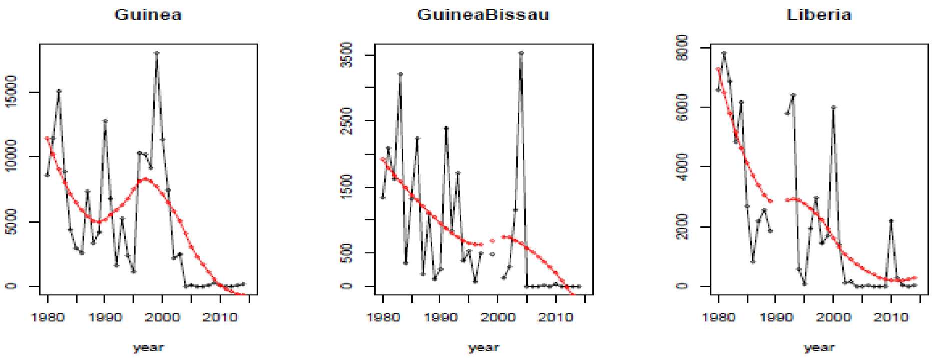

Figure 11c: Loess Fit for Benin, Burkina Faso and Cape Verde.

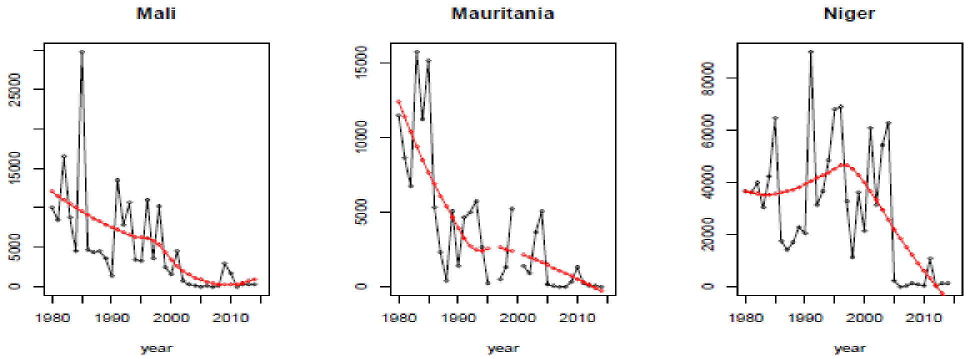

Figure 11d: Loess Fit for Mali, Mauritania and Niger.

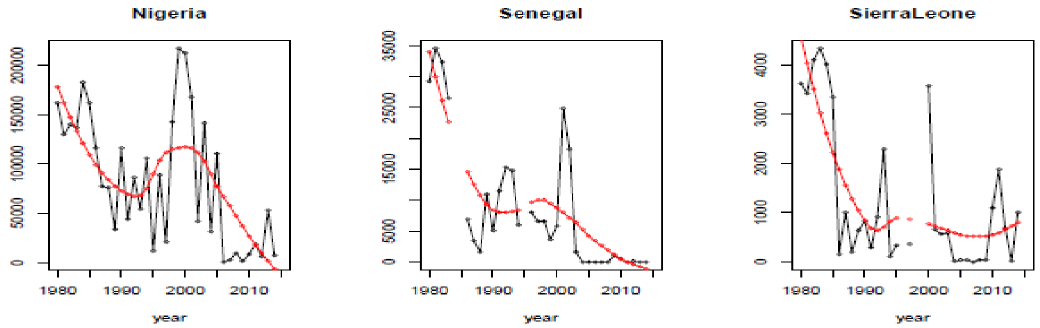

Figure 11e: Loess Fit for Nigeria, Senegal and Sierra Leone.

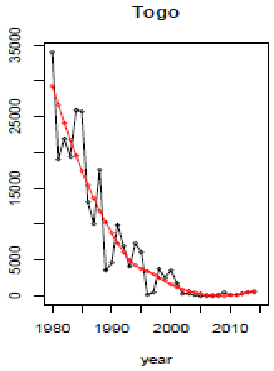

Figure 11f: Loess Fit for Togo.

