



Mathematical Modeling of Infectious Disease with Special Reference of Statistical Based Methods for Epidemic Surveillance

Shukla Uma Shankar

Department of Mathematics, L.B.S.S., P.G.College, Anandnagar, Maharajganj, U.P., India.

To Cite this Article

Shukla Uma Shankar, "Mathematical Modeling of Infectious Disease with Special Reference of Statistical Based Methods for Epidemic Surveillance", *International Journal for Modern Trends in Science and Technology*, Vol. 07, Issue 04, April 2021, pp.:107-112.

Article Info

Received on 16-March-2021, Revised on 31-March-2021, Accepted on 09-April-2021, Published on 15-April-2021.

ABSTRACT

In the last years, an intensive world wide effort is speeding up the developments in the establishment of a global surveillance network for combating pandemics of emergent and re-emergent infectious diseases. Scientists from different fields extending have teamed up for rapid assessment. Towards this aim mathematical modeling plays an important role in efforts that focus on predicting, assessing and controlling potential outbreaks. To better understand and model the contagious dynamics, the impact of numerous variable ranging from the micro host-pathogen level to host-to-host interactions, as well as prevailing ecological, social, economic and demographic factors have to be analyzed and studied thoroughly. In the present study, I discuss the main approaches that are used for the surveillance and modeling of infectious disease.

KEY WORDS: Epidemic surveillance, Infectious disease, Mathematical modeling, Statistical based methods

INTRODUCTION

The rapid technological and theoretical progress has dramatically enhanced our arsenal in fighting epidemics and we are getting better on it. The global surveillance network is growing under an intensive vaccines and antiviral drug and knowledge goes deep in details such as the molecular structure of a variety of viruses (Abeku, et.al. 2002)¹. A large and intensive research is evolving for the design of better drugs and vaccines, yet, studies warns us that a new pandemic-influenza-type is the most worrisome one is sooner or later on the way (Mishra, 2020a)²³. The problem stems mainly from two reasons:

1.The continuous and ever-lasting mutations of the viruses.

2.The complexity in the disease transmission mechanism.

Unfortunately the odds are that in a real crisis, even if researches succeed to come up with a vaccine tailor made for an emerged virus strain, it is doubtful that it would stop a pandemic (Mishra, 2020b)²⁴.

Mathematical, statistical models and computational engineering are playing a most valuable role in shedding light on the problem and for helping make decisions (Acuna-Soto, et.al. 2005)².

THE BEGINNING OF MATHEMATICAL MODELING IN EPIDEMIOLOGY:

The first mathematical modeling of epidemics dates back in 1766. In this paper "Essaid'une nouvelle analyse de la mortalite cause par la petite verole, Daniel Bernoulli developed a mathematical

model to analyze the mortality due to smallpox in England, which at that time was one in 14 of the total mortality. Bernoulli used his model to show that inoculation against the virus would increase the life expectancy at birth by about three years (**Bernoulli, 1766**)⁴. A translation in English and review of this work can be found in (**Blower and Bernoulli, 2004**)⁶, while a revision by D' Alembert appears in **Dietz and Hesterbeek (2002)**¹⁴. **Lambert, (1772)**²⁰, followed up the work of Bernoulli extending the model by incorporating age-dependent parameters. Laplace has also worked on the same concept. However this line of research has not been developed systematically until the bench mark paper of **Ross (1911)**²⁷, which actually establishes modern mathematical epidemiology. In this work, Ross addressed the mechanistic a priori modeling approach using a set

of equations to approximate the discrete-time dynamics of malaria through the mosquito-borne pathogen transmission (for a discussion and a review of this model see also **Smith, et. al. 2012**)²⁹.

MATHEMATICAL MODELING METHODOLOGIES

Mathematical modeling and simulation allows for rapid assessment. Simulation is also used when the cost of collecting data is prohibitively expensive, or there are a large number of experimental conditions to have been proposed looking at the problem from different perspectives (**Akaike, 1974**)³. The category of approach used in this study statistical methods for surveillance of outbreaks and identification of spatial patterns in real epidemics (**Best et.al., 2005**)⁷. Here I, draw the map of this approach and try to describe their basic underpinning concepts (**Fig. 1**).

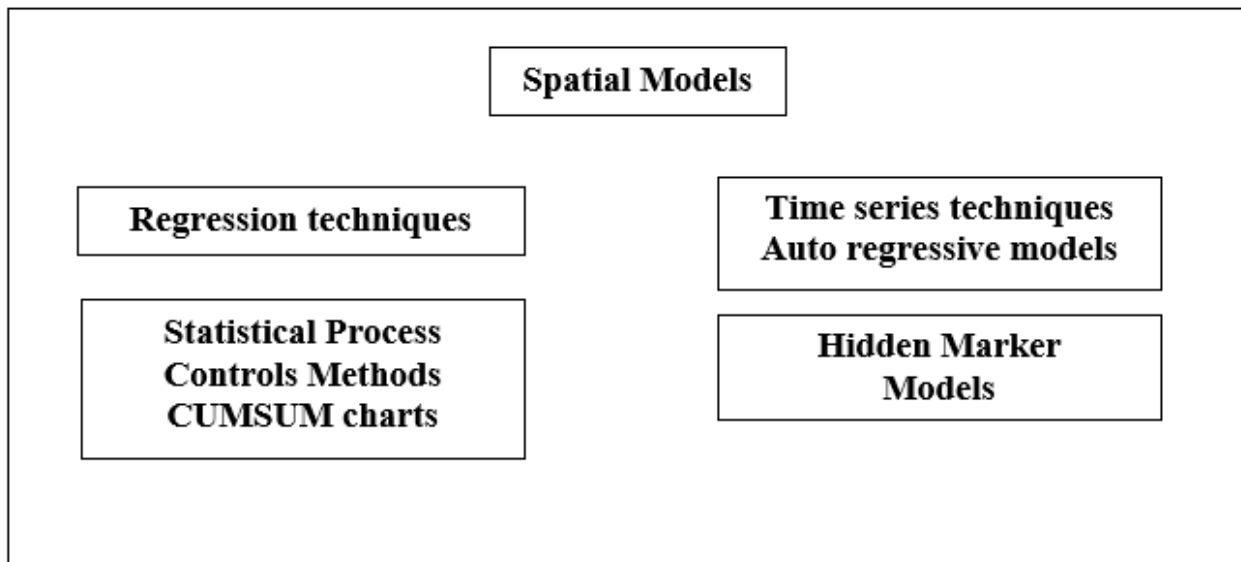


Figure 1: Statistical - based methods for Epidemic Surveillance

STATISTICAL - BASED METHODS FOR EPIDEMIC SURVEILLANCE:

One of the most important aspects in epidemics revolves around the surveillance, early detection of possible outbreaks and patterns that may help controlling a spread (**Carrat and Valleron, 1992**)⁸. One of the very first success stories in the area is the modeling of cholera epidemic that swept through London in 1854. In the dawn of 20th century Greenwood an epidemiologist and statistician was the first professor of Epidemiology and Statistics at the London School of Hygiene and Tropical Diseases establishing a rigorous mathematical connection between fields.

An intensive worldwide effort led by World Health Organization and Centers for Disease control is speeding up surveillance network (**Choi and Thacker, 1981**)⁹. New emerged pandemics

such as the AIDS, the severe acute respiratory syndrome (SARS) of 2002-2003 and the H₁N₁ swine flu of 2009 pandemics reminds is about the importance of surveillance and prompt outbreak detection (**Costagliola, et.al. 1994**)¹¹. Towards this aim, statistical methods have enhanced our potential in fighting epidemic allowing for rapid assessment of emerging situations (**Dietz, 1988**)¹³. Obviously, the correctness of the data and the selection of the appropriate methodology are crucial for the construction of statistical models that can capture in an efficient robust way the communicable disease characteristics (**Cowling, et.al. 2006**)¹². In this study, author present and discuss the most common schemes that can be classified as follows:

REGRESSION METHODS:

Regression models try to detect an outbreak from time-series of epidemic-free periods by monitoring a statistic of reported infected cases, say $y(t)$. An epidemic alert is raised when a certain threshold, say k , is surpassed, defined by $|y(t) - \mu| > k$, (μ being the mean value of the time-series distribution) within a confidence interval (usually of 95%).

