Use an Improved Chicken Swarm Algorithm to Determine Optimal Stratigraphic Boundaries for a Threshold Geometric Stochastic Process for the Epidemic Covid-19

Omar R. Jasim
College of Administration and Economics, University of Hamdaniya.
Qutaiba N. Nauef
College of Administration and Economics, University of Baghdad.

Abstract

The number of daily cases of a specific infection, such as the emerging acute respiratory syndrome Corona (Covid-19), frequently reveals several patterns during the outbreak of a particular epidemic disease: a monotonic increase during the growing stage or the outbreak of the epidemic, stationary in the number of daily cases called in Some sources are the stabilization stage, that is, controlling the epidemic to eliminate it, and then decreasing during the declining stage. In this research, an artificial intelligence technique is used, represented by the improved chicken swarm optimization algorithm to determine the optimum stratigraphic limits for modeling data on the numbers of daily cases of the Corona virus for the three Iraqi governorates (Baghdad, Erbil and Basra). For this purpose, a stochastic model called the Geometric stochastic process model with an intelligent threshold was proposed, Since the turning points (inversions) of the data were determined using an intelligent technique to reach the best representation of the data under study. It was concluded that the proposed model outperformed to the traditional model by (23.41%) in modeling epidemic data in Baghdad governorate, while the superiority in epidemic data for Basra and Erbil governorates (23.96%) and (26.40%), respectively, and this indicates That validity of the theoretical postulates that have been addressed in the theoretical side.

Keywords: Chicken Swarm Algorithm, Monotone Trend, turning points, Threshold Geometric Stochastic Process, Corona virus Data (Covid-19), Optimal Stratigraphic Boundaries, Moving Window Method, Root Mean Squares Error.

استخدام خوارزمية الدجا ج المحسنة لتحديد الحدود الطبقية المثلى للعملية التصادفية الهندسية ذات العتبة لوباء كوفيد ١٩

عمر رمزي جاسم

كلية الادارة والاقتصاد، جامعة الحمدانية، موصل – العراق.

قتيبة نبيل نايف

كلية الادارة والاقتصاد، جامعة بغداد، بغداد – العراق.

المستخلص

اثناء تفشي مرض وبائي معين على سبيل المثال، المتلازمة التنفسية الحادة المستجدة كورونا (*Covid-19*)، غالباً ما تظهر عدد حالات الاصابة اليومية لوباء معين اتجاهات متعددة: زيادة رتيبة خلال مرحلة النمو أو تفشي الوباء، يتبعها مرحلة ثبات عدد حالات الاصابة اليومية وتُسمى في بعض المصادر مرحلة الاستقرار أي السيطرة على الوباء للقضاء علية ثم تتناقص خلال مرحلة الانخفاض او الانحسار او التلاشي للوباء. تم في هذا البحث استخدام تقنية من تقنيات الذكاء الاصطناعي والمتمثلة باستخدام خوارزمية أمثلة سرب الدجاج المحسنة لتحديد الحدود الطبقية المثلى لنمذجة البيانات الخاصة بأعداد الاصابات اليومية لوباء فايروس الكورونا (*Covid-19*) وللمحافظات العراقية الثلاثة (بغداد، اربيل والبصرة).اذ تم لهذا الغرض اقتراح نموذج تصادفي عشوائي يُسمى بنموذج العملية التصادفية الهندسية ذو العتبة الذكائية، أذ تم تحديد نقاط التحول (الانقلاب) للبيانات باستخدام تقنية ذكائية للوصول الى أفضل تمثيل للبيانات قيد الدراسة. أذ تم التوصل الى أن النموذج المقترح تفوق على النموذج التقليدي بنسبة(٢٣,٤١٪) في نمذجة بيانات الوباء في محافظة بغداد، في حين كان التفوق في بيانات الوباء لمحافظتي البصرة وأربيل بنسب (٢٣,٩٦٪) و(٢٦,٤٠٪) على التوالي، وهذا يُشير الى صحة المسلمات النظرية التي تم التطرق اليها في الجانب النظري.

الكلمات المفتاحية: خوارزمية الدجاج، الحدود المثلى للعملية الهندسية، فاروس كورونا– ١٩.

(1) Introduction

Monotone Trends is used in a wide variety of functional, engineering, and health-related applications. The Single trend data was previously modeled using the Non-Homogenous Poisson Process (NHPP) model with a Single Hazard Rate. The Cox-Lewis model and the Weibull Process Model are also used in this model. The Ordinary Geometric Stochastic Process (OGSP), which was proposed by [8], is a direct approach to modeling data with monotonic trend. The Ordinary geometric stochastic process $[Y_i = a^{i-1}X_i \, ; i = 1,2, \dots]$ has a parameter called the ratio of geometric stochastic process, it measures the direction and the strength of this direction for geometric stochastic process [1]: it is increasing when the ratio (a <1), and it is Decrease when the ratio (a> 1) while it is Stationary when the ratio (a = 1), then it is called Renewal Stochastic Process (RSP). [11]

The Geometric stochastic process often deals the data with Single Trend. However, the number of real applications show several Trends. During an outbreak of a specific infectious disease, for example, the number of regular infections always goes through a growth period, then a stabilization stage, and finally a decreasing stage. This statistical model aids in determining the epidemic's peak and end points, as well as assisting health-care authorities in determining the phases of the epidemic's progression and evaluating the activities and policies used in therapy to resolve and eradicate the epidemic. Furthermore, it provides them with the knowledge they need to make a decision in a rapidly changing situation.[4]

The geometric stochastic process's ratio parameter (a) is the ratio of the predicted current outcome to the previous result, which shows the trend's evolution after random noise is removed. As a result, the ratio geometric stochastic process (a) is important in epidemiological modeling because it is related to the Basic Reproductive Number (R 0), which reflects the number of new infections caused by each infected person and the virus's aggressive capacity to infect new people for each infected person. The value of the aggressive capacity of the epidemic (R_0> 1), then the epidemic will attack more and more people in each round of infection so that the expected number of daily infections increases, which implies that the stochastic process is increasing with time and that the proportion of the geometric stochastic process is defined as follows (a <1). On the other hand, when the value of the hostile capacity of the epidemic is given in the following form (R_0<1), this means that the hostility of the epidemic decreases until it fades with time and that the ratio of the geometric stochastic process is defined as follows (a> 1)[9][2].

The relationship between the value of the antagonism ratio (R_0) and the geometric stochastic process (a) is a key indicator of whether or not the epidemic is under control[10].

Because of the importance of the topic and its impact on human life as well as the national economy of the country, extensive research and numerous studies have been conducted to understand these epidemics and the extent of their aggressive power in spreading and outbreaks. Among these epidemics are viral hepatitis of multiple types and severe acute respiratory syndrome (SARS) as well as the associated Corona virus epidemic. With severe acute respiratory syndrome type 2. (SARS-COV-2)

As a result, several researchers analyzed the statistical model for these epidemics' data and attempted to understand how they spread and faded, as each of [3] proposed a basic mathematical model to describe the data of severe acute respiratory syndrome (SARS) as follows:[10]

$$C(nt) = \sum_{i=0}^{n} R_0^i \qquad \& \qquad D(nt+d) = \sum_{i=0}^{n} C(nt) \times f \qquad \ldots (1)$$

Where the total number of infected cases by days is given by C(s), total number of deaths by day s is denoted by D(s), reproductive number R_0 is the simplest (number of new infectious cases from one infectious case in one incubation period), fatality rate is denoted by the symbol f, and (t) is the time it takes for an egg to hatch (number of days from infection to symptom). and (d) refers to the period that you're experiencing symptoms (number of days from symptom to recovery or death).

Through the above model, we note that it is not practically applicable, because it requires model assumptions represented in determining the early stage of transmission of the epidemic, as it is difficult to determine this because of the lack of a good monitoring system to identify these data[6].

On the other hand, geometrical shaped stochastic processes models focus on a series data with a single-trend, as they are considered simple models in determining the single-trend of the data, and to take into account the random nature of the data, a threshold geometric stochastic process model was proposed to address multiple trends in the data under study.

(2) Target:

This study is the first attempt to apply each of the proposed models in the theoretical aspect represented by the use of intelligent techniques on epidemiological data to model data on daily casualty numbers for the three Iraqi governorates (Baghdad, Erbil and Basra). One of the defects identified on the (OGSP) model is the inability to deal with multiple monotone trend data, as well as the fact that the shape parameter does not change, meaning it remains constant for all periods without Therefore, some suggested techniques were used to overcome this defect, including the Threshold Geometric Stochastic Process (TGSP). The idea of this model is based on determining the turning points (TP) to determine the threshold. The threshold and from it we suggest the Moving Window Technique (MWT) to determine the best turning point (s).

(3) Threshold Geometric Stochastic Process (TGSP):

In this section, we will be suggested to use the Threshold Geometric Stochastic Process Model (TGSP) by dividing the process into (k) sections (layers), as it is considered one of the best methods in processing multi-trend data (MT)[2].

Therefore, the urgent need for this topic is that it is included in the treatment and identification of epidemics spread in our real life. Therefore, the monotone stochastic process with a single-trend has been studied and an attempt to expand it to take into account the random nature of the data, which is multi-trend, as its application to the epidemiological data makes us search for a method through which the single-trend is generalized into multiple monotone trends. This is done by using the monotonic stochastic model. If the generalization of the normal geometric stochastic process model to the geometric stochastic process model, this is done using the monotone random model. If the generalization of the geometric random process model to the geometric random process model with a threshold is to divide the process into (k) sections or parts as follows:

$$GP_k = \{X_i \,; T_k \le i \le T_{k+1}\} \qquad k = 1,2,\dots,K \qquad \dots (2)$$

To each (n_k) observational pattern that starts with a turning point (T_k) where:

$$T_k = 1 + \sum_{j=1}^{k-1} n_j \qquad k = 1,2,\dots,K \qquad \dots \quad (3)$$

$$T_{k+1} = n + 1 \quad ; \quad \sum_{k=1}^{K} n_k = n \quad ; \quad T_1 = 1 \qquad \dots \quad (4)$$

Then the model can be represented as the renewal stochastic process as follows :

$$RP_k = \{a_k^{\,i-T_k} X_i \,;\, T_k \le i \le T_{k+1}\} \qquad k = 1,2,\dots,K \qquad \dots \quad (5)$$

where $(a_k > 0)$ called the ratio of kth geometric stochastic process, which measures the trend and strength of the trend of the process, while the expected value and variance of the threshold geometric stochastic process can be given as follows:

$$\mathbb{E}(X_{T_k+t}) = \frac{\mu_k}{a_k^t} \quad \& \quad Var(X_{T_k+t}) = \frac{\sigma_k^2}{a_k^{2t}} \qquad t = 0,1,\dots,n_k - 1 \qquad \dots \quad (6)$$

Thus $\{a_k, \mu_k \;\&\; \sigma_k^2\}$ are important threshold geometric stochastic process parameters that fully determine both the mean and variance of (X_{T_k+t}) in the kth geometric stochastic process.