In this study the method of detrending called differencing was used; which is a method of making a data set (time series) that is non stationary in mean stationary by taking the first difference. The value of the trend line was then subtracted from the original data, giving a time series of residuals from the trend. This “difference” option is attractive for simplicity, and for giving a convenient breakdown of the variance. The residual series is in the same units as the original series, and the total sum of squares of the original data is the same as the trend sum-of-squares plus the residual sum-of-squares. In addition to this, we normalized and ‘centred’ the detrended data on zero (mean = 0, variance = 1). Finally we plotted the standard deviation of points from zero versus year as shown in Figures 12a-12f

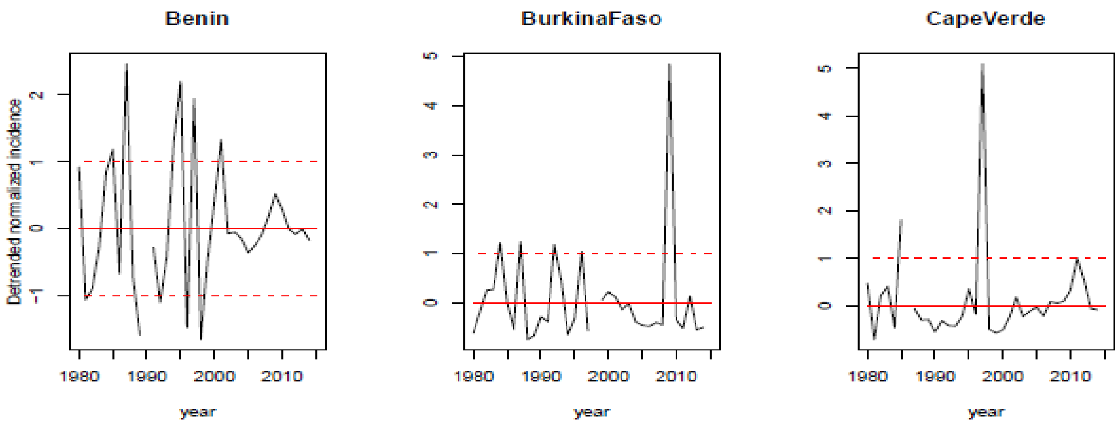

Figure 12a: Profile of Standard deviation of Detrended-normalized-centred data for Benin, Burkina Faso and Cape Verde.

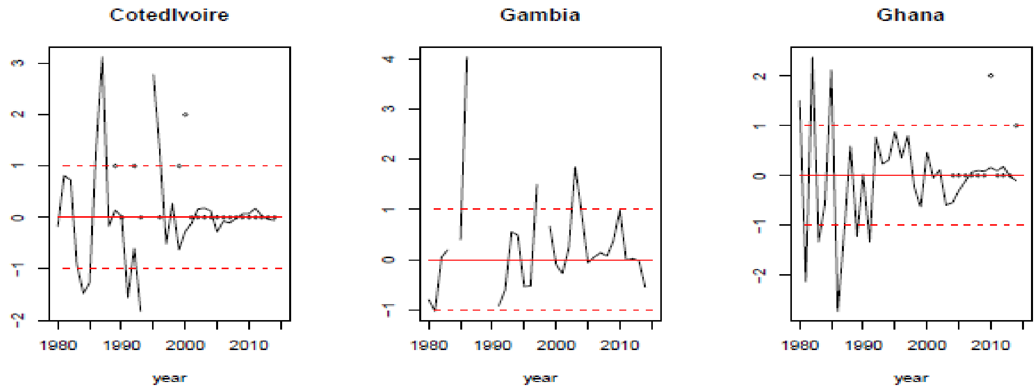

Figure 12b: Profile of Standard deviation of Detrended-normalized-centred data for Cote d'Ivoire, Gambia and Ghana.