A basic regression model is that proposed from Serfling which was initially constructed to monitor the deaths of influenza based on the seasonal pattern of pneumonia and influenza deaths. Due to the seasonal behavior of the disease the following cyclic regression model has been addressed:

$$y(t) = a + bt + \sum_{i=1}^m c_i \cos\theta + \sum_{i=1}^m d_i \sin\theta + e(t)$$

Where, θ is a linear function of time t while the coefficients are to be determined by a parameter identification technique. The cosine and sine terms are used to approximate cyclical seasonal patterns; $e(t)$ is the noise (assumed that is Gaussian distributed with mean zero and variance σ^2) which is estimated from the time-series.

Today, the above approach is used by the Centers for Disease Control in the US, Australia, France, and Italy for the detection of influenza outbreaks. While the approach is very popular among epidemiologists for predicting and surveillance purposes, one has to be cautious about their use as the form of the equation relies usually on adhoc assumptions on the dependence between the dynamics of a disease and the independent factors (variables) that determine its spread (**Kermack and McKendrick, 1927, 1932, 1933**)^{16,17,18}.

Times series analysis based on autoregressive models such as the autoregressive integrated moving average model (ARIMA) and seasonal ARIMA (SARIMA) as well as neural net-works. These models relax the hypothesis of autocorrelation of regression models as well as the hypothesis of simple autoregressive moving) in which past disturbances are not modeled. In this category, ARIMA models are the most commonly used. Their general form reads:

$$A(Z^{-1})[\Delta^d y(t) - \mu] = B(Z^{-1})e(t)$$

where $y(t)$ denotes a stationary stochastic process at time t with mean value $E(y(t)) = \mu$; Z^{-1} is the backward shift operator defined by $Z^{-k}y(t) = y(t - k)$ and Δ^d is the differencing operator of order d defined by $\Delta^d = (1 - Z^{-1})^d$; $A(Z^{-1})$ is the autoregressive

operator defined as $A(Z^{-1}) = 1 + a_1Z^{-1} + \dots + a_{n_a}Z^{-n_a}$; $B(Z^{-1})$ is the moving-average operator defined by $B(Z^{-1}) = 1 + b_1Z^{-1} + \dots + b_{n_b}Z^{-n_b}$; $e(t)$ is the residual (noise) at time t representing the part of the measurement that cannot be predicted from previous measurements. For $d = 0$ and $n_a = n_b = 0$ one gets the random walk with drift. Seasonal differencing enters naturally in the above framework by considering the seasonal differencing operator $\Delta_s^d = (1 - Z^{-1})^d (1 - Z^{-k})^s$ where k is the length of seasonal cycle and S is the degree of seasonal differencing producing series of changes from one season to the next.

Statistical process control methods including cumulative sum (CUSUM) charts and exponentially weighted moving average (EWMA) based methods. CUSUM is probably the most common used technique for the detection of disease out-breaks. This is achieved by monitoring a cumulative performance measure over time. Let us consider the number of infected cases $y(t)$ as observed at different time instances $t, i = 1, 2 \dots n$. In its simple representation, for a single parameter process, CUSUM is defined as

$$CUSUM(i) = \sum_{t=1}^i (y(t_1) - k)$$

OR in a recursive form as

$$CUSUM(0) = 0$$

$$CUSUM(i) = \max(0, CUSUM(i-1) + y(t_1) - k), i \geq 0$$

Where, k is a reference value corresponding to the difference between to the in control and the out of control mean.