We propose in this paper to use the chicken swarm optimization algorithm with the moving window technique to estimate the turning points $\{T_k \,;\, k = 1,2,\dots,K\}$, which represent the time when the trend changes its direction. the threshold geometric stochastic process models are applied successively to a subset of data of fixed length (L) starting from time $\{s = 1; s = 2\}$ and so on up to $\{s = n - L\}$. Since the ratio $(a_s > 0)$ of a (GSP) will change with each (s) as the window moves, turning point s (T_s) can be located by a change of (a_s) from 'less than 1' to 'greater than 1' and vice versa which occur at s=T_s; k =1, 2,…, K. Then the GSP model are given by:

$$GP_k = \{X_i \,; T_k \le i \le T_{k+1}\} \qquad k = 1,2,\dots,K \qquad \dots \quad (7)$$

As a result, depending on the nature of the data, we can detect multiple turning points. In certain experiments, the ratio parameter (a_s) appears to fluctuate about one, resulting in short patterns and unstable parameter estimates. To prevent "noisy" shifts, short trends are combined with the next trends and given some constraints so that they are not as short and unpredictable, and among these restrictions is that the number of observations in each point is not less than seven {h = 7}. Also, the change in the value of the geometric stochastic process ratio (a) for each point exceeds the value of (0.005), meaning that:[2]

$$\left|a_{(T_k)} - a_{(T_{k-1})}\right| > d = 0.005 \qquad \dots \quad (8)$$

Moreover, different values of $\{h, d\}$ will be tested until reasonable results are obtained with high accuracy. the window widths (L) are great importance in determining the level of accuracy of the results, since different window widths (L) give a different set of turning points (T_i). Whereas, if the window width (L) is short, then the change is measured more accurately. Meanwhile, the noise may also be captured in various changes.

On the other hand, if the window width (L) is very wide or large, then the average changes will be calculated very large, which leads to the loss of a lot of information and changes in trends because the increase at certain times may be offset by the decrease in other periods, so this will lead to a bias. Inaccuracy in the results. If the window length is equal to the total sample size (n), then only one trend can be identified, meaning that there is one threshold and there are no turning points (T_s).

For the purpose of analyzing the data set, different values were assumed for the basic parameters, as the length or width of the window (L) was set within a range of values (20-30) and for each window there is a set of corresponding trends as well as a set of parameters for each trend or turning point.

(4) Chicken Swarm Optimization Algorithm:

The chicken swarm algorithm is one of the inspired algorithms by nature was first proposed by[12] which simulates the hierarchical order of the chicken swarm and the behaviors of the chicken swarm including roosters, chickens, and chicks.

Optimization can also be defined as one of the branches of knowledge that deals with the discovery or investigation of optimal solutions to a particular issue within alternatives group, or it can be viewed as one of the main quantitative tools in a decision-making network, since decisions must be taken to improve one or more of targets in a specific set.[13]

Optimization has been an active field of research for several decades, and the scientific and technological growth of recent years has enabled the emergence of an abundance of difficult optimization problems that have been the reason to development more efficient algorithms.

In the real world, the optimization suffers from several problems, the most important which are:

1. Difficulty in identifying the overall optimal solutions from local.
2. Noise in evaluating the solution.
3. The curse of the existence of dimensions or the large presence of dimensions (such as the exponential growth of the search space with the dimension of the problem).
4. Difficulties associated with restrictions .

Today, there is a rich variety of algorithms established for most types of problems, however, different cases of the same problem may have different computational requirements and this has given way to developing new algorithms and improving the existing ones, and as a result there will be a continuous need for new and more advanced ideas in optimization theory and its applications.

The most recent development over the past two decades is the tendency to use mats heuristic algorithm. In fact, the vast majority of modern optimization techniques are usually Heuristic Algorithm.

The Chicken Swarm Optimization (CSO) algorithm is a new algorithm that was proposed to address optimization problems, and that idea is based on imitating the hierarchical arrangement in the chicken swarm and the behavior, including (Roosters, hens, and chicks, as the CSO can efficiently extract the intelligence of a chicken swarm to a computer. The chicken swarm algorithm has demonstrated high speed and efficiency in finding the best solution through practical experiments with a variety of optimization problems. [12]

We may mathematically improve the CSO algorithm based on the aforementioned definitions. The following rules were used to idealize the chickens' actions for the sake of convenience.

1. There are many groups of chickens in a swarm. A dominant rooster, a couple of hens, and chicks make up each group.
2. The fitness values of the chickens decide how to split the swarm into many classes and determine the identity of the chickens (roosters, hens, and chicks). Chickens will be given the highest possible physical education. Roosters, each of which will be the leader of a group of roosters. Chicks are chickens that have the lowest health scores around the board. The hens will be the rest of the group. The hens choose which community they will reside in at random. Furthermore, the mother-child relationship between chickens and chicks is formed at random.
3. In a group, the hierarchical order, dominance relationship, and mother-child relationship would remain unchanged. Only these cases are modified after a certain number of time measures (G).
4. Chickens may prevent individuals from eating their own food by following the group of their mate rooster in search of food. Assume the chickens snatch healthy food that other humans have already discovered. Food is sought by the females in the vicinity of their mother (hen). Individuals that are dominant have an advantage in the food competition.

Assume that (RN), (HN), (CN), and (MN) respectively reflect the number of roosters, hens, chicks, and mother hens. The best (RN) chickens are thought to be roosters, while the worst (CN) are thought to be chicks. The rest is done as if it were a swarm of hens. All (N) virtual chickens, represented by their positions $\{x_{i,j}^{t}\ ; i \in [1, \dots, N], j \in [1, \dots, D]\}$ at time step t, search for food in a dimensional space. The optimization problems in this work are the most basic. As a result, the best (RN) chickens are those with the lowest (RN) fitness values.