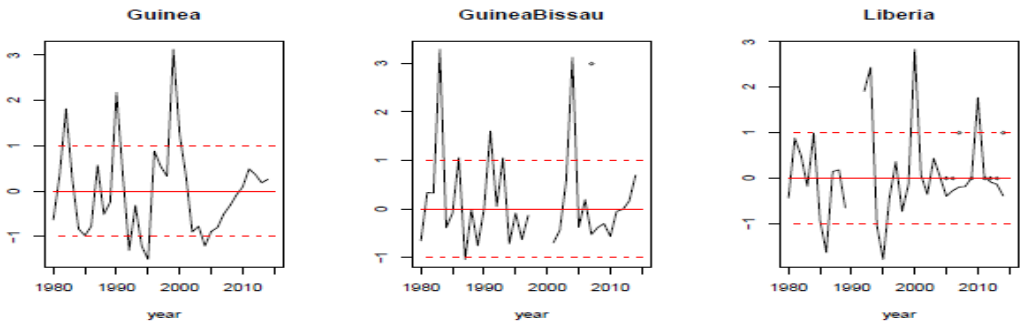

Figure 12c: Profile of Standard deviation of Detrended-normalized-centred data for Guinea, Guinea Bissau and Liberia.

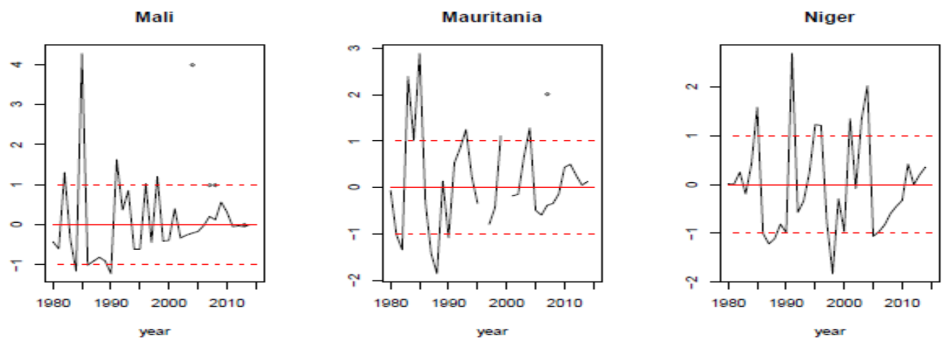

Figure 12d: Profile of Standard deviation of Detrended-normalized-centred data for Mali, Mauritania and Niger.

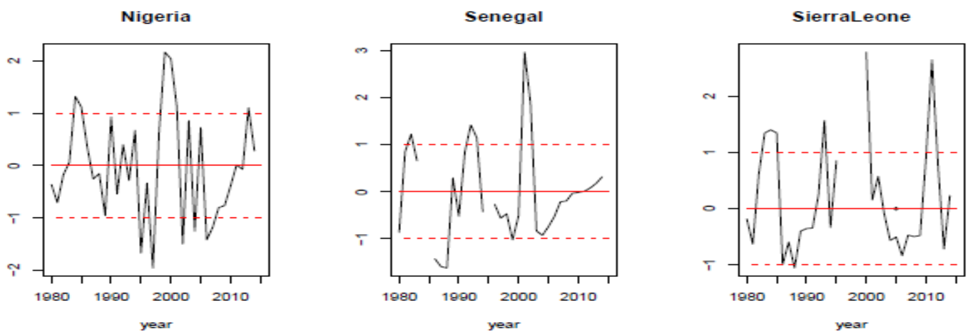

Figure 12e: Profile of Standard deviation of Detrended-normalized-centred data for Nigeria, Senagal and Sierra Leone.

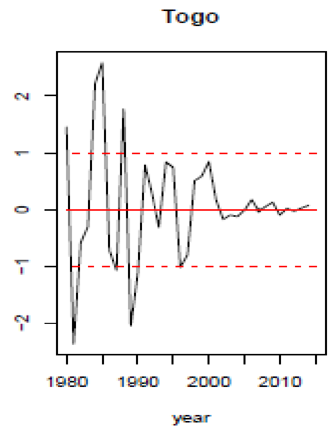

Figure 12f: Profile of Standard deviation of Detrended-normalized-centred data for Togo.

Synchrony

An interesting feature of the data is the degree to which peaks and troughs in cases appear coincident in time. A randomization test of the pairwise correlations in oscillations of the time series found significant evidence of phase synchrony (*p=0.0 39*), implying some synchrony in measles epidemics over extremely large spatial scales.

The degree of phase synchrony tends to decay (if existing) with distance between countries (pairwise coherence), although this relationship is not significant at the 5% level (*P=0.1414*) when tested with appropriately conservative randomization procedures for the time series.

To investigate synchrony, the approach of Hampson et al. [11] was followed. The pairs of detrended, normalized-'centred' time series were compared by using the sum of their products through time. Positive products indicate phase synchrony, and negative products indicate phase asynchrony. Simulated pairs were generated with the time series shuffled by an offset of between 1 and 36 years for all possible variations, and the sum of the product was calculated. An overall statistics, and significance level, was therefore generated from the sum of the pairwise products of all countries included in the analysis. Identical results were obtained by using cross-correlation functions and measures of region wide synchrony. Additionally, the pairwise test statistics were plotted against the log of the distance between countries (from the central point in each country, with distance calculated by using the Haversine formula). To test the a priori hypothesis that correlation between countries decreases with distance, we calculated the slope of a weighted regression of the pairwise correlation statistic and the log distance between countries. We then randomized the distance between countries and recalculated the slope for 10,000 runs to estimate the probability of generating such a slope by chance (shuffle slope). Though there is a phase synchrony (positive values indicated that). there is no sufficient evidence of correlation between geographical distance and phase synchrony.