Hidden Markov models (HMM) used to explain statistical correlation in time series:

The question that the HMMs come to answer in epidemiology is the following: how can we infer about the dynamics of particular infectious disease and forecast its outbreak when we cannot monitor/record explicitly the characteristics of the disease but we can observe some possible indicators of the disease? For example, can we forecast the evolution of an influenza epidemic by monitoring for example the number of reported cases as recorded through a surveillance network of physicians or in hospital units? HMM models are exploited exactly under these limitations/constraints. Within this context, let us denote by $Y(t)$ the stochastic process of the unobserved (hidden) state, e.g., the number of cases of the disease in the population at time t and

with $O(t)$ the stochastic process of the observable states.

Formally, HMMs are Markov processes, i.e., stochastic processes which satisfy the so called Markov property (here for the sake of presentation we assume discrete in time Markov processes defined by :

$$P(Y(t) = y(t) | Y(t-1) = y(t-1), Y(t-2) = y(t-2) \dots Y(1) = y(1)) = P(Y(t) = y(t) | Y(t-1) = y(t-1))$$

Spatial models for monitoring, identifying and forecasting disease outbreaks in different locations. Most of the infectious disease result to strong spatio-temporal pattern whose systematic analysis is of outmost importance for better understanding, prediction and combating outbreaks (MeLocd, 2000)²¹. Spatial surveillance requires the use of multivariate techniques (Greenland, 2005)¹⁵. Most of the multivariate methods can be viewed as extensions of standard univariate methods- as the ones described above - ; however there are others such as clustering, principal component analysis (PCA) based methods that do not have a common ancestor with univariate ones. Kleinshmidt et al. (2000)¹⁹ used a two tier approach for the surveillance of malaria. They used regression analysis on the larger scale and kriging interpolate the count data at an unobserved location in order to forecast the prevalence of the disease in the local scale. Cohen

et al. (2010)¹⁰ exploited PCA to create a single surveillance index that can be used to summarize temporal temporal and spatial trends of malaria in India (MacNab, 2003)²².

AGENT-BASED SIMULATIONS:

In contemporary mathematical epidemiology, agent-based modeling represents the state-of-the-art for reasoning about and simulating complex epidemic systems. These take into account details such as the transportation infrastructure of the simulated area, the mobility of the population, demographics, and epidemiological aspects such as the evolution of the disease within a host and transmission between hosts (Fig. 2). Public-health epidemiologists, researchers, and policy makers are turning to these detailed models for reasons of ethics, cost, timeliness and appropriateness. In epidemic systems, testing experimental conditions would put the safety of people at risk, creating an ethical problem. In other cases, real-time evaluation of an existing system may be prohibitively long. For example, in a disaster, simulation can be used to rapidly evaluate many previously unexamined alternatives. In all of these cases, since the real-world system under study is a complex system, multi-agent simulations are used as they are considered to incorporate the appropriate level of complexity.