(4-1) Chicken Action:

Roosters with higher fitness values have preference over those with lower fitness values when it comes to food. For the sake of convenience, this condition may be simulated by cocks with higher fitness values being able to forage in a wider variety of locations than cocks with lower fitness values. This can be expressed as follows:

$$x_{i,j}^{t+1} = x_{i,j}^{t} * (1 + Randn(0, \sigma^2)) \quad \dots \quad (9)$$

$$\sigma^2 = \begin{cases} 1 & ; \quad if\ f_i \le f_k \\ \exp\left(\frac{(f_k - f_i)}{|f_i| + \varepsilon}\right) & ; \quad otherwise \end{cases} \qquad k \in [1, N], k \neq i \quad \dots \quad (10)$$

Where Randn $(0, \sigma^2)$ represents a normal distribution with a zero mean and variance (σ^2). The smallest constant in the computer is (ε), which is used to prevent zero-division-error. A rooster's index (k) is chosen at random from the rooster's group, and the fitness value of the corresponding x is (f).

The hens, on the other hand, will hunt for food alongside their roosters. Furthermore, they would steal good food found by other chickens at random, while being repressed by the other chickens. The dominant hens will have an advantage over the submissive hens when vying for food. These phenomena can be mathematically expressed as follows:

$$x_{i,j}^{t+1} = x_{i,j}^{t} + S_1 * Rand * \left(x_{r1,j}^{t} - x_{i,j}^{t}\right) + S_2 * Rand * \left(x_{r2,j}^{t} - x_{i,j}^{t}\right) \quad \dots \quad (11)$$

$$S_1 = \frac{\exp(f_i - f_{r1})}{(abs(f_i) + \varepsilon)}$$

$$S_2 = \exp((f_{r2} - f_i)) \quad \text{....} \quad (13)$$

Rand is a uniform random number between 0 and 1, and $\{r1 \in [1, \dots, N]\}$ is the index of the rooster, which is the ith hen's group-mate, while $\{r2 \in [1, \dots, N]\}$ is the index of the chicken (rooster or hen), which is selected at random from the swarm. $(r1 \neq r2)$.

Obviously, $(f_i > f_{r1}), (f_i > f_{r2})$, resulting in $(S_1 < 1 < S_2)$. If $S_1 = 0$, the ith hen will forage for food first, followed by the other chickens. The greater disparity between the two chickens' fitness values, the smaller S_2 and the greater distance between their positions. As a result, the hens will be less likely to snatch food from other chickens. There are competitions in a group, which is why the formula form of S_1 varies from that of S_2. The fitness values of the chickens in relation to the rooster's fitness value are simulated as competitions between chickens in a group for simplicity. If $(S_2 = 0)$, the ith hen will look for food within its own territory. The fitness value of the rooster is special in this category. As a result, the smaller fitness value of the ith hen, the closer S_1 approaches to (1) and the smaller difference between the ith hen's location and that of its group-mate rooster. As a result, the dominant hens are more likely to consume the food than the submissive hens.

To forage for food, the chicks move around their mother. This is how it's put together:

$$x_{i,j}^{t+1} = x_{i,j}^{t} + FL * \left(x_{m,j}^{t} - x_{i,j}^{t}\right) \quad \dots \quad (14)$$

Where $\left(x_{m,j}^{t}\right)$ represent the position of the i^th^ chick's mother $(m \in [1, N])$. $\left(FL \in (0,2)\right)$ is a parameter that indicates that the chick will forage for food with its mother. Taking into account individual variations, each chick's FL will fluctuate between 0 and 2.

(4-2) Selection of Improved Chicken Swarm Algorithm Parameters

The chicken swarm algorithm has six basic parameters that will be discussed, as humans primarily preserve chickens as a source of food, Chickens can also be used to provide the requisite eggs, which can be used as a food source. As a result, raising chickens is better for humans than raising roosters. As a result, (HN) would be higher than (RN), and since chickens have different personalities, not all hens will hatch their eggs at the same time. As a result, (HN) is greater than (MN), despite the fact that each chicken can raise several chicks. We assume that the number of adult hens will outnumber the number of chicks (CN), and that (G) should be set to a value that is suitable for the problem. If the value of (G) is very large, the algorithm will not be able to quickly converge to the best global level. The algorithm may be similar to the optimal local algorithm, despite the fact that (G) is very weak. It could be found inside $(G\ [2,20])$. Furthermore, if both (RN) and (MN) are set to zero, the formula for chick movement can be applied to the corresponding section in (DE), as well as through experiments. It is possible to calculate the value of $(DE = 2)$. If the parameter value (FL) is chosen within the range $\{FL \in [0.4,1]\}$, the algorithm performs well. [13]

(5) Use The Chicken Swarm Algorithm to Determine the Optimum Stratigraphic Boundaries for Geometric Stochastic Process

The chicken swarm optimization algorithm includes a number of basic steps, these steps are interrelated with each other, and this algorithm cannot be applied to any issue unless all of these steps are applied, otherwise the algorithm will lose its value and usefulness in finding or improving the solution. The proposed chicken swarm optimization algorithm to find the optimum stratigraphic boundaries for a threshold geometric stochastic process can explain as follow:

1. **Initial Data**: It is a reading of the data values for problem under study.
2. **Initialize data**: In the beginning, sudden fluctuations in the data are eliminated by taking the moving mean for stochastic series, and then the Moving Window Technology (MWT) is used, whose idea is to take a set of data with a specific size (30 observation for example) which represents displaying that window and finding the parameters for that group, then leaving the first value from the group and taking another value after the last value and calculating its parameters and continuing this process until the data is completed. After that, the differences between the ratio of the geometric stochastic process parameter (a_i) between each group and the next group calculated by applying equation No. (8), and in the event that the difference is greater than the specified threshold value, the specified point is considered a turning point.
3. **Find Inversion Points**: This step is one of the most important steps in the proposed method, and it is based on finding inflection points in the data and dividing the points into

increasing, decreasing and stable depending on the ratio of the geometric stochastic process parameter (a).