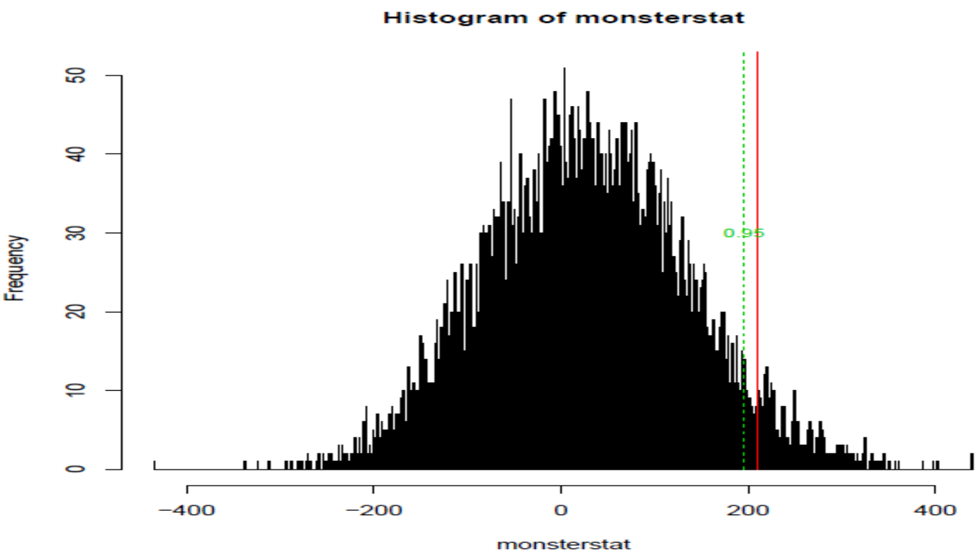

Figure 13: Overall Statistics

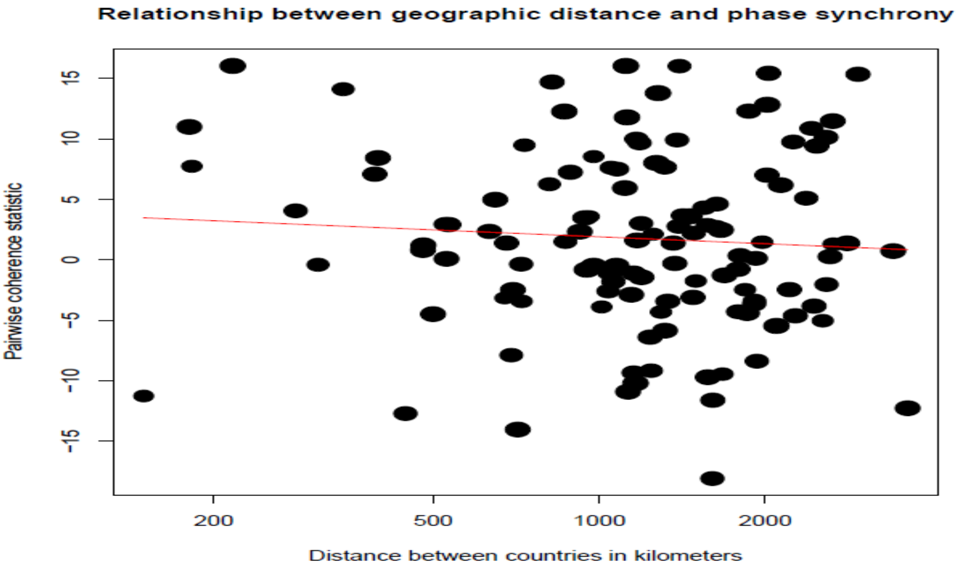

Figure 14: Relationship between geographic distance and phase synchrony. The red line is the weighted regression line (using the square root of the number of points in the time series) of the correlation statistic against the log distance between countries.

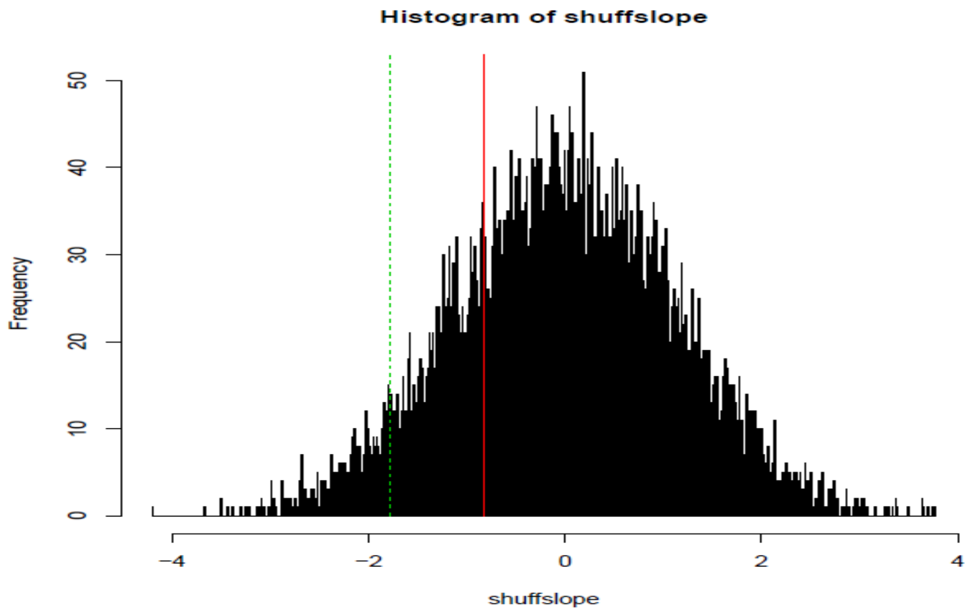

Figure 15: Histogram of 10,000 shuffles.

Effect of Vaccination

We examined the effect of vaccination on the data set for each country and the findings are presented in Figures 16 -21. The findings revealed that the era of vaccination had really impacted on the number of cases reported after the introduction of vaccination (though there were sparks but vaccination had been able to keep the cases low).

