

UNIVERSIDAD NACIONAL DEL LITORAL



# **Modelado de epidemias utilizando sistemas auto-organizados**

Leonardo López

**FICH**

FACULTAD DE INGENIERIA  
Y CIENCIAS HIDRICAS

**INTEC**

INSTITUTO DE DESARROLLO TECNOLÓGICO  
PARA LA INDUSTRIA QUIMICA

Tesis de Doctorado **2017**





UNIVERSIDAD NACIONAL DEL LITORAL  
Facultad de Ingeniería y Ciencias Hídricas  
Instituto de Desarrollo Tecnológico para la Industria Química

**MODELADO DE EPIDEMIAS  
UTILIZANDO  
SISTEMAS AUTO-ORGANIZADOS**

**Leonardo López**

Tesis remitida al Comité Académico del Doctorado  
como parte de los requisitos para la obtención  
del grado de  
DOCTOR EN INGENIERIA  
Mención Inteligencia Computacional, Señales y Sistemas  
de la  
UNIVERSIDAD NACIONAL DEL LITORAL

**2017**

Comisión de Posgrado, