

Estimación de parámetros, análisis de sensibilidad y validación del modelo SIR aplicado a datos históricos del virus A(H1N1) en México

Hamilton Gómez Osorio
hsgomezo@eafit.edu.co
Universidad EAFIT
Medellín, Colombia

Santiago Isaza Cadavid
sisazac@eafit.edu.co
Universidad EAFIT
Medellín, Colombia

Pablo Osorio Marulanda
paosorion@eafit.edu.co
Universidad EAFIT
Medellín, Colombia

10 de junio de 2021

Resumen

Este trabajo busca analizar el comportamiento epidemiológico de una enfermedad transmisible como la influenza A(H1N1), a partir de un modelo SIR, aplicado a un grupo de personas diagnosticadas con la enfermedad en México en el año 2009. En este se realiza una respectiva estimación de parámetros poblacionales, como lo es la tasa de contagio y la tasa de recuperación, y un análisis de sensibilidad utilizando la herramienta Simulink de MATLAB.

Palabras clave: SIR, A(H1N1), Simulink, Epidemiología

Abstract

This paper seeks to analyze the epidemiological behavior of a communicable disease such as influenza A (H1N1), based on an SIR model, applied to a group of people diagnosed with the disease in Mexico in 2009. Making a respective estimate of Population parameters, such as the contagion rate and the recovery rate, and a sensitivity analysis using the MATLAB Simulink tool.

Keywords: SIR, A(H1N1), Simulink, Epidemiology

1. Introducción

La construcción de modelos matemáticos es en la actualidad una de las herramientas más utilizadas para la descripción, explicación y predicción de sistemas dinámicos. La función principal de un modelo para una enfermedad infecciosa consiste en proveer un medio que posibilita entender la dispersión de la enfermedad a través de una población bajo diferentes escenarios. Los modelos matemáticos para enfermedades infecciosas permiten extraer de ellos propiedades y características de las relaciones entre los elementos que podrían permanecer ocultas a simple vista. Además, en la mayor par-

te de los problemas de enfermedades infecciosas del mundo real, no es factible experimentar con la realidad, ya que puede ser muy costoso, peligroso, inmoral o incluso imposible. Por lo tanto, es natural intentar superar esta dificultad con la construcción de un modelo que describa de manera adecuada las características básicas de la epidemia, y entonces usar el modelo para predecir las consecuencias de introducir cambios específicos [1].

En general, los modelos epidemiológicos se basan en dividir a la población expuesta a la infección en pequeños grupos compartimentados, cada uno formado por individuos con un idéntico estado con respecto a la infección

en cuestión. El modelo implementado en este paper es el modelo SIR, propuestos por el epidemiólogo y físico Anderson Gray McKendrick y el matemático William Ogilvy Kermack [7], el cual es uno de los modelos epidemiológicos más simples y al mismo tiempo capaz de capturar las características típicas de los brotes infecciosos. Este modelo tiene en cuenta un grupo de Población susceptible (S), Población infectada (I), Población recuperada (R). Estos conforman lo que se conoce como Población total (N).

La influenza A(H1N1), o influenza porcina, es un virus que se le atribuye a los cerdos y que ha sido muy común en el último siglo ocasionando fuertes infecciones respiratorias alrededor del mundo. Esta enfermedad es un tipo de gripe la cual comienza al entrar en contacto directo con cerdos o con alguna superficie que tenga el virus. Las principales vías de infección son la nariz, la boca y los ojos, y esta puede ser transmitida de persona en persona. La influenza porcina es una enfermedad que aunque controlada pueden desaparecer sus síntomas luego de una semana, en el caso de no ser tratada, dado a sus síntomas como usualmente son diarrea, náuseas, dolor de garganta, tos, entre otros, pueden convertirse en enfermedades más graves como neumonía o insuficiencias respiratorias, lo que es mucho más riesgoso en niños y adultos mayores de 65 años [6] [2].

Dada la magnitud de las epidemias y las implicaciones que estas pueden tener en una población, esta investigación tiene como objetivo estudiar el comportamiento de infestación de la influenza A(H1N1) en un grupo diagnosticado con la enfermedad en México en el año 2009 por medio del modelo epidemiológico SIR que represente el fenómeno de infección observado en esta población y, adicionalmente, hacer una comparación respecto a estudios de esta enfermedad en la población mexicana para estos periodos. Para esto se utilizan múltiples fuentes como estudios anteriores y desarrollos teóricos del modelo implementado. Se realiza una simulación del problema con su respectiva estimación de parámetros, entregando como principal aporte un análisis de sensibilidad, y una evaluación del ajuste del modelo al comportamiento monitoreado del virus a través de validación. Finalmente se concluye a partir de los resultados encontrados.