4. **Individual Representation**: The composition and form of the individual in the issue of stratification is two parts, the first part represents the inflection points that are selected from step (3) and at which the data is divided, and the second part represents the parameters of the geometric stochastic process that have been assessed for the Different stratification.
5. **Primary community creation:** The size of the algorithm group (number of particles) and the method for generating the primary population have a significant effect on the algorithm's output and the quality of the results. each individual must contain a number of sites equal to $((L - 1) + L * 2)$, where L represents the number of classes based on which the community is to be divided. In general, an ideal community should ensure the greatest possible diversity of individuals (solutions) in the search space.
6. **Objective Function**: In this step, we calculate the objective function for the problem and for each particle in the community, and the objective function in the problem of stratification is a criterion of Root Mean Square Error (RMSE), it is considered the criterion of the root mean square error (RMSE) is One of the most important criteria used in the comparison between the models used as well as in measuring the accuracy for the used model, as this criterion is used in the case of one-way data, as well as in the case of more than one trend, meaning there is a limit to the threshold between the data, and it is calculated through the following formula:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \hat{\mu}_k \hat{a}_K{}^{-(i-T_K)}\right)^2} \quad for \;\; T_k \leq i \leq T_{k+1} \qquad \dots \; (15)$$

7. **Applying chicken swarm optimization algorithm procedures (Applying CSO Procedures)**: After creating an elementary community of size (N_{pop}) individuals, the procedures of the chicken swarm optimization algorithm will be applied according to the equations shown on (9) to (14).

(6) Application Side:

(6-1) Covid-19 Data

Corona virus epidemic (Covid-19) is one of the most dangerous global health threats due to the ease of transmission as well as the long incubation period of the epidemic, which makes it one of the most widespread epidemics, the first case of the epidemic was recorded in Wuhan, China, on December 31, 2019 and began in the spread of a global epidemic. The first infection in Iraq was identified on February 24, 2020, in the city of Najaf, when a sample of an Iranian religious student was tested and the result was positive for his infection with Corona virus disease associated with extreme acute respiratory syndrome type II (SARS-CoV-2). Other cases of the Corona outbreak have been recorded in Iraq's various regions. Until December 31, 2020, the total number of confirmed total cases in Iraq was (595,291), whereas the total number of deaths in Iraq was (12,813).

Our research focuses on modeling data on daily infection numbers in Iraq for the Corona epidemic and for the three governorates (Baghdad, Erbil and Basra). Figure (1) shows the number of daily cases of the Corona virus epidemic in Iraq and Baghdad Governorate. Being increasing in a certain period (the period of the onset of the outbreak of the epidemic) and stable in another period (the period of taking the necessary measures to control the epidemic) in addition to decreasing in other periods (periods of recession of the epidemic to reach its fading out).

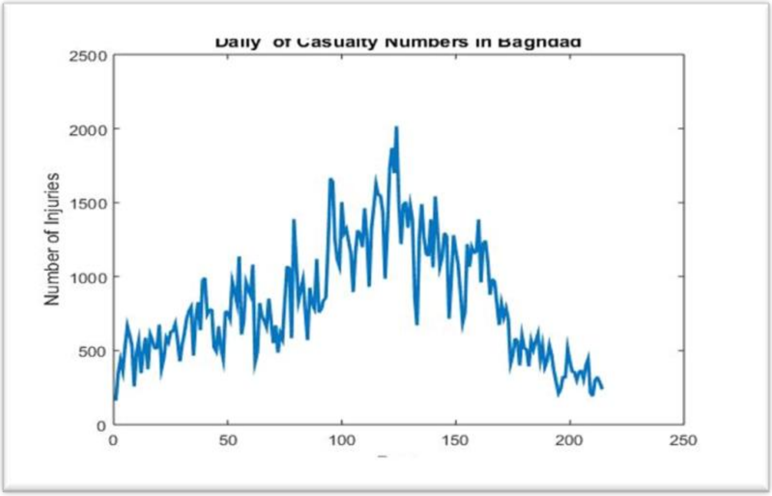

Figure (1)
The number of daily infected in Baghdad governorate

While Figure (2): shows the number of daily cases of the Corona virus epidemic in Iraq and Erbil Governorate. It is also evident by drawing the randomness of the data during the different periods and the passage of registration cases in cycles, each of which represents a change in the aggressiveness of the epidemic and its strength in infecting the largest possible number of people.

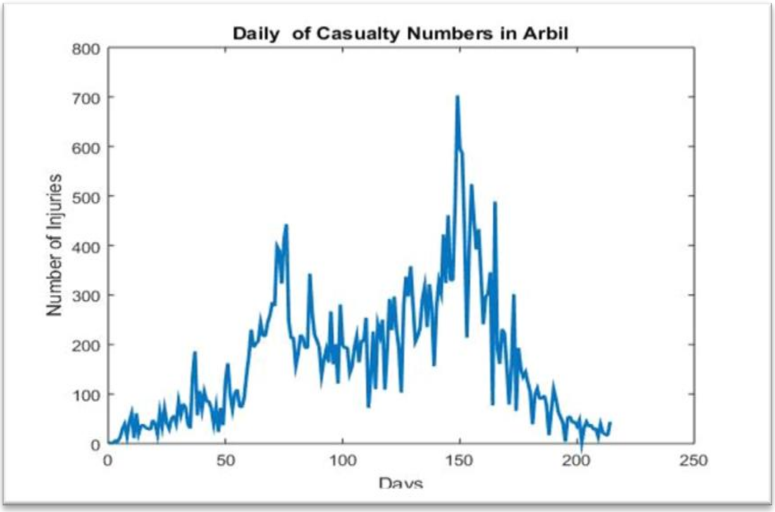

Figure (2)
The number of daily infected in Erbil Governorate

Figure No. (3) shows the number of daily cases the Corona virus epidemic in Iraq and Basra Governorate. As it is noticed through the figure as well as the passage of cases of Infected registration in the governorate in three stages, which is the increase, followed by stability, and then decreasing numbers of the data under study.

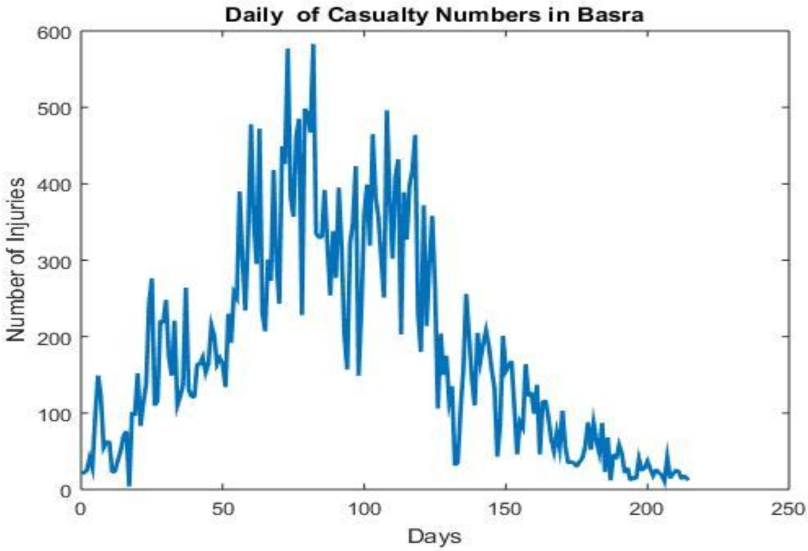

Figure (3)
The number of daily injuries in Basra Governorate

To illustrate more information the basic summary of the data under study, which represents the number of daily infected, the table (1) shows the basic information for data preparation of daily infected in epidemic covid-19 in Iraq, the three provinces (Baghdad, Erbil and Basra) for the period from 15/3/2020 to 31/12/2020, the first months from the date of recording the injury were neglected due to the randomness of the data, as well as the lack of equipment necessary for laboratory testing procedures.

Table (1)
Basic information of covid-19 data for the three regions

Governorate	Basra	Erbil	Baghdad
Start date	15/3/2020	15/3/2020	15/3/2020
End date	31/12/2020	31/12/2020	31/12/2020
n = number of data	214	214	214
Number of cases in total Sn	38980	35987	177628
Number of cases per day$\left({}^{Sn}/{}_{n}\right)$	182.149	168.163	830.037

(6-2) The procedures used to test stochastic processes of geometric form:

When applying the geometric shape stochastic process in a set of real data, the researcher faces several problems, the most important of which is measuring the appropriateness of data for this stochastic process. To solve this problem, the following procedures used to test the geometric shape stochastic process are applied by the following steps:

First: Test monotony trend in the data set

This test is the first process in analyzing monotone-trend data, and in order to test the monotonic trend in the data set, the graphical technique (GT) was used as follows:

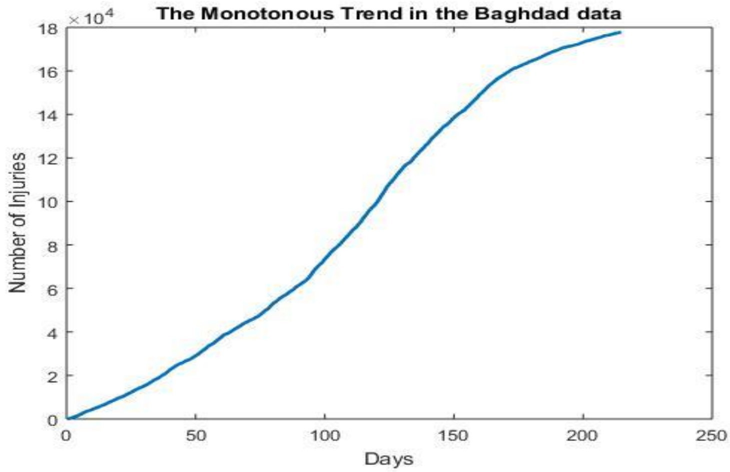

Figure (4)
Monotonic trend in daily infected figures data in Baghdad

From Figure (4), it is noticed that there is a monotonic trend in the data on the numbers of daily Coronavirus infections in Baghdad Governorate. Figure (5) shows the existence of a monotonic trend in the data on the numbers of daily Corona virus infections in Erbil Governorate.

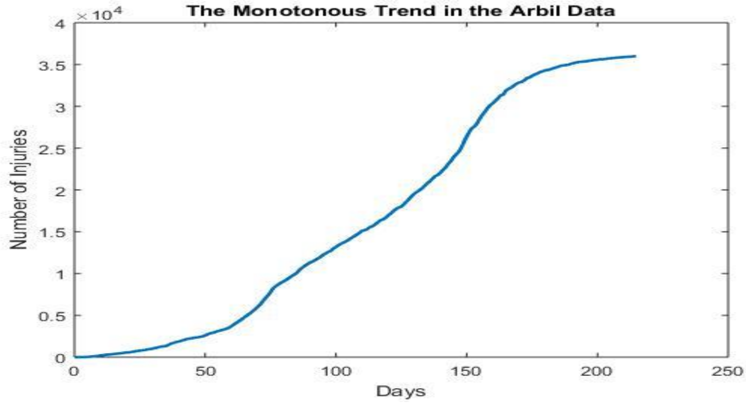

Figure (5)
Monotonic trend in daily infected figures data in Erbil Governorate

While Figure (6) shows the existence of a monotonous trend in the data on the numbers of daily Corona virus infections in Basra Governorate.