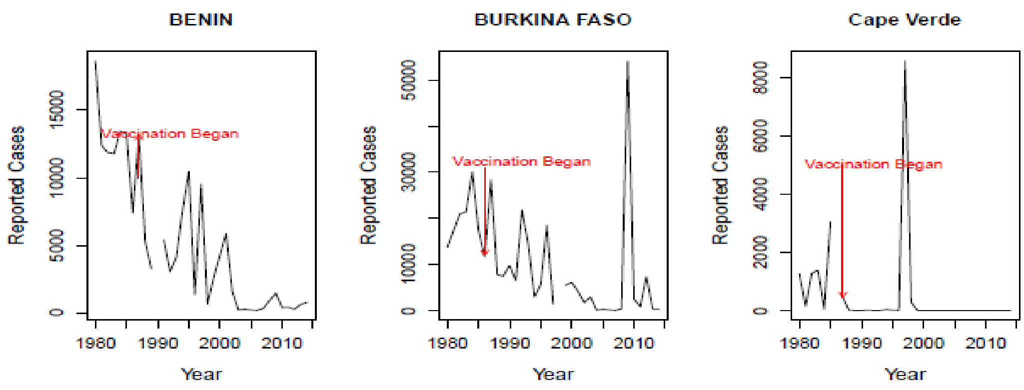

Figure 16: Vaccination profile of Benin, Burkina Faso and Cape Verde.

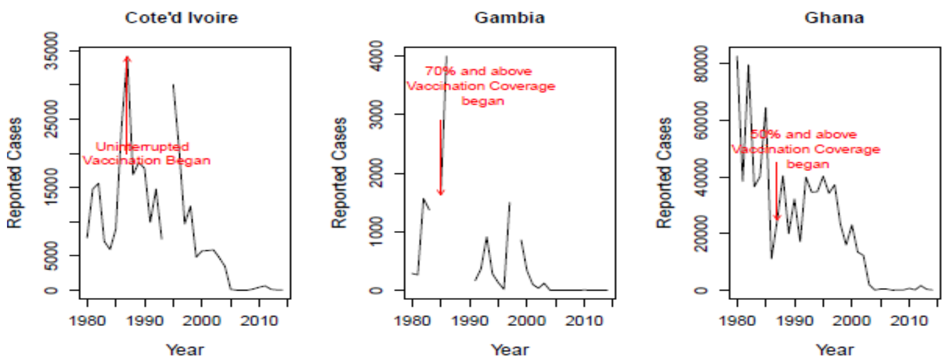

Figure 17: Vaccination profile of Cote d'Ivoire, Gambia and Ghana.

Cote d'Ivoire was able to embark on an uninterrupted vaccination coverage (unlike some countries); Gambia was able to record vaccination coverage that was above 70% since mid 80s and that has helped in keeping the number of reported cases low; Ghana was able to embark on vaccination coverage that average 50% since late 80s.

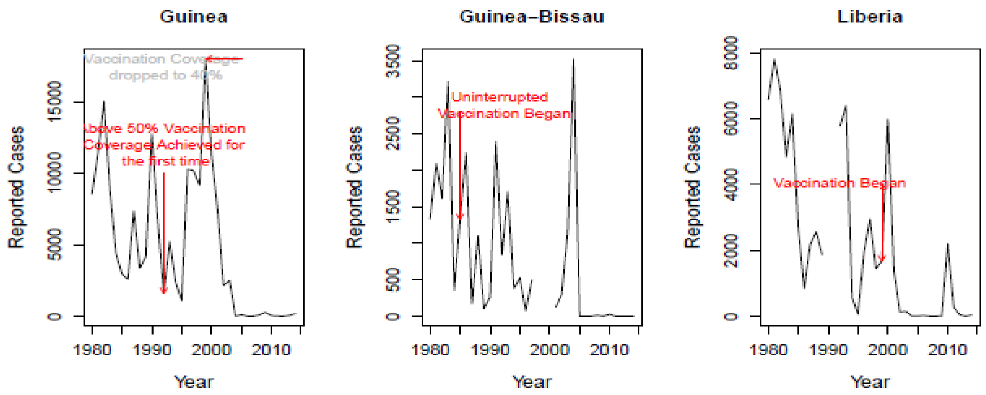

Figure 18: Vaccination profile of Guinea, Guinea-Bissau and Liberia.

