

Primeros pasos en la implementación de modelos epidemiológicos con metapoblaciones

Lázaro Daniel González Martínez¹

Abstract

Análisis de las dinámicas de movimientos euleriano y lagrangiano. Introducción de estos en el modelo epidemiológico clásico SIR. Implementación de modelos epidemiológicos con metapoblaciones. Comparación y análisis crítico dentro de los modelos clásicos, de los modelos con metapoblaciones, y entre ambos.

Palabras Clave

modelos epidemiológicos, metapoblaciones, dinámica, Runge-Kutta

¹ Facultad de Matemática y Computación, Universidad de la Habana, La Habana, Cuba, lazarodaniel.glezmtnz@gmail.com

1. Introducción

La Epidemiología es el estudio de la distribución y los determinantes de la prevalencia de enfermedades en humanos; entre sus funciones se encuentran describir la distribución de la enfermedad, identificar las causas o los factores de riesgo de las enfermedades, construir y probar teorías, y planificar, implementar y evaluar programas de detección, control y prevención, según se explica en [1]. Una forma de modelar epidemias es a través de ecuaciones diferenciales. Los modelos epidemiológicos clásicos tienen una característica en común, y es que asumen generalmente homogeneidad en la población. De ahí que factores como la tasa de contagio y de recuperación del individuo sean tratados uniformemente para todos.

Sin embargo algunas variaciones de estos modelos han incluido el empleo de metapoblaciones para modelar las epidemias. Una metapoblación es un grupo de poblaciones de la misma especie que viven en áreas espacialmente aisladas pero que interactúan en algún nivel [2]. En otros términos nos referimos a localidades separadas geográficamente cuya interacción es mediante la migración. Cuando se añaden los factores de aislamiento espacial y migración a un modelo epidemiológico este tendrá en cuenta la movilidad entre las zonas en estudio y la heterogeneidad de la población en general. Esto brinda mayor robustez a la modelación pero como cualquier modelo, no todo es ventajoso. Hay aspectos computacionales que se deben tener en cuenta como el aumento de parámetros, ecuaciones, complejidad temporal, entre otros, así como la información que se maneja, las interpretaciones previas y posteriores a la solución del modelo, etc.

Esta investigación tiene como objetivos:

- Hacer un estudio de los modelos clásicos SIR, SEIR y SIS, los modelos de dinámica de movimientos eulerianos y lagrangianos, y el modelo epidemiológico SIR con metapoblaciones.
- Ofrecer una primera implementación funcional de los modelos anteriores, teniendo en cuenta la eficiencia de estos dentro de lo posible, principalmente el orden de la complejidad temporal.
- Comparar soluciones tanto de los modelos epidemiológicos clásicos (MEC) entre ellos, los modelos de movimientos (MM), los modelos epidemiológicos con metapoblaciones (MEM), así como realizar comparaciones MEC-MEM.

Señalar además, que este reporte constituye la antesala de un Trabajo de Diploma, que se propone hacer una implementación eficiente de los modelos epidemiológicos con movimiento euleriano y lagrangiano, así como la estimación de los parámetros de estos. Aunque este último punto no se trata con profundidad en este reporte, se recomiendan acciones a tener en cuenta en el futuro en varios sentidos. Los principales artículos empleados fueron [3] y [4] los cuales cuentan con los modelos clásicos y los modelos con metapoblaciones respectivamente. En ambos se analizan profundamente desde una perspectiva más analítica y matemática estos modelos, como sería las aproximaciones de los valores de R_0 , el tamaño de la epidemia, la máxima prevalencia, entre otros. Pero en este trabajo nos enfocaremos más en los detalles propios que marcan las diferencias entre estos tipos de modelos. Aunque estos sean los artículos centrales, también se tienen en cuenta otros.

2. Modelos Epidemiológicos Clásicos

En la bibliografía clásica sobre Matemática Aplicada en la Epidemiología destaca el uso de los modelos compartimentales, para modelar epidemias; en estos modelos se divide en clases epidemiológicas a la población bajo estudio, y se describe el flujo de transiciones entre estas [5]. Existen varios modelos para emular una pandemia, pero como hemos aclarado anteriormente, solo haremos un pequeño análisis de los modelos SIR, SEIR y SIS.

2.1 Modelo SIR

El Modelo SIR es el clásico modelo para epidemias, que considera 3 clases poblacionales. Los *susceptibles*: las entidades que pueden adquirir la enfermedad; los *infestados*: aquellos que ya tienen la enfermedad y pueden transmitirla; los *recuperados*: los individuos que ya no portan la enfermedad y ganan inmunidad a esta. En este último caso suelen interpretarse las muertes a causa de la enfermedad como *recuperados*. El flujo entre clases es bastante cercano a la realidad de manera general. La tasa de crecimiento de *infestados* está favorecida por la cantidad de contactos entre *infestados* y *susceptibles* principalmente, mientras que los *recuperados* crecerán a medida que los *infestados* pasen la enfermedad. En este caso el modelo epidemiológico podría quedar de la siguiente forma:

$$\dot{S} = -\beta \frac{SI}{N} \quad (1)$$

$$\dot{I} = \beta \frac{SI}{N} - \gamma I \quad (2)$$

$$\dot{R} = \gamma I \quad (3)$$

Se debe destacar que por lo general en este modelo se considera $\beta^* = \frac{\beta}{N}$, en vez del parámetro β que utilizamos. Este cambio es posible ya que β y N son constantes, y resulta práctico para comprender mejor las modificaciones que se hagan a estos modelos al añadir metapoblaciones. Note que N hace referencia al tamaño de la población que se estudia la cual se considera constante, $\frac{\beta}{N}$ es la tasa de contagio y γ la tasa de recuperación

2.2 Modelo SEIR

Este modelo añade una nueva clase al modelo SIR: los expuestos. Esta clase simula el tiempo de incubación de la enfermedad de un paciente, donde aún no posee la enfermedad ni la puede transmitir. En este caso se añade la variable $E(t)$ que indica la cantidad de expuestos que hay en el instante de tiempo t . De aquí podemos deducir el modelo:

$$\dot{S} = -\beta \frac{SI}{N} \quad (4)$$

$$\dot{E} = \beta \frac{SI}{N} - \sigma E \quad (5)$$

$$\dot{I} = \sigma E - \gamma I \quad (6)$$

$$\dot{R} = \gamma I \quad (7)$$

En este caso se incorpora un nuevo parámetro σ representando la tasa de incubación de la enfermedad. Note que esta vez el aumento de infestados se debe a la cantidad de personas que van rebasando el tiempo de incubación. Este modelo tiene sentido ser considerado para enfermedades como la viruela o la varicela que tienen entre una y tres semanas de incubación. En [6] se propone un modelo de este tipo con la peculiaridad de separar a los recuperados en 2 subgrupos: los recuperados vivos y los muertos por la enfermedad. Además que le incorporan metapoblaciones y utilizan los datos que se tenían hasta el 16 de abril, para estimar en regiones de Ohio y Michigan, EEUU, los parámetros del modelo y realizar predicciones. Otra variante de este modelo incluye la prevalancia del patógeno de la enfermedad en el ambiente y cómo los susceptibles presentan otra forma de contagio [7].

2.3 Modelo SIS

Por otra parte este modelo elimina del modelo SIR la clase de recuperados. En este caso se asume que los individuos una vez que superan la enfermedad no ganan inmunidad a esta, sino que vuelven a ser susceptibles. El modelo entonces puede ser:

$$\dot{S} = -\beta \frac{SI}{N} + \gamma I \quad (8)$$

$$\dot{I} = \beta \frac{SI}{N} - \gamma I \quad (9)$$

$$(10)$$

Note que la diferencia radica en que los susceptibles ahora aumentan a medida que los infestados se curan.

3. Modelos Epidemiológicos con Metapoblaciones

Los modelos compartimentales ya vistos, no tienen en cuenta una característica primordial para describir la enfermedad: la movilidad. En su lugar asumen que se puede tener un contacto con todas las personas. Por lo que la determinación de β está dada por la probabilidad de contacto con otra persona sin importar su condición. Sin embargo los modelos basados en metapoblaciones tienen en cuenta la movilidad de las distintas clases compartimentales por regiones. Además una característica de estos modelos es que permiten reflejar cierta heterogeneidad en la población, ya que algunos de los parámetros de los modelos clásicos como la probabilidad de contagio y la tasa de recuperación pueden particularizarse para cada una de las regiones. Recordemos que en los modelos explicados en secciones anteriores solo se tiene en cuenta la transición entre clases y la población se considera homogénea, de lo cual se puede inferir que todas las personas susceptibles tienen la misma probabilidad de ser contagiadas por cualquier persona infestada.

Estos modelos por lo general se trabajan con grafos dirigidos, que ilustran la movilidad de la población en diferentes regiones. Entiéndase movilidad como la migración de personas de una zona a otra, en sus diferentes variantes. En la práctica las regiones por la que puede transitar una persona serán representadas con los nodos del grafo, mientras que la tasa de movilidad de una región a otra, será representada por las aristas entre estos nodos. Formalmente podemos definir estos grafos con una o más funciones de costo asociadas a la tasa de movilidad entre los nodos.

Primero debemos entender algunos de los modelos de movimiento que existen independientes de la modelación epidémica. En este trabajo solo tendremos en cuenta dos variantes: movimiento euleriano y movimiento lagrangiano. Para hacer más amigable la explicación a partir de ahora consideraremos como población a las personas que habitan un país, y los municipios como las zonas en que estas se mueven. E indistintamente haremos uso de *municipio* y *nodo* como conceptos iguales.

3.1 Movimiento Euleriano

El movimiento euleriano intenta reflejar la migración indefinida de la población entre municipio y municipio. Es decir, el movimiento de personas de una municipio a otro por una temporada grande de tiempo, sin retorno o con retorno esporádico. En la práctica esto se visualiza generalmente con las mudanzas que puede realizar una familia, o el alquiler de rentas para vivienda, o incluso las visitas prolongadas que hacemos a familiares en otras provincias por problemas personales. En resumen, es el movimiento que realiza una persona hacia otro nodo, por una temporada indefinida de tiempo.

Para el modelo matemático debemos definir la cantidad de nodos K y una matriz de movilidad $F_{K \times K}$ tal que f_{ij} representa la tasa de movilidad del nodo i hasta el nodo j es decir la migración unilateral desde la localidad i hasta la j . Luego este modelo quedaría:

$$\dot{N}_i = - \sum_{j=1}^K f_{ij} N_i + \sum_{j=1}^K f_{ji} N_j \quad (11)$$

La tasa de cambio de la población en una locación disminuye por la cantidad de personas que emigran a otras localidades $\left(-\sum_{j=1}^K f_{ij} N_i\right)$ y aumenta por la cantidad de personas que inmigran de otras locaciones a la actual $\left(\sum_{j=1}^K f_{ji} N_j\right)$. Además puede notar que este sistema de ecuaciones diferenciales tiene K ecuaciones y $K(K-1)$ parámetros relacionados a la movilidad.

3.1.1 Modelo SIR con Movimiento Euleriano

Cuando combinamos ambos modelos debemos tener en cuenta que las clases compartimentales se deben fraccionar por localización para que tenga sentido la inclusión de una dinámica de movimiento. Es decir, como mismo en el modelo no epidemiológico los habitantes de una locación podían *mudarse* a otra, entonces esas personas pudieran bien ser susceptibles, y/o infestadas, y/o recuperadas. En este sentido habría que considerar tener *susceptibles*, *infestados*, y *recuperados*, por cada uno de los nodos. Además como en un principio queremos que este modelo no tan solo incorpore movilidad, sino más heterogeneidad debemos contar con los vectores $\beta_{1 \times K}$ y $\gamma_{1 \times K}$ que representan las tasas de contagio y recuperación por localidad. De esta forma el sistema quedaría:

$$\dot{S}_i = -\beta_i \frac{S_i I_i}{N_i} - \sum_{j=1}^K f_{ij} S_i + \sum_{j=1}^K f_{ji} S_j \quad (12)$$

$$\dot{I}_i = \beta_i \frac{S_i I_i}{N_i} - \gamma_i I_i - \sum_{j=1}^K f_{ij} I_i + \sum_{j=1}^K f_{ji} I_j \quad (13)$$

$$\dot{R}_i = \gamma_i I_i - \sum_{j=1}^K f_{ij} R_i + \sum_{j=1}^K f_{ji} R_j \quad (14)$$

$$\dot{N}_i = - \sum_{j=1}^K f_{ij} N_i + \sum_{j=1}^K f_{ji} N_j \quad (15)$$

Puede notar que al añadir la dinámica de movimiento obtenemos $4K$ ecuaciones, las K que trae el modelo de movimiento, más K ecuaciones por cada una de las clases según las localizaciones. Además se le suman a los $K(K-1)$ parámetros de movilidad, los K de β y los K de γ , para un total de $K(K+1)$ parámetros.

3.2 Movimiento Lagrangiano

Por otra parte el movimiento lagrangiano intenta proyectar los desplazamientos temporales de las personas, las cuales van hacia otra locación y luego regresan. Ejemplo de esto, son las personas que van a trabajar fuera de su municipio, o van a hacer compras, o salen con motivos de ocio. Esta dinámica intenta recordar las personas que siendo de una locación están de visita en otra. Aunque a nivel micro no se puede especificar cada persona dónde está, como en una microsimulación computacional, sí es posible llevar el conteo de la cantidad de personas que están desplazadas temporalmente, lo cual logra enriquecer la modelación.

Similar al movimiento euleriano debemos primero definir las tasas de desplazamiento de movimiento entre localidades, pero en este caso contamos con la salida y el retorno. Por lo tanto definamos las matrices $\phi_{K \times K}$ y $\tau_{K \times K}$ tales que:

- ϕ_{ij} denota la tasa de movimiento de los habitantes de la locación i hacia la locación j .
- τ_{ij} denota la tasa de movimiento de los habitantes de la locación i que están en la locación j y regresan hacia la i .

Dicho esto el sistema de ecuaciones nos queda da la siguiente forma:

$$\dot{N}_{ii} = - \sum_{j=1}^K \phi_{ij} N_{ii} + \sum_{j=1}^K \tau_{ij} N_{ij} \quad (16)$$

$$\dot{N}_{ij} = -\tau N_{ij} + \phi N_{ii} \quad (17)$$

3.2.1 Modelo SIR con Movimiento Lagrangiano

Cuando introducimos este sistema dentro de un modelo SIR clásico, volvemos a fraccionar cada una de las clases compartimentales pero esta vez cada clase se descompone en K^2 grupos, debido a que se necesita controlar para cada locación las cantidad de personas que hay en esta, y fuera de esta, en otra de las locaciones. Pero igual que en el modelo SIR con movimiento euleriano, solo necesitamos dos vectores β y γ para las tasas de contagio y recuperación. De esta forma obtenemos:

$$\dot{S}_{ii} = -\beta_i \frac{S_{ij} \sum_{k=1}^K I_{ki}}{\sum_{k=1}^K N_{ki}} - \sum_{k=1}^K \phi_{ik} S_{ii} + \sum_{k=1}^K \tau_{ik} S_{ik} \quad (18)$$

$$\dot{S}_{ij} = -\beta_j \frac{S_{ij} \sum_{k=1}^K I_{kj}}{\sum_{k=1}^K N_{kj}} - \tau_{ij} S_{ij} + \phi_{ij} S_{ii} \quad (19)$$

$$\dot{I}_{ii} = \beta_i \frac{S_{ij} \sum_{k=1}^K I_{ki}}{\sum_{k=1}^K N_{ki}} - \gamma_i I_{ii} - \sum_{k=1}^K \phi_{ik} I_{ii} + \sum_{k=1}^K \tau_{ik} I_{ik} \quad (20)$$

$$\dot{I}_{ij} = \beta_j \frac{S_{ij} \sum_{k=1}^K I_{kj}}{\sum_{k=1}^K N_{kj}} - \gamma_j I_{ij} - \tau_{ij} I_{ij} + \phi_{ij} I_{ii} \quad (21)$$

$$\dot{R}_{ii} = \gamma_i I_{ii} - \sum_{k=1}^K \phi_{ik} R_{ii} + \sum_{k=1}^K \tau_{ik} R_{ik} \quad (22)$$

$$\dot{R}_{ij} = \gamma_j I_{ij} - \tau_{ij} R_{ij} + \phi_{ij} R_{ii} \quad (23)$$

$$\dot{N}_{ii} = - \sum_{j=1}^K \phi_{ij} N_{ii} + \sum_{j=1}^K \tau_{ij} N_{ij} \quad (24)$$

$$\dot{N}_{ij} = -\tau N_{ij} + \phi N_{ii} \quad (25)$$

Puede notar que al añadir la dinámica de movimiento obtenemos $4K^2$ ecuaciones, las K^2 que trae el modelo de movimiento, más K^2 ecuaciones por cada una de las clases. Además se le suman a los $2K(K-1)$ parámetros de movilidad, los K de β y los K de γ , para un total de $2K + 2K(K+1) = 2K^2$ parámetros. Es evidente que mientras mayor sea el número de localizaciones más robusto se torna el sistema de ecuaciones pero aumenta la complejidad temporal para resolver el mismo.

4. Implementación de los Modelos

Con este trabajo proponemos una implementación sencilla de para la resolución de los distintos modelos presentados. Se utilizó como lenguaje de programación **Python** porque cuenta con módulos prácticos para resolver este tipo de problemas, tales como **numpy** y **scipy**. Más específicamente se utilizó la función **solve_ivp** del módulo

scipy.integrate para solucionar el sistema, empleando el algoritmo *Runge-Kutta* de grado 4. Los resultados se visualizan en gráficas generadas con el módulo **matplotlib.pyplot**.

Es importante señalar aspectos que se tuvieron en cuenta durante la implementación, y otros que se deben tener en cuenta para una futura versión mejorada. Recordemos que el algoritmo *Runge-Kutta* de grado 4, al que nos referiremos a partir de este momento sin especificar el grado, tiene una complejidad temporal de $4pT(y)$, siendo $T(y)$ la complejidad de evaluar una solución y en $y' = f(t, y)$, y p la cantidad de iteraciones del algoritmo. Note que y puede ser un vector de varias variables, al igual que y' .

La complejidad de evaluar los modelos compartimentales sin dinámica de movimiento suele ser bastante sencilla y de orden lineal en la cantidad de clases. De esta forma si un modelo tiene c clases, la complejidad será $T(y) = O(c)$ y nuestro algoritmo de solución quedará $4pO(c)$ haciendo abuso de notación. Esto sigue siendo de orden lineal en p por lo general ($O(cp) \approx O(p)$) ya que se saben la cantidad de clases c fijas, además de que estos valores son pequeños (de 3 a 6 clases).

Sin embargo estos modelos con metapoblaciones se complejizan mucho más, empezando por el hecho de que los vectores y y y' aumentan en dimensión. Para los modelos SIR con la dinámica añadida, los vectores solución, como vimos anteriormente, son de tamaño $4K$ y $4K^2$, para el movimiento euleriano y lagrangiano respectivamente. Por lo tanto a lo mínimo que podemos aspirar es a una complejidad de tal orden, es decir $O(K)$ y $O(K^2)$ respectivamente. En el primer caso la implementación resulta en $O(K^2)$ ya que las sumatorias de migración nos obligan a realizar K iteraciones por cada una de las variables. En este caso no resulta posible optimizar el orden de la complejidad ya que hay sumatorias que no tienen ni factores comunes, ni son comunes entre ecuaciones. En el segundo caso sí es posible que la implementación tenga orden $O(K^2)$, pero debemos hacer varias observaciones:

- Las sumatorias que implican a ϕ , tienen un factor común Q que no depende de la sumatoria, es decir, $\sum_{j=1}^K \phi_{ij}Q = Q \sum_{j=1}^K \phi_{ij}$, con lo cual es posible calcular estas una vez antes de iterar por la variables en cada paso de *Runge-Kutta*. Incluso como cada ϕ_{ij} es fijo, esta operación se puede realizar una única vez antes de hacer la evaluación y son sumatorias que nos ahorramos para cada iteración. Esto último no influye en el orden de la complejidad pero aún así sí influye positivamente en esta.
- Las sumatorias de las fracciones que implican a los I_{ki} y N_{ki} , se pueden calcular a parte antes de iterar por las variables sumando una complejidad temporal $O(K)$ a la ya existente $O(K^2)$ lo cual mantendría la complejidad en $O(K^2)$. Si esto no se tiene en cuenta, lo trivial sería calcular la sumatoria en ese momento y se incrementaría la complejidad a $O(K^3)$, lo cual resultaría más ineficiente.
- En el caso de las sumatorias $\sum_{k=1}^K \tau_{ik}Q$, donde Q puede ser S_{ik} , I_{ik} , R_{ik} y N_{ik} estas aparentemente incrementarían la complejidad a orden $O(K^3)$, pero si hacemos un análisis más detallado, solo se realizan cuando se evalúan las variables de índices (i, i) . Por lo tanto en total se están evaluando $4K^2$ veces, lo cual sigue siendo orden $O(K^2)$. Note que si estas sumatorias las sacamos y calculamos antes, similar al caso anterior, tendríamos $O(K^2) + O(K^2) = O(K^2)$, quedando invariante el orden. Sin embargo debido a lo explicado no es necesario hacer estas sumatorias aparte, ya que no afectan la complejidad.

Entonces teniendo en cuenta estas observaciones la solución del problema del valor inicial para un modelo SIR con movimiento lagrangiano se puede implementar con una complejidad temporal de orden $O(pK^2)$ siendo p la cantidad de iteraciones del algoritmo *Runge-Kutta*.

5. Resultados

A continuación se ilustrarán los resultados obtenidos de las diferentes implementaciones de cada uno de los modelos. De cada resultado obtenido se dará una breve explicación de los parámetros y valores iniciales usados, y en la medida que amerite, se irán haciendo comparaciones. Por una cuestión de comodidad comenzaremos mostrando los resultados de los modelos epidemiológicos clásicos, luego los modelos de movimiento, y finalmente los modelos epidemiológicos con metapoblaciones. Para hacer comparaciones justas, nos basaremos dentro de lo posible en los mismos valores iniciales, e iguales parámetros o parecidos.

Iniciamos comparando algunos modelos epidemiológicos clásicos. Haremos uso de las implementaciones realizadas para los modelos: SIR, SEIR y SIS. Utilizaremos los parámetros $\beta = 0.07$ y $\gamma = 0.2$. Como variables iniciales del modelo SIR se emplea $y_0 = (49955, 5, 0)$, del SEIR $y_0 = (49955, 0, 5, 0)$ y del SIS $y_0 = (49955, 5)$. Puede apreciar las diferencias entre los modelos y además sus similitudes. En primer lugar los modelos SIR, y SEIR alcanzan resultados en este caso muy parecidos, pero note que el pico de infestados es menor en el segundo y tarda más en ser alcanzado. Esto se debe al tiempo de incubación de la enfermedad, donde se posee el virus y no se transmite, posibilitando que durante este tiempo hayan infestados recuperándose, y de esta forma el aumento de casos positivos crezca con más lentitud. Además note que luego de alcanzar el pico de expuestos, se alcanza el pico de infecciosos, lo cual tiene sentido en un modelo SEIR. Por otra parte puede observar cómo en los tres modelos a partir de cierto valor t^* se mantiene constante la cantidad de susceptibles a la enfermedad. En los primeros 2 casos es debido a que la cantidad infestados comienza a tender a 0, es decir, se acaba la pandemia. Sin embargo en el segundo caso, lo que se logra es una estabilidad entre susceptibles e infestados, es decir, que se infestan tantas personas como las que se recuperan. Esta última gráfica se puede malinterpretar asumiendo que las personas infestadas en el punto