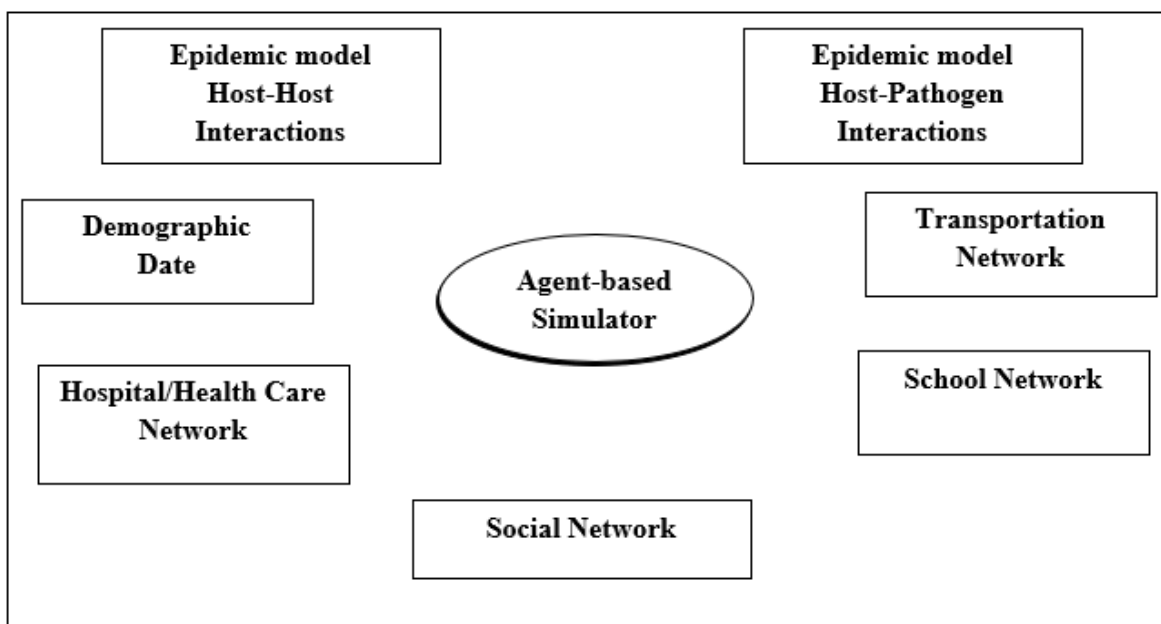


Figure 2: Schematic of the components of an agent-based epidemic simulation.

The author investigated the efficiency of various contagion control scenarios such as vaccination of households, children at school, isolation of infected persons and vaccination of medical staff in hospital. **Balcan et. al. (2009)⁵** investigated how short-scale and long-scale contacts due to air can influence the spatio-temporal pattern of pandemic (**Watkins, et.al. 2008**)³⁰.

CONCLUSION:

In this paper I discussed and presented Key modeling method used for the surveillance presented key modeling method used for the surveillance and forecasting of infectious disease outbreaks (**Nunes, et.al. 1980**)²⁵. The epidemiological model used in this study is the statistical based method for epidemic surveillance. The public health organization throughout the world used such model to evaluate and develop intervention disease outbreak policies for ever-emerging epidemics (**Pelat, et.al. 2007**)²⁶. Simulation allows for rapid assessment and decision making, providing quantification and insight in the the spatiotemporal dynamics of the spread (**Shiryaev, 1963**)²⁸. An intensive inter and multi-disciplinary research effort is speeding up the developments in the field integrating advances from epidemiology (**Sroup, et.al., 1988**)²⁹, molecular biology, computational engineering and science and applied mathematics and well as sociology. Now days, molecular, sociological, demographic and epidemiologic data are exploited to develop state of the art detailed very large scale bottom up agent-based models aspiring to approximate the dynamics of real world cases (**Zhou, et.al., 2010**)³¹.

REFERENCES

- [1] Abeku, T.A.; de Vlas, S.J.; Borshboom, G.; Teklehaimanor, A.; Kebede, A. and Olana, D. et. al., (2002): Forecasting malaria incidence from historical morbidity patterns in epidemic prone areas of Ethiopic seasonal adjustment method performs best. *Trop Med Int Health* 2002; 7:851-7; PMID : 12358620; <http://dx.doi.org/10.1046/j.1365-3156.2002.00924x>
- [2] Acuna-Soto, R.D.; Stahle, D.W.; Therrell, M.D.; Gomcz, C.S. and Cleaveland, M.K. (2005): Drought, epidemic disease, and the fall of classic period cultures in Mesamerica (AD 750-950). Hemorrhagic fevers as a cause of massive population loss. *Med Hypotheses* 2005; 65:405-9; PMID:15922121; <http://dx.doi.org/10.1016/j.mehy.2005.02.025>
- [3] Akaike H.A.(1974): new look at the statistical model identification. *IEEE Trans Amutomat Contr.* 1974:19:716-723. <http://ds.doi.org/10.1109/TAC.1974.1100705>.
- [4] Bernoulli D.(1766): Essai d'une nouvelle analyse de la mortalite cause par la petite veroble. *Mem, Math Phys Acad Roy Sci Paris* 1766 ; 1:1-14.