Guinea was able to achieve a 50% coverage for the first time in early 90s but drops along the line and the drop coincides with the year the highest number of cases was reported; Guinea-Bissau was able to begin uninterrupted vaccination in mid 80s; Liberia was able to start vaccination in the late 90s.

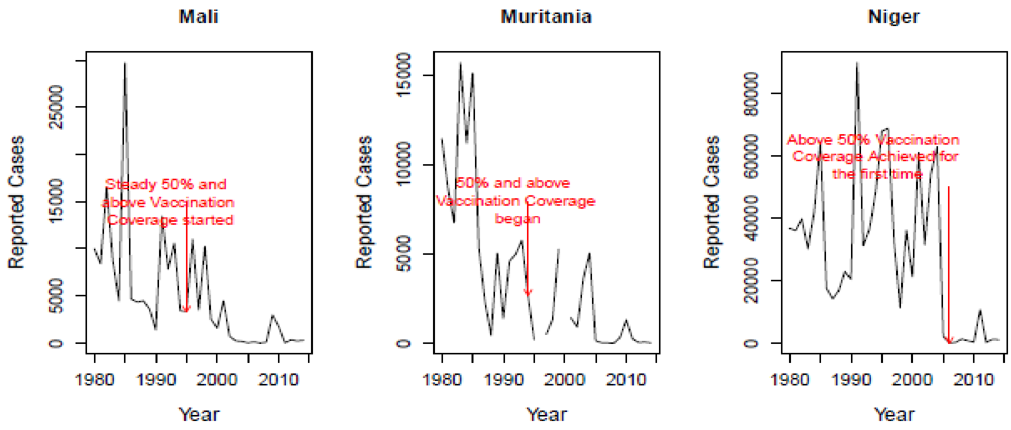

Figure 19: Vaccination profile of Mali, Mauritania and Niger.

Mali was able to begin a steady 50% and above coverage in the mid 90s, though they have been able to attain the percentage before then but it was not sustained; Mauritania was able to achieve 50% coverage in mid 90s and did not go down since then; Niger could not achieve up to 50% vaccination coverage until 2006.

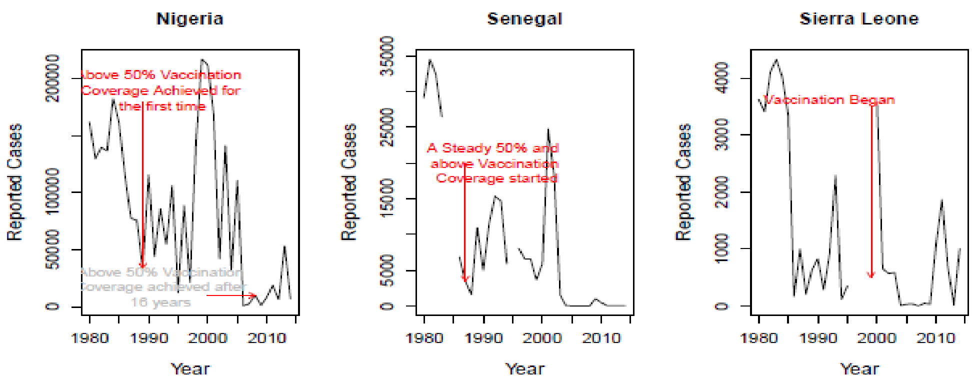

Figure 20: Vaccination profile of Nigeria, Senegal and Sierra Leone.