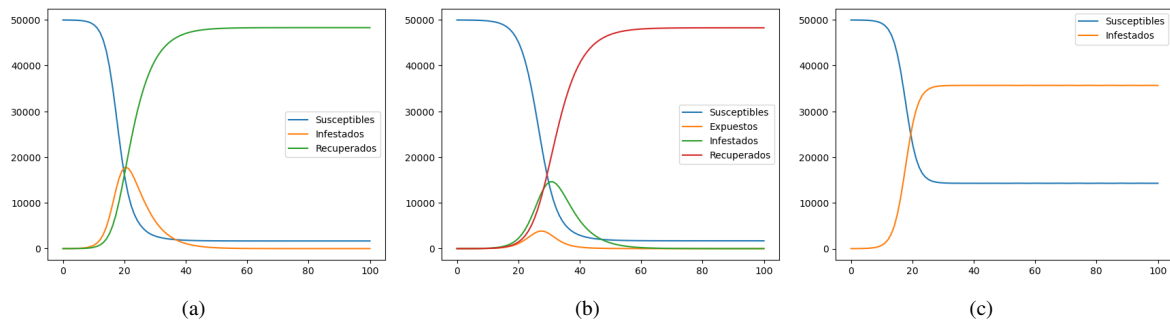


Figure 1. Modelo SIR (a), Modelo SEIR (b), y Modelo SIS (c).

t^* se mantienen en ese estado, al igual que las susceptibles, pero lo que realmente se debe entender es que siempre habrá una cantidad de personas infestadas y susceptibles constantes. Esto último tiene sentido debido a la forma en que se construye el modelo, donde en todo momento hay flujo entre las clases infectados y susceptibles.

Seguimos con los modelos de dinámicas de movimientos. La matriz de movilidad que estaremos usando en el caso del movimiento euleriano es:

$$F = \begin{pmatrix} 0 & 0.1 & 0.2 \\ 0.2 & 0 & 0.3 \\ 0.13 & 0.14 & 0 \end{pmatrix} \quad (26)$$

Para el movimiento lagrangiano emplearemos $\phi = F$, y 2 variantes de τ : $\tau = 0_{3 \times 3}$ y $\tau = \phi$. Además se utilizan como valores iniciales para el primero $y_0 = (1500, 2000, 1000)$, mientras que para el segundo la matriz diagonal, con esta igual al y_0 definido anteriormente.

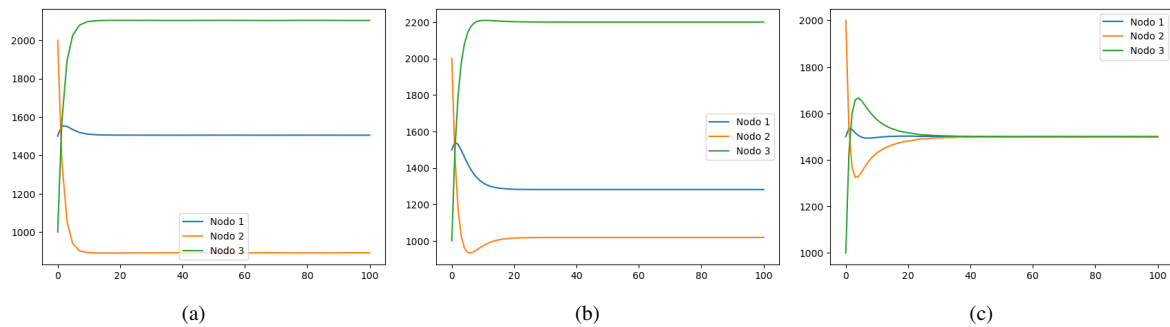


Figure 2. Movimiento Euleriano como matriz de movimiento F (a). Movimiento Lagrangiano con $\phi = F$ y $\tau = 0_{3 \times 3}$ (b), y con $\tau = \phi$ (c).

Es evidente las diferencias en los 3 ejemplos. Aunque estos 2 modelos pudieran tener semejanzas debido a que intentan modelar el movimiento, hay que tener presente que ambos modelos no son equivalentes [4]. Algo que se debe notar es la tendencia al equilibrio a partir de cierto punto. La interpretación de esto puede ser confundida con que a partir de cierta unidad de tiempo t^* , las personas se quedan en su localidad actual, o se detiene el desplazamiento: NO, este pensamiento está errado. Lo que se puede inferir es que a partir de t^* la cantidad de personas en cada locación se mantiene constante. Note la diferencia, y cómo realmente sigue ocurriendo la movilidad, lo que ha llegado a un punto de equilibrio donde cada región recibe un número de personas igual a la cantidad de personas que envía a otras regiones.

Finalmente se añaden metapoblaciones a los modelos epidemiológicos. Para el caso euleriano (caso 1) se toma como matriz de movimiento la matriz F ya definida. Se utilizaron los valores de los parámetros $\beta = (0.7, 0.5, 0.3)$ y $\gamma = (0.2, 0.2, 0.2)$. En el caso lagrangiano se mantuvieron estos parámetros pero se utilizaron dos juegos de matrices ϕ y τ : el primero tomando $\phi = F$ y $\tau = \phi$ (caso 2), y el segundo con $\phi = F$ y $\tau = 0_{3 \times 3}$ (caso 3). Los valores iniciales usados fueron en los 3 casos una población de 1500, 2000 y 1000 para los 3 nodos, con 15, 20 y 10 infestados respectivamente, así como 0 recuperados en cada nodo, y el resto de personas susceptibles. Para el caso lagrangiano además, no se considera en un inicio que haya personas de un nodo i en un nodo j con $i \neq j$. En la figura (3) se ve el comportamiento por nodos de las diferentes clases, para estos 3 casos, en función del tiempo.