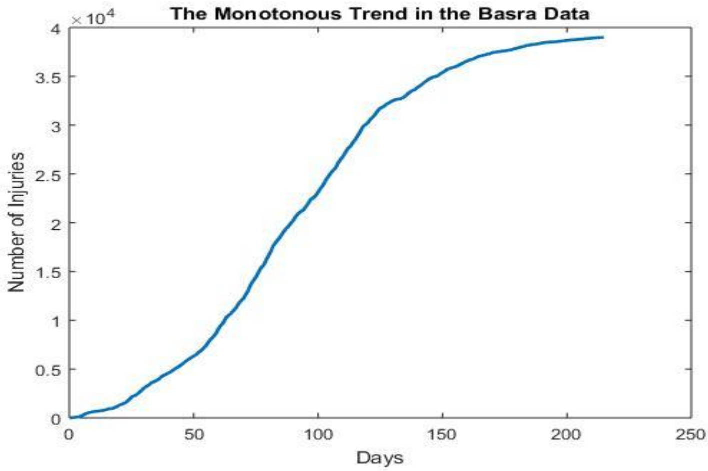

Figure (6)
Monotonic trend in daily casualty figures data in Basra

On this basis, this study relied on employing the advantages of evolutionary algorithms represented by the chicken swarm optimization algorithm in stratified random sampling and specifically in finding the optimal stratified limits depending on the method of the moving window to determine the threshold limit, as the stratigraphic boundaries were found for the data of the daily numbers of infections of the Corona epidemic (Covid-19) in Iraq for the three governorates (Baghdad, Erbil, Basra) using the Chicken Swarm Optimization (CSO) algorithm**.**

Second: Estimate and Determine the Optimal Boundaries

In this part of the research, a comparison will be made between the proposed model represented by the geometric stochastic process with an intelligent threshold and based on the algorithm of chicken swarm optimization to determine the stratigraphic boundaries and the ordinary geometric stochastic process (OGSP) model to reach the best model that represents the data under study, if a proposal is also made. Use of the chicken swarm optimization algorithm to estimate the parameters of the typical geometric stochastic process model. Table No. (2) represents the estimates of the intelligent threshold geometric stochastic process model parameters with using the chicken swarm optimization algorithm as well as the best stratification based on the best width of window (h) depend on the Root mean Squares Error (RMSE) criterion for the model.

Table (2)
The parameter estimates for the proposed model using chicken swarm optimization algorithm

Cities	Baghdad	Erbil	Basra
Number of layers(L)	3	3	3
width Window (h)	30	20	20
Class boundaries	164,173	58,186	126,194
$\hat{a}$	[0.994,0.997,1.186]	[0.99,1.0001,1.145]	[0.9964,1.01,1.2]
$\hat{\mu}$	[588.16,491.45, 566.59]	[154.66,241.39,689.44]	[197.99,700,233.28]
$RMSE$	297.94	99.01	99.18

While Table (3) represents the parameters estimates of the ordinary geometric stochastic process (OGSP) using the chicken swarm algorithm, as well as finding the value of the root mean squares error criterion for the model for a whole of the purpose model to comparing the two models and arriving the best models, which represent the best representation of the data under study.

Table (3)
The parameters estimate for the geometric stochastic process using the chicken swarm optimization algorithm

Cities	Baghdad	Erbil	Basra
$\hat{a}$	0.9999	0.9983	1.0028
$\hat{\mu}$	819.6147	139.4078	244.3201
$RMSE$	389.0015	130.2086	134.7548

While the following figures represent the extent to which the proposed models conform to the real data, which also obtained an idea about the preference of one model over another.

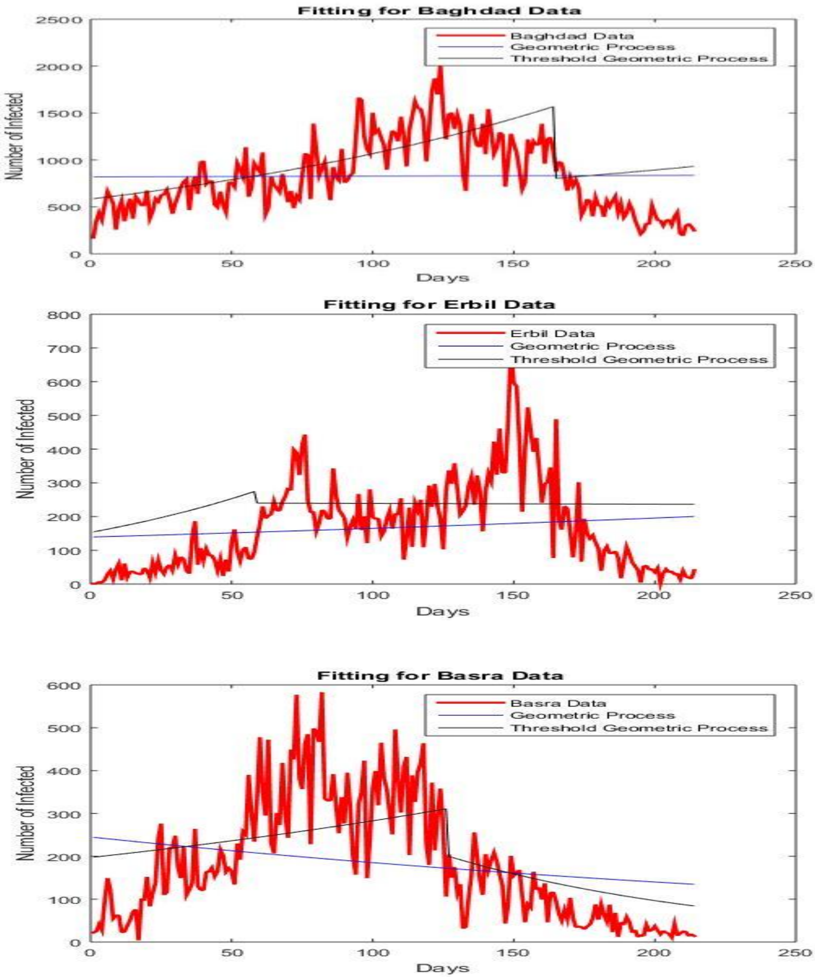

Figure (7)
represents the extent of fitting data in the three governorates and to the two stochastic models.