Nigeria was able to achieve 50% vaccination coverage in late 80s but could not sustain it, they were able to achieve 50% again after 16 years from the first time they achieved it; Senegal began a sustained 50% coverage in late 80s; Sierra Leone could not begin vaccination until late 90s.

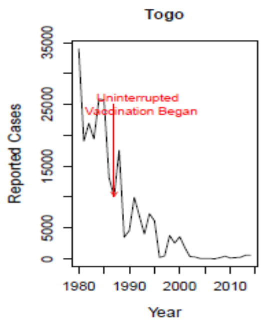

Figure 21: Togo began sustained vaccination coverage in 1987

Conclusion

An extensive analysis on West African Measles data set was performed and some interesting results were obtained. It wasfound out that vaccination has serious effect on the number of reported cases in each of the counties considered. The occurrence of measles outbreak depicts synchrony across countries under consideration but there was no sufficient information to show that there is correlation between geographical distance and phase synchrony. The work was based on secondary data, but similar analysis could be carried out on primary data provided the data capturing mechanism is well designed.

REFERENCES

Anderson, O. (1976). Time series analysis and forecasting: the Box-Jenkins approach: London, Butterworths, 182.

Bjørnstad, O. N., Finkensta¨dt, B., and Grenfell, B. T. (2002), Dynamics of measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecological Monographs*, 72, 169–184.

Dietz, K. (1976). The incidence of infectious diseases under the influence of seasonal fluctuations. *Lecture Notes in Biomathematics*, 11, 1–15.

Earn, D. J. D., Rohani, P., Bolker, B. M. & Grenfell, B. T. (2000). A simple model for complex dynamical transitions in epidemics. *Science*, 287, 667–670.

Earn, D. J. D., Rohani, P. & Grenfell, B. T. (1998). Persistence, chaos and synchrony in ecology and epidemiology. *Proceedings of the Royal Society of London B*, 265, 7–10.

Erlander, S. & Stewart, N. F. (1990). The gravity model in transportation analysis: theory and extensions. *International Science*, Netherlands.

Finkensta¨dt, B. F., & Grenfell, B. T. (2000). Time series modeling of childhood diseases: a dynamical systems approach. Journal of the Royal Statistical Society C, *Applied Statistics* 49, 187–205.

Finkensta¨dt, B.F., Keeling, M. J., & Grenfell, B. T. (1998). Patterns of density dependence in measles dynamics. *Proceedings of the Royal Society of London B*, 265, 753–762.

Grenfell, B. T., Bjørnstad, O. N., & Finkenstadt, B. F. (2002). Dynamics of measles epidemics: scaling noise, determinism and predictability with the time series SIR model. *Ecological Monographs*, 72, 185–202.

Grenfell, B. T. & Harwood, J. (1997). (Meta)population dynamics of infectious diseases. *Trends in Ecology & Evolution*, 12, 395–399.

Hampson, K., Dushoff, J., Bingham J., Bruckner G., Ali Y. H. & Dobson, A. (2007). Synchronous cycles of domestic dog rabies in sub-Saharan Africa and the impact of control efforts, PNAS, 104 (18), 7717–7722.

Hanski, I. (1998), Metapopulation dynamics. *Nature*, 396, 41–49.

McLean, A. R. & Anderson, R. M. (1988). Measles in developing countries. I. Epidemiological parameters and patterns. *Epidemiology and Infection*, 100, 111–133.

Mollison, D., Isham, V. & Grenfell, B. T. (1993). Epidemics: models and data, *Journal of Royal Stastical Society A* ,157, 115–149.

Okubo, A. (1980). Diffusion and ecological problem: mathematical models. *Springer*, Berlin.

Swinton, J. (1998). Extinction times and phase transitions for spatially structured closed epidemics. *Bulletin of Mathematical Biology*, 60, 215–230.

Xia, Y., Bjørnstad, O. N., & Grenfell, B. T. (2004). Measles Metapopulation Dynamics: A

Gravity Model for Epidemiological Coupling and Dynamics, *The American Naturalist*, 164, (2), 267-281