- [5] Balcan, D.; Vittoria, C.; Bruno, G. and Hao, H. (2009): Multiscale mobility networks and the spatial spreading of infectious diseases. *PNAS* Vol. 106(51):21484-21489. <https://doi.org/10.1073/pnas.0906910106>.
- [6] Blower, S. and Bernoulli, D.(2004): An attempt at a new analysis of the mortality cause by smallpox and of the advantage of inoculation to prevent it. *1766 rev Med Virol* 2004 ; 14:275-88; PMID : 15334536; <http://dx.doi.org/10.1002/nmv.443>.
- [7] Best, N.; Richardson, S. and Thomson, A. (2005): A comparison of Bayesian spatial models for disease mapping. *Stat Methods Med Res.* 2005; 14:35-59; PMID:15690999;<http://dx.doi.org/10.1191/0962280205sm388oa>
- [8] Carrat, F. and Valleron, A.J. (1992): Epidemiologic mapping using the "kriging" method: application to an influenza-like illness epidemic in France. *Am J Epidemiol.* 1992; 135:1293-300; PMID 1626546.
- [9] Choi, K. and Thacker, S.B. (1981): An evaluation of influenza mortality surveillance, 1962-1979. I. Time series forecasts of expected pneumonia and influenza deaths. *Am J Epidemiol* 1981 : 113:215-26; PMID : 6258426.
- [10] Cohen, A.A.; Dhingra, J.; Jorkar, R.M.; Rodriguez, P.S.; Sharma, V.P. and Jha, P. (2010): The Summary Index of Malaria Surveillance (SIMS): a stable index of malaria within India. *Popul Health Mctr* 2010;8;1 PMID 20181218; <http://dx.doi.org/10.1186/1478-7954-8-1>.
- [11] Costagliola, D.; Flahault, A.; Galinec, D.; Garnerin, P.; Menares, J. and Valleron, A.J. (1994): When is the epidemic warning cut-off point exceeded; *Eur J Epidemiol* 1994;10:4756; PMID:7843360; <http://dx.doi.org/10.1007/BF01719680>.
- [12] Cowling, B.J.; Wong, I.O.L.; Ho, L.M.; Riley, S. and Leung, G.M. (2006): Methods for monitoring influenza surveillance data, *Int J Epidemiol* 2006, 35 : 1314-21; PMID : 16926216; <http://dx.doi.org/10.1093/ije/dy1162>
- [13] Dietz, K. (1988): The first epidemic model : a historical note on P.D. En'ko. *Aust J Stat* 1988;30A 56-65; <http://dx.doi.org/10.1111/j.1467-842X.1988.tb00464.x>
- [14] Dietz, K. and Hesterbeek, J.A.P. (2002): Baniel Bernoulli's epidemiological model revisited. *Math Biosci* 2002; 180:1-21; PMID : 12387913 ; [http://dx.doi.org/10.1016/S00025-5564\(02\)00122-0](http://dx.doi.org/10.1016/S00025-5564(02)00122-0).
- [15] Greenland, S. (2005): Regression Methods for Epidemiologic Analysis, *Handbook of Epidemiology. Part II*, 2005. 625-91;http://dx.doi.org/10.1007/987-3-540-265771_17
- [16] Kermack, W.O. and McKendrick, A.G. (1927): Contribution to the mathematical theory of epidemics, *Proc. R. Soc. London, A Contain Pap Math Phys Character* 1927; 115-700-21; <http://dx.doi.org/10.1098/rspa.1927.011>
- [17] Kermack, W.O. and McKendrick, A.G. (1932): Contributions to the mathematical theory of epidemics, part II. *Proc. R. Soc. London*, 1932; 138-83; <http://dx.doi.org/10.1098/rspa.1932.0171>
- [18] Kermack, W.O. and McKendrick, A.G. (1933): Contributions to the mathematical theory of epidemics, part III. *Proc R Soc Lond* 1933; 141 : 94-112; <http://dx.doi.org/10.1098/rspa.1933.0106>.
- [19] Kleinschmidt, I.; Bagayoko, M.; Clarke, G.P.Y.; Craig, M. and Le Sueur D. (2000): A spatial statistical approach to malaria mapping. *Int. J. Epidemiol.* 2000; 29:355-361; PMID:10817136; <http://dx.doi.org/10.1093/ije/29.2.355>.
- [20] Lambert, J.H. (1772): Die Gebraucher der Kinderblattern. *Beyrage zum Behraucher der Mathematik und deren Anwendung. Buchhandlung der Realschule* 1772; 3:568.
- [21] MeLocd, K.S. (2000): Our sense of Snow : the myth of John Snow in medical geography. *Soc. Sci. Med* 2000;