2. Metodología

2.1. Descripción del sistema

Para el año 2009 se presentaron cifras alarmantes de pacientes con neumonía y enfermedades respiratorias en los diferentes hospitales de México entre finales del mes de febrero y principios de marzo. Este fenómeno alertó a las autoridades de salud y rápidamente empezaron una investigación para esclarecer las razones de esta problemática. Situación similar se venía presentando en los países norteamericanos en los que múltiples pacientes tenían síntomas de lo que parecía ser una enfermedad respiratoria grave. México, luego de tomar diferentes muestras quiso verificar la compatibilidad del virus que se presentaba en su país con las patologías presentadas en estos otros estados, por lo cual decidió analizar sus muestras en Canadá y la relación de las enfermedades quedó confirmada, enfermedad conocida como influenza A(H1N1), que para ese momento ya presentaba casos en Europa.

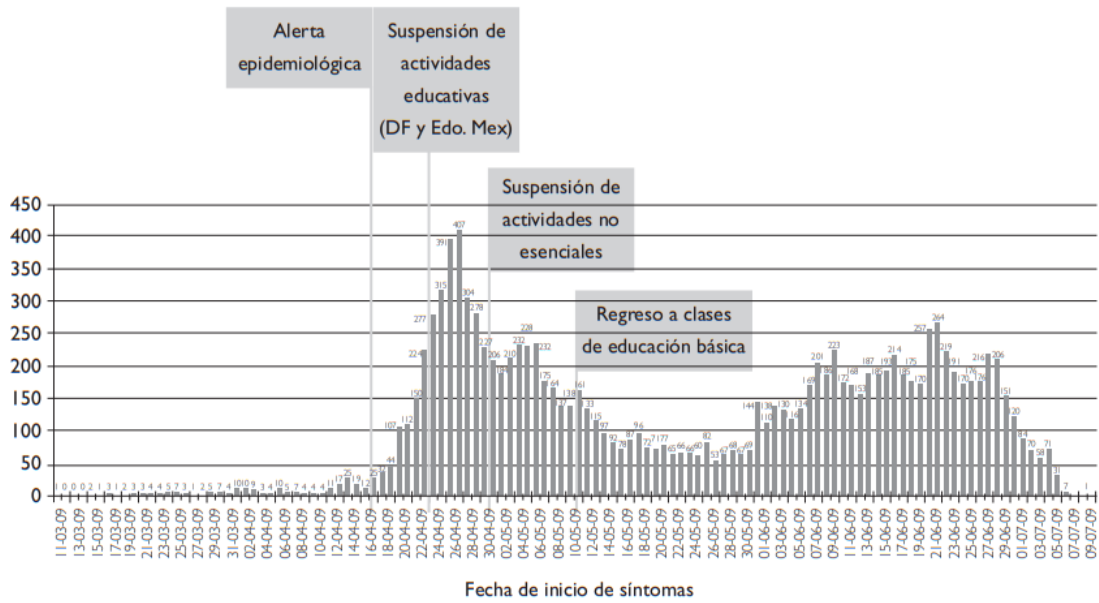
Transcurrido Abril los casos de esta gripe estaban en disminución en Estados Unidos por lo que se esperaba un comportamiento similar en los casos mexicanos; lo que no se esperaban es que los casos en México siguieran en aumento y que estos, contrario a lo conocido, se dieran regularmente en pacientes entre edades de 5 a 14 y 25 a 44 años. Luego de una observación detallada de los casos se pudo ver que esta población estuvo más afectada al no ser vacunados oportunamente y el estado mexicano promovió las campañas de prevención y estudio de dicha enfermedad en el país [4].

2.2. Métodos experimentales

Según estudios realizados por el Instituto Nacional de Salud Pública de México sobre los casos de infección y mortalidad de personas con influenza A(H1N1), se evidenció un muestreo de nuevos casos registrados cada dos días de pacientes con diagnóstico positivo para este virus entre los meses marzo y junio. Dichos datos fueron tomados antes de que se iniciara la alerta epidemiológica registrada en abril y se extendió hasta junio para ver su progresividad en el tiempo dado el aumento de casos registrados. Los datos experimentales tomados por esta entidad se muestran en la Figura 1.

2.3. Descripción del modelo matemático

El modelo SIR tiene en cuenta tres grupos: **Población susceptible (S)**, refiriéndose a los



Fuente: Base de datos INDR. N = 12 645 casos confirmados

Figura 1: Casos de pacientes diagnosticados con influenza A(H1N1) en México entre marzo y junio de 2009.

individuos que pueden ser infectados por el virus; **Población infectada (I)**, refiriéndose a los individuos que ya fueron infectados por el virus y pueden transmitirlo a individuos de la población susceptible con los que entren en contacto; y finalmente la **Población recuperada (R)**, que se refiere a los individuos que adquieren la inmunidad al virus y, por lo tanto, no afectan la transmisión al interactuar con otros individuos. La **Población total (N)** es

$$N = S + I + R$$

[3].

El modelo en ecuaciones diferenciales está planteado como sigue:

$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \alpha I$$

$$\frac{dR}{dt} = \alpha I$$

Los parámetros β y α , se refieren a la tasa de propagación de la enfermedad (tasa de infección) y a la tasa de recuperación (tasa de alivio), respectivamente.

2.4. Diagrama de bloques

Se utilizó la herramienta Simulink de MATLAB para simular el problema planteado.

Para esto se hizo el respectivo diagrama de bloques del sistema de ecuaciones diferenciales representado en la Figura 2.

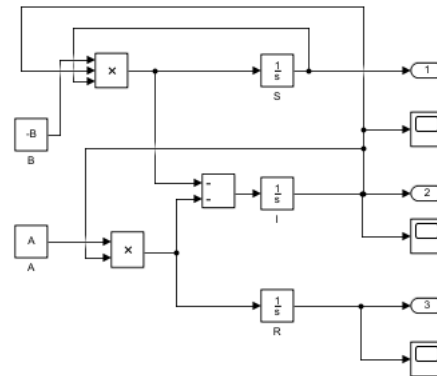


Figura 2: Diagrama de bloques en Simulink

2.5. Métodos de simulación, identificación y análisis

Para la simulación del modelo trabajado utilizamos el método de integración de Euler(ode1) con un tamaño de paso fijo de 0.01.

Para los valores iniciales considerados, tales como la tasa de infección(β), tasa de recuperación(α) y el número de individuos inicial (N), se escogieron valores aleatorios entre 0 y 0.1, al ser promedios. Además, teniendo en

cuenta que estas tasas son bajas, y el valor de S_0 se encuentra entre 0 y 10000, respecto al intervalo posible de la cantidad de personas que trabajaban con cerdos en México y podrían estar susceptibles en ese momento al virus. Este último corresponde al intervalo considerado en trabajos anteriores [5]. Los valores iniciales obtenidos y posteriormente usados fueron:

$$S(0) = 1034$$

$$\alpha = 0,0501$$

$$\beta = 0,0005$$

Para estos valores se obtuvo el comportamiento inicial de la población de susceptibles, infectados y recuperados, ilustrado en la Figura 3, 4 y 5, respectivamente.

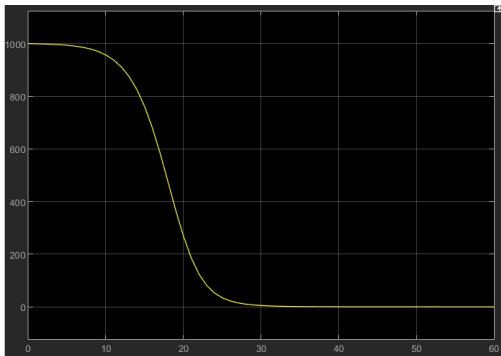


Figura 3: Comportamiento de los susceptibles con condiciones iniciales

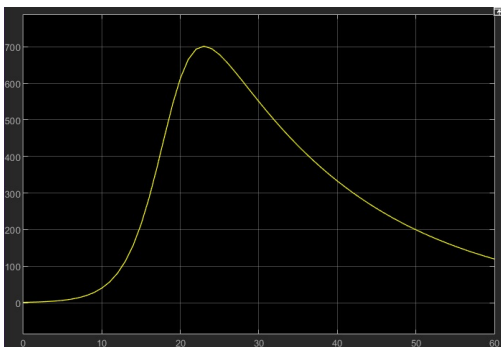


Figura 4: Comportamiento de los infectados con condiciones iniciales

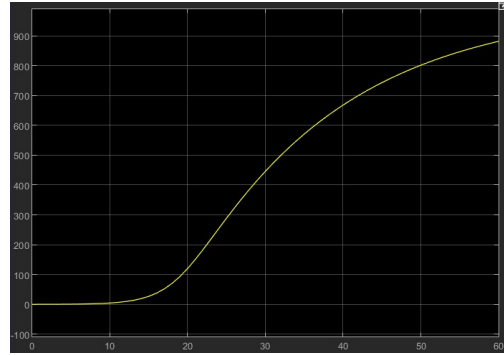


Figura 5: Comportamiento de los recuperados con condiciones iniciales

3. Resultados

3.1. Identificación de parámetros y análisis de sensibilidad

El análisis de sensibilidad se realiza generando 50 valores aleatorios con distribución uniforme por el método de Montecarlo para ambos parámetros, los cuales para el caso de α se encuentra acotada entre $[0,0645, 0,08]$. Para el caso de β se encuentra entre $[0,0017, 0,0021]$. Adicionalmente el análisis se realizará aplicando el método de correlación, regresión estandarizada y correlación parcial.

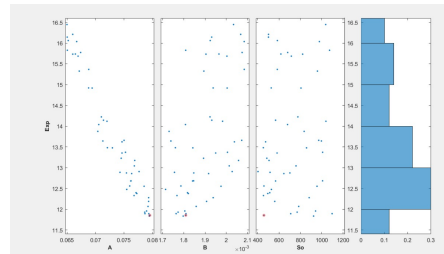


Figura 6: Valores aleatorios generales

La Figura 6 indica que, en efecto, el intervalo otorgado se acerca mucho al punto óptimo de estimación. Para este análisis, los valores de mejor ajuste para los parámetros fueron de:

$$\alpha = 0,0624$$

$$\beta = 0,0028$$

$$S_0 = 458,6749$$

Con una estimación cerca del valor 11,85.

Estos valores serán evaluados en el modelo y comparados con los datos experimentales.

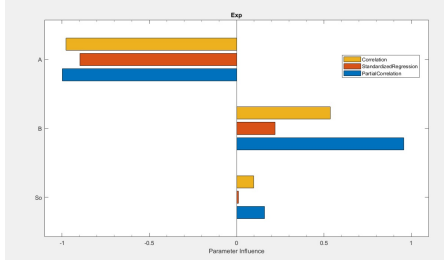


Figura 7: Análisis de sensibilidad para los parámetros α , β y S_0

La Figura 7 da una idea de cuáles serían los parámetros que mayor efecto tendrían en el modelo. Para este caso en concreto, respecto los tres métodos de evaluación, el parámetro (β) es el que más modificaría la estructura y el comportamiento del sistema.

Para la estimación de parámetros se dará un tratamiento al modelo en el que se busca que la curva respectiva a los susceptibles (casos observados de la enfermedad) se comporte de manera respectiva a los datos experimentales escogidos.

Los datos experimentales mencionados se presentan en la Figura 8.

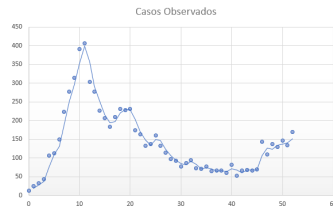


Figura 8: Distribución de los casos confirmados de influenza

Estos representan la distribución de los casos confirmados de influenza A(H1N1) de acuerdo a la fecha de inicio de los síntomas. Los datos comprenden desde el 16 de abril de 2009 hasta el 7 de junio del mismo año, donde cada una de las observaciones hace referencia al total observadas en cada dos días de este intervalo. El modelo fue nuevamente evaluado con los datos que lanzaron el mejor ajuste en el análisis de sensibilidad. Los resultado se presentan en la Figura 9.

Para encontrar los valores que mejor se ajustan se realizó un proceso de optimización buscando así minimizar el valor de la suma de errores cuadrados obtenidos en cada iteración. Así, se obtuvieron nuevos valores óptimos para los parámetros. En la Figura ?? se pueden apreciar los resultados.