(7) Results Discuss:

Through Table (2), which shows the parameters estimation for the intelligent threshold geometric stochastic process. The results show that the data was divided into three layers, and these layers are the same that represent the behavior of any epidemic in terms of spread at the beginning of the period, and this is indicated by the value of $(a < 1)$ as the process is increasing, while the second layer indicates the stability of the process (that is, the start of controlling the epidemic), so we notice that the best model for this period can be represented by the Renewal Process (RP), because the parameter value $(a \cong 1)$, finally it can be observed that the behavior of data begin to fall and fade, and this is through the value of ratio geometric stochastic parameter $(a > 1)$.

By comparing the proposed model with the ordinary geometric stochastic process model, the results show that great prevalence of the proposed model, because it’s works on studying each period separately and evaluating its parameters, which results in accurate estimates and a high fitting for the original data, while the ordinary geometric stochastic process model fails to estimate thus, because the increase in the first period fades with the decrease in the last period, and the result is the average for all periods.

Therefore, the proposed model can be used in Multi Trend data, while the ordinary geometric stochastic process model is used for single trend data only.

(8) Conclusion:

The theoretical conclusions are considered an introduction to the interpretation of the researcher's perceptual framework, in addition to being a guide in strengthening the theoretical principle of the research and an extension of his theoretical assumptions, and in this way the most important conclusions (practical and theoretical) are presented as follows:

1. Through the study, it was found that the main cause of the outbreak of Coronavirus is the spread of infected people in the community, as it is the direct cause of many cluster infections in hospitals and local communities in Baghdad, Erbil and Basra.
2. The results indicate that the spread of the disease may be influenced by a variety of environmental factors such as temperature and humidity, as well as other external factors such as precautionary measures that may change over time.
3. We note through the theoretical superiority of the proposed model in the parameters estimation as well as in the fitting model with results showed superiority of the proposed model by 23.41% in the epidemic data modeling in the province of Baghdad, while the excellence of the epidemic data for the provinces of Basra and Erbil with rates of 23.96% and 26.40%, respectively. This indicates the validity of the theoretical postulates that been addressed.

(9) Recommendations:

The completion of the approved conclusions in building the content and content of the recommendations with the following:

1. Use other stochastic models for modeling data of the numbers of daily infections in the Corona virus epidemic. Among these models, the stochastic alpha-series process model and the (HPP) or (NHPP) after dividing the community into several layers and treating each layer separately according to the best model Represented.
2. Use another intelligent algorithm to determine the optimal class boundaries of the threshold geometric shape stochastic process model. These intelligent algorithms include: The bee

optimization algorithm, the ant optimization algorithm, and the blue whale optimization algorithm.

(10) References:

[1] Braun, W. J., Li, W. and Zhao, Y. Q.(2005),"**Properties of the Geometric and Related Processes**". Naval Research Logistics, 52, 607-616.

[2] Chan, J. S. K., Yu, P. L. H., Lam, Y. and Ho, A. P. K. (2006), **'Modeling SARS Data Using Threshold Geometric Process''**. Statistics in Medicine, 25, 1826-1839.

[3] Choi BCK, Pak AWP**." A simple approximate mathematical model to predict the number of severe acute respiratory syndrome cases and deaths**". Journal of Epidemiology and Community Health 2003; 57:831–835.

[4] Cheng, G. Q. and Li, L. (2011), **''Two Different General Monotone Process Models for a Deteriorating System and Its Optimization '',** International Journal of System Science ,42:1,57-62.

[5] Cox, D.R. and Lewis, P.A. (1966), **''Statistical Analysis of Series of Events"**. Chapman and Hall, London, United Kingdom.

[6] Lam, Y. (1992b), **''Nonparametric Inference for Geometric Processes''**. Communications in Statistics Theory and Methods, 21, 2083-2105.

[7] Lam, Y., Zhu, L. X., Chan, S. K. and Liu, Q. (2004), **''Analysis of Data from a Series of Events by a Geometric Process Model''**. Acta Mathematica Applicate Sinica, 20, 263-282.

[8] Lam, Y. (2007a), **''The Geometric Process and its Applications''**. World Scientific, Singapore.

[9] Malti Baghel, Shikha Agrawal, and Sanjay Silakari, “**Survey of Metaheuristic Algorithms for Combinatorial Optimization**", International Journal of Computer Applications, Vol.58 (2012), pp.21-31.

[10] R. Andrea, M. Blesa, C. Blum, and S. Michael, “**Hybrid Metaheuristics an Emerging Approach to Optimization"**, (Springer, 2008).

[11] Slaman,Muthana S.,Jasim,Omar R.(2013), “Comparison of the geometric stochastic process with a related process with an application” ,Iraqi Journal of Statistical Sciences.

[12] Xianbing Meng, Yu Liu, Xiaozhi Gao, and Hengzhen Zhang, **“A New Bio-Inspired Algorithm: Chicken Swarm Optimization**", Vol.8794 (2014), pp.86-94.

[13] Xin-She Yang," **Nature-Inspired Metaheuristic Algorithms 2edn**", Beckington, Uk (Luniver Press, 2008), pp.242-46.