Se puede apreciar comportamientos generales similares pero se notan las diferencias que hay entre los resultados obtenidos. Según se expresa la simulación cambia el comportamiento en cada modelo. Es curioso ver cómo cada nodo puede ser interpretado como un modelo SIR aparentemente independiente del resto. Note que este desglose por nodos nos brinda más información y permite tomar acciones particulares sobre los nodos. Aunque son llamativos estos gráficos, debemos analizarlos a un nivel macro, para tener una idea del comportamiento general

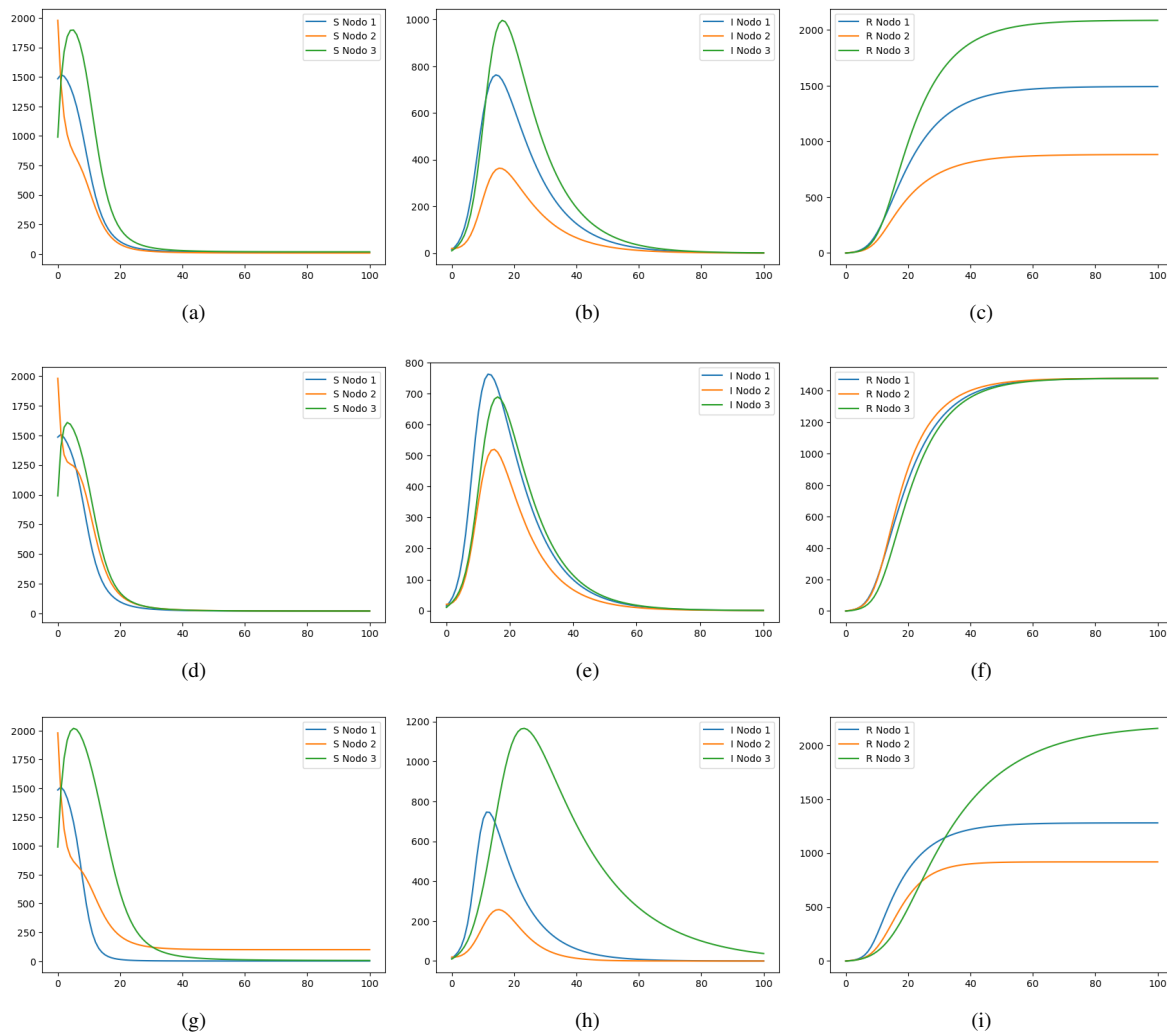


Figure 3. Desglose por nodos y por clases Susceptibles (S), Infestadas (I), y Recuperadas (R) de los modelos epidemiológicos con dinámica de movimiento en función del tiempo. Modelo SIR con movimiento Euleriano caso 1 (a-c). Modelo SIR con movimiento Lagrangiano caso 2 (d-f) y caso 3 (g-i).

de la enfermedad y poder realizar comparaciones válidas sobre un modelo SIR sin metapoblaciones. Para ello primero debemos solucionar un modelo SIR bastante cercano a estos, para ellos utilizamos $\gamma = 0.2$ tal y como fijamos para todos los nodos, y $\beta \approx 0.52222$ que sería el promedio de los β en la población. Además como valores iniciales tendríamos $1500 + 2000 + 1000 = 4500$ personas como población total, de las que $15 + 20 + 10 = 45$ están infestadas, no hay ninguna recuperada, y el resto son susceptibles. Los resultados se pueden visualizar en (4).

Note que la principal diferencia del modelo SIR clásico con el resto es la cantidad de personas susceptibles que quedan una vez eliminada la enfermedad. Esto se debe a que en este caso la población es considerada homogénea, mientras que en el resto, cada localización sufre la epidemia en paralelo y hay una crecimiento particular de casos en cada zona. Para los casos (c) y (d) de la figura (4), note que la disminución de casos infestados es marcadamente desacelerada en un caso, a pesar de que se trata del mismo modelo SIR con movimiento lagrangiano. De esta forma se evidencia el abanico de posibilidades que puede traer este modelo con la introducción de los parámetros τ , ya que recordemos que para estos 2 casos fueron seleccionados diferentes parámetros. En [4] se hace un especial estudio de la influencia de los parámetros escogidos en ambos tipos de movimientos, donde clasifican al lagrangiano de cierta forma caótico, debido a que es muy susceptible a cambios más drásticos bajo ligeras modificaciones en los parámetros de movilidad; principalmente cuando la tasa de contagio es mayor en una nodo que se suele visitar con estancias pequeñas de tiempo. Y por último note que a pesar de ser muy similares a nivel macro los modelos del caso 1 y el caso 2, no son iguales, y tienen comportamientos diferentes a nivel micro como se vio en la figura (3).

6. Discusión

Los modelos compartimentales son muy útiles para describir las epidemias, ya que tenemos constancia en cada instante de tiempo de la cantidad de personas que hay en cada clase estudiada, al igual que se describe de una manera fácil el flujo entre clases y con ciertos parámetros se puede controlar esta. En la práctica los parámetros deben ser sometidos a una interpretación real de la situación epidemiológica que exista, para poder hacer predicciones de la

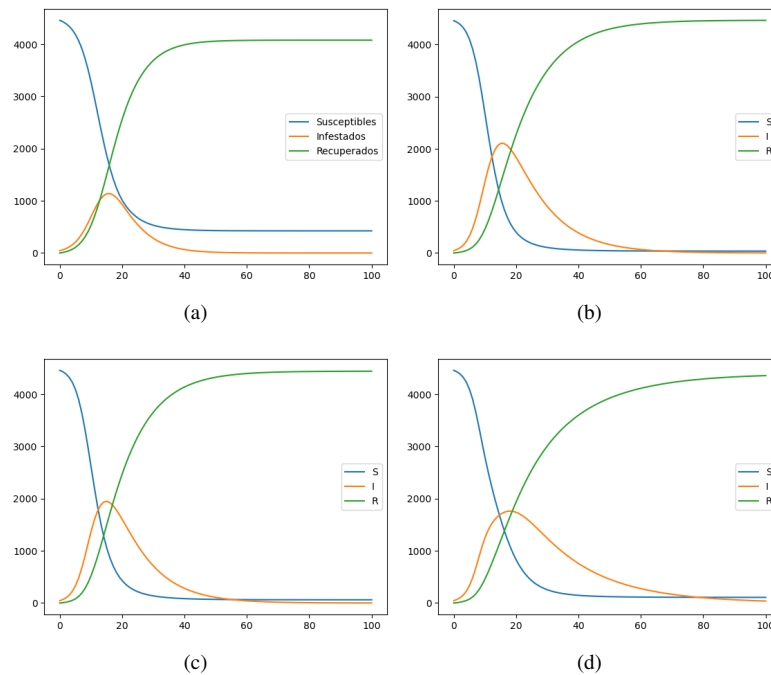


Figure 4. Vista general de los modelos epidemiológicos por clases Susceptibles (S), Infestadas (I), y Recuperadas (R) en función del tiempo. Modelo SIR clásico (a). Modelo SIR con movimiento euleriano (b). Modelo SIR con movimiento lagrangiano caso 2 (c) y caso 3 (d).

epidemia y tomar control de esta.

Las variantes sin metapoblaciones del modelo SIR, añaden, modifican o eliminan clases en la modelación, lo cual tiene un impacto generalmente positivo en la misma, reflejando mejor la realidad, lo que hay que tener presente es cuándo usar cuál. Esto último está relacionado directamente con la enfermedad en estudio y debido a sus características se debe usar la variante de modelo SIR correspondiente. No tiene sentido que una enfermedad sin inmunidad se modele con un SIR clásico, ya que no reflejará la realidad y haremos predicciones erradas. Por otra parte estos modelos funcionan mejor en poblaciones homogéneas debido a la naturaleza misma de su definición. Esto no es que sea desventajoso siempre, sino que tiene más sentido aplicarlo en escenarios donde la población tiene comportamientos, rutinas, etc., similares. Pero cuando hay diferencias notables dentro de la población estos pierden precisión. Una ventaja de estos modelos es que a nivel computacional la implementación es sencilla, y su solución suele tener un orden $O(cp)$ en cuanto a complejidad temporal, usando el algoritmo Runge-Kutta, donde c es la cantidad de clases y p son las iteraciones del algoritmo.

Por otra parte las variantes con metapoblaciones brindan más robustez al incluir dinámicas de movimientos, aún respetando las clases que se quieran modelar. En este caso cada locación se fracciona en estas clases y se puede particularizar en los comportamientos por área y clase. En este sentido aportan más información que los anteriores. Además un pilar importante es que permite modelar la heterogeneidad de la población lo cual suele ser mucho más común [6]. Por ejemplo, la provincia de Matanzas, en Cuba se distingue por tener municipios muy apartados geográficamente con una dinámica propia en cada uno; en este sentido es preferible utilizar un modelo de este tipo para simular el comportamiento de una epidemia. Otra ventaja de este modelo es que ilustra mucho mejor la toma de algunas medidas reales de movilidad. Por ejemplo, cerrar el tránsito de un municipio a otro, o de una provincia a otra, se traduce a alterar las matrices de movilidad F , ϕ y τ . De esta forma aislamos un nodo, o incluso varios nodos, lo cual permite independizar los comportamientos para cada región. En la práctica el cierre de localizaciones no es absoluto sino con cierto grado (leve, gradual, severo), y estos modelos matemáticos pueden modelar esto ya que los parámetros no son binarios, sino escalares en \mathbf{R} . En general si se tienen los datos de movilidad real, emplear estos modelos podría ser muy ventajoso.

Sin embargo, los modelos con metapoblaciones tienen varias desventajas. La primera es que requieren más ecuaciones y a su vez la evaluación en estas suele ser polinómica. Por ejemplo en la implementación realizada se logra que la complejidad temporal de evaluación sea de orden cuadrado $O(K^2)$, pero es muy fácil caer en una implementación $O(K^3)$. Evidentemente esto no solo presupone más tiempo sino más capacidad espacial también. Lo segundo es que requiere muchísimos parámetros, los cuales no se obtienen trivialmente, y se necesitan en el mejor de los casos, hacer estimaciones, ya que pueden ser datos a los que no se tenga acceso. Particularmente estamos hablando de las tasas de movilidad entre un conjunto de localizaciones. Lo tercero sería la representación de la población en nodos, lo cual trae cierta ventaja ya que se refleja la realidad, pero hay aspectos que debemos tener en cuenta. Cuando se construye el grafo de las localizaciones es común trabajar con áreas de salud, municipios, provincias, etc. Se debe tener presente qué se quiere modelar y con qué datos se va a modelar. Para particularizar

y detectar el problema, nos situaremos en la provincia de La Habana en Cuba. Esta cuenta con municipios muy cercanos que comúnmente personas pueden ir de uno a otro caminando (como Centro Habana - Cerro o Centro Habana - Habana Vieja), y visitarlo con frecuencia; entonces tendrá sentido separar estos 2 municipios? las tasas de contagio y recuperación son muy diferentes? Otro ejemplo es cuando se desea controlar dentro de un municipio una epidemia, pudiéramos hacer una partición por Consejos Populares que es lo que suele tener una mejor organización, pero no sería mejor hacer una división por Áreas de Salud? Estas son incógnitas interesantes. Y muy relacionado con esto está la parte de dónde se extraen los datos? Las tasas de movilidad están relacionadas por los nodos que seleccionamos, pero si en la práctica los datos de movilidad que tenemos son proporcionados por ejemplo, por la cantidad de recepciones de señales telefónicas de un receptor, entonces lo que realmente tenemos es la movilidad entre zonas alrededor de cada receptor y esto no tiene que coincidir necesariamente con el área de un municipio, ya que pueden haber receptores que colinden entre dos o más municipios. Por lo tanto, se debe hacer un buen trabajo de interpretación de la realidad antes de construir un modelo de este estilo.

Por último debemos tener en cuenta que también los modelos con metapoblaciones generalmente requieren estimar muchos parámetros y se deben utilizar técnicas especializadas para esto. Por lo tanto esto aumenta la complejidad de la modelación sumado a la complejidad de solucionar el modelo, lo que a simple vista parece un menor problema, pero sin una buena optimización puede ser algo desastroso o imposible de computar

7. Conclusiones

Con este trabajo se logró hacer un estudio sobre los modelos matemáticos epidemiológicos más usados en la práctica como es el caso de los modelos SIR, SEIR, y SIS, además de los modelos de dinámicas de movimiento euleriano y lagrangiano, y la introducción de estos últimos en modelos SIR. Se explicó el funcionamiento general de los primeros modelos para entender los últimos mientras que se profundizó en los modelos de movimientos atendiendo a sus peculiaridades e interpretaciones. Se ofreció una primera implementación eficiente en cuanto a complejidad temporal en el algoritmo de evaluación del sistema de ecuaciones. Se ejecutaron algunos ejemplos y se compararon marcando las principales diferencias, y revisando las similitudes para concluir que como modelos son diferentes. Por último se discutió las ventajas y desventajas de cada familia de modelos, constatando que no se debe colocar un modelo por encima de otro apartado del problema que se tiene.

Todos estos modelos tienen sus aspectos buenos y malos, y atienden a diferentes casos de uso, por lo que no es justo, ni práctico enunciar la superioridad de alguno. Sin embargo, una vez tenemos un problema y analizamos la situación que gira entorno a este, sí podríamos seleccionar un modelo que supere al resto, para realizar una simulación cercana a la realidad. Para casos donde la población tiene comportamientos bastantes comunes es preferible hacer uso de los modelos clásicos. Cuando la población se encuentra geográficamente apartada es preferible usar modelos con metapoblaciones, y en caso de que aún así la población se comporte de manera similar, se puede mantener la homogeneidad de esta fijando iguales valores β_i y γ_i para todos los nodos. Para casos como el anterior pero donde la heterogeneidad es fundamental, sí deberíamos establecer β_i particulares para cada nodo. Por último y no menos importante, a la hora de decidir tipos de movimiento, como el movimiento lagrangiano es más dinámico, debe usarse cuando entre nodos habrá mucho tráfico de este tipo, dígame por ejemplo, entre municipios de una provincia; y en los casos donde el tráfico por períodos largos sea superior a viajes rutinarios, y a la vez indefinido, es preferible emplear el movimiento euleriano, como por ejemplo, entre provincias si sabemos que la cantidad de mudanzas a nivel nacional es grande, o por ejemplo entre país y país ya que es muy común que las personas migren en este caso por tiempo indefinido o plazos más largos de tiempo.

La principal conclusión de este proyecto se reitera e insiste que está en la necesidad de interpretar bien el problema que se tiene, los datos que sabemos de este, y seleccionar el modelo de estos que más sentido analítico y práctico tenga. Adicionalmente se propone seguir trabajando en la implementación eficiente de estos sistemas, revisando aspectos como la reducción de ecuaciones, técnicas de estimación de parámetros y mecanismos computacionales para agilizar los cálculos, como pueden ser cambiar de herramientas matemáticas, utilizar otras técnicas de compilación, o incluso emplear otro lenguaje de programación.

References

- [1] Stefan Ma and Yingcun Xia. *Mathematical understanding of infectious disease dynamics*, volume 16. World Scientific, 2009.
- [2] Richard Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the ESA*, 15(3):237–240, 1969.
- [3] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.
- [4] Daniel T Citron, Carlos A Guerra, Andrew J Dolgert, Sean L Wu, John M Henry, Héctor M Sánchez C, and David L Smith. Comparing metapopulation dynamics of infectious diseases under different models of human movement. *Proceedings of the National Academy of Sciences*, 118(18):e2007488118, 2021.

- [5] Howard Howie Weiss. The sir model and the foundations of public health. *Materials matematics*, pages 0001–17, 2013.
- [6] Daniela Calvetti, Alexander P Hoover, Johnie Rose, and Erkki Somersalo. Metapopulation network models for understanding, predicting, and managing the coronavirus disease covid-19. *Frontiers in Physics*, 8:261, 2020.
- [7] Samuel Mwalili, Mark Kimathi, Viona Ojiambo, Duncan Gathungu, and Rachel Mbogo. Seir model for covid-19 dynamics incorporating the environment and social distancing. *BMC Research Notes*, 13(1):1–5, 2020.