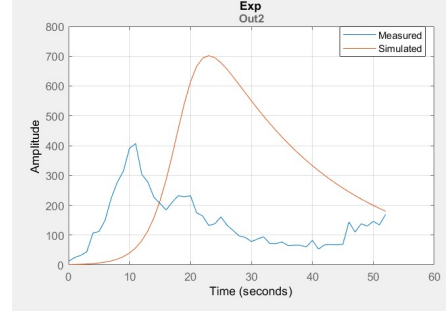


Figura 9: Relación entre los datos teóricos y experimentales iniciales

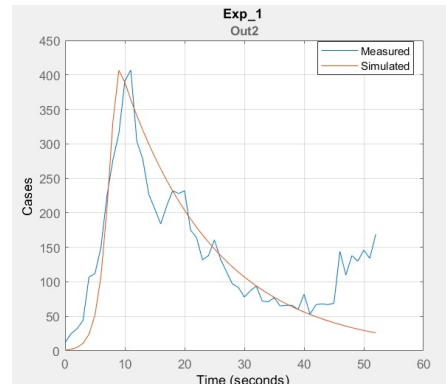


Figura 10: Relación entre los datos teóricos y experimentales óptimos

Donde:

$$S(0) = 458,68$$

$$I(0) = 12$$

$$R(0) = 0$$

$$\alpha = 0,0624$$

$$\beta = 0,0028$$

$$SSE = 0,86$$

Las iteraciones respectivas para cada uno de los parámetros se realizó como se muestra en la Figura 11.

Se entiende entonces que, dado el análisis de sensibilidad, las iteraciones para la optimización son mucho menores.

Con los valores estimados es posible entonces determinar el comportamiento de los demás estados, esto es, los individuos susceptibles y los individuos recuperados.

La Figura 12 representa los individuos susceptibles a través del tiempo y la Figura 13 representa los individuos recuperados a través del tiempo.

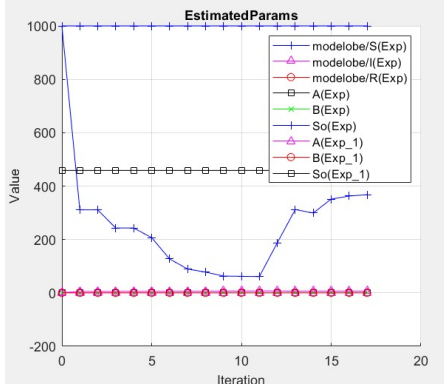


Figura 11: Iteraciones en el proceso de optimización

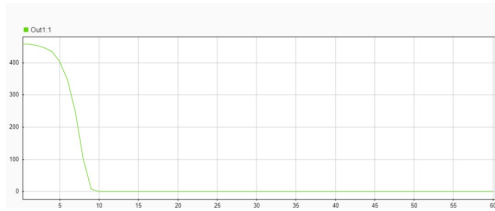


Figura 12: Individuos susceptibles

3.2. Validación del modelo

Para la validación del modelo se obtienen los residuales, por medio de los valores experimentales y aquellos hallados a través de el modelo con los parámetros estimados. Estos se presentan en la Figura 14.

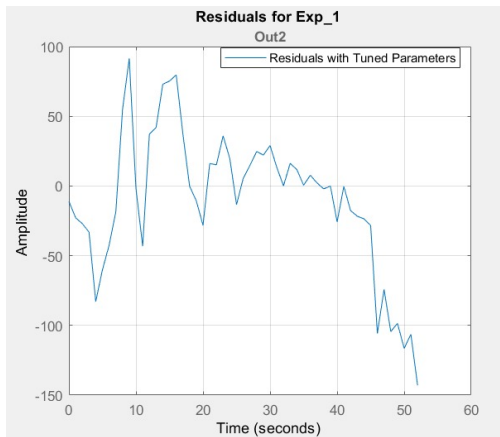


Figura 14: Gráfico de los residuales del modelo

Con base en datos obtenidos teóricamente y los obtenidos experimentalmente, el coeficiente de correlación de Pearson es:

$$Coe_f_R^2 = 0,4766$$

Adicionalmente, el coeficiente de correlación para los valores experimentales y teóricos es de:

$$Coe_f - corr = 0,6904$$

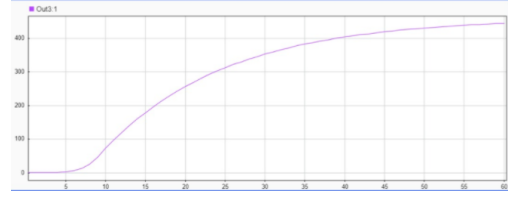


Figura 13: Individuos recuperados

3.3. Otros análisis y resultados

Otro análisis que se consideró fue una comparación con un trabajo similar, que también evalúa un modelo SIR de la influenza A(H1N1) [5]. En este trabajo se utilizaron diferentes funciones de MATLAB con un método de integración ode45 y con intervalos para las tasas entre 0 y 100. Adicionalmente, se consideró una población susceptible entre 0 y 10000. Luego se hizo todo el proceso de optimización y obtuvieron los siguientes valores óptimos:

$$S(0) = 514$$

$$I(0) = 1$$

$$R(0) = 0$$

$$\alpha = 0,0812$$

$$\beta = 0,0011$$

Comparado con los valores que se obtuvieron en el presente trabajo, se puede ver que aunque la estimación de parámetros para el estudio realizado en el 2009, donde se utilizó un intervalo de estimación mucho mayor, un método de integración diferente y además se utilizaron funciones establecidas por MATLAB para estos procesos, los resultados finales son muy similares. Esto indica que el modelo es muy estable, y por consiguiente que los parámetros estimados son correctos.

4. Conclusiones

El Modelo SIR es un modelo adecuado para simular el comportamiento de una enfermedad en el tiempo dado que estos se ajustan a los datos experimentales. Las epidemias son muy difíciles de controlar sin la tasa de infectados crece mucho ya que la recuperación es más lenta que la infección. La estimación de parámetros es muy sensible a las diferencias para algunos valores, por eso es que dependiendo del tipo de métodos que se utilice para estimarlos las diferencias entre los resultados de los modelos pueden ser muy diferentes. A pesar de que la estimación de parámetros puede ser muy sensible, el proceso de optimización hace que los resultados sean mucho más parecidos ya que se están encontrando

valores más ajustados.

Aunque el modelo SIR se ajuste a las fechas para las que se evaluaron los datos experimentales, es claro que el modelo es poco sensible y no se alcanza a adaptar a cambios posteriores ocasionados por factores adversos, por ejemplo, aquellos casos ocurridos desde el 26 de mayo de 2009. Por esto, para futuros desarrollos es recomendado tener en cuenta el fenómeno de manera estacionaria.

Es importante tener en cuenta que en este modelo no hay retroalimentación, pues los recuperados siempre serán recuperados y no pasan a otros estados, además de que la población de susceptibles S_0 no está tan bien definida en su intervalo, pues en estudios anteriores, no se especifica si esta población es rural o urbana y cuántos de ellos pudieron estar verdaderamente expuestos a condiciones de infección.

Referencias

- [1] Niels Becker. The uses of epidemic models. *Biometrics*, pages 295–305, 1979.
- [2] Mayo clinic. Influenza porcina: (influenza h1n1), Recuperado en noviembre, 2019.
- [3] D. S. Fernández del Viso. Modelos mecanísticos determinísticos con ecuaciones diferenciales, Recuperado en noviembre, 2019.
- [4] Germán E Fajardo-Dolci, Francisco Hernández-Torres, Javier Santacruz-Varela, Javier Rodríguez-Suárez, Philippe Lamy, Heberto Arboleya-Casanova, Rafael Gutiérrez-Vega, Gabriel Manuell-Lee, and José Ángel Córdova-Villalobos. Perfil epidemiológico de la mortalidad por influenza humana a(h1n1) en México. *salud pública de México*, 51(5):361–371, 2009.
- [5] Juan Camilo Vargas Jerez, Helbert Eduardo Espitia Cuchango, and Carlos Humberto Galeano Urueña. Sir y seir empleando optimización para el modelamiento de la transmisión de la gripe a (h1n1). *Ingenium*, (21), 2010.
- [6] Rama Judicial. Preguntas y respuestas: Virus a(h1n1), Recuperado en noviembre, 2019.
- [7] Julián Mauricio Fajardo Patino. Modelos determinísticos y estocásticos si y sir para difusión de enfermedades contagiosas. 2012.