- 50:923-35; PMID : 10714917 ; [http://dx.doi.org/10.1016/S02779536\(99\)00347-7](http://dx.doi.org/10.1016/S02779536(99)00347-7)
- [22] MacNab, Y.C. (2003): Hierarchical Bayesian modeling of spatially correlated health service outcome and utilization rates: *Biometrics* 2003:59:305-16; PMID:12926715; <http://dx.doi.org/10.1111/1541-0420.00037>.
- [23] Mishra, S. P. (2020a): A Review on Corona Virus Disease 2019 (COVID-19): Current Status and Future Perspectives. *Int. J. Creat. Res. Thoughts* Vol. 8 (6): 3542-3551.
- [24] Mishra, S. P. (2020b): A Review on Current Status of Corona Virus Disease 2019 (COVID-19) Pandemic in India. *IJRAR* Vol.7(3):621-630. DOI: 10.1729/Journal.24032.
- [25] Nunes, B.; Viboud, C.; Machado, A.; Ringholz, C.; Rebelode-Andrade, H. and Nogueira, P. et.al. (1980): Excess mortality associated with influenza epidemics in Portugal, 1980 to 2004, *PLoS One* 2011: 6: e20661 ; PMID : 0020661
- [26] Pelat, C.; Boclle, P.Y.; Cowling, B.J.; Carrat, F.; Flahault, A. and Ansart, S. et.al. (2007): Online detection and quantification of epidemics, *BMC Med Inform Decis Mak* 2007; 7:29; PMID: 17937786: <http://dx.doi.org/10.1186/1472-6947-29>.
- [27] Ross, R. (1911): *The prevention of malaria*, London: John Murray, 1911:651-686.
- [28] Shiryaev, A.N. (1963): On optimum methods in quickest detection problem. *Theory Probab Appl.* 1963. 8:22-46; <http://dx.doi.org/10.1137/1108002>.
- [29] Smith, D.I.; Battle, K.E.; Hay, S.I.; Barker, C.M.; Scott, T.W. and McKenzie, F.E. (2012): Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS Pathog* 2012: 8:c1002588: PMID: 22496646: <http://dx.doi.org/10.1371/journal.ppat.1002588>.
- [30] Sroup, D.E.; Thacker, S.B.; Herndo, J.L. and Herndo, J.L. (1988): Application of multiple time series analysis to the estimations of pneumonia and influenza mortality by age 1962-1983. *Stat MED* 1988.7:1045-59; PMID : 3264612; <http://dx.doi.org/10.1371/7-e40310> PMID-22905091; <http://dx.doi.org/10.1371.journal.pone.0040310>
- [31] Watkins, R.E.; Eagleson, S.; Vccnendaal, B.; Wright, G. and Plant, A.J. (2008): Applying cusum-based methods for the detection of outbreaks of Ross River virus disease in Western Australia : *BMC Med Inform Decis Mak* 2008; 8:37; PMID:18700044; <http://dx.doi.org/10.1186/1472-6947-37>.
- [32] Zhou, Q.; Luo, Y. and Wang, Z. (2010): control chart based on lick childhood ratio test for detecting parterned mean and variance shifts. *Compur Stat Data Anal* 2010; 54: 1634-45; <http://dx.doi.org/10.1016/j.csda.2010.01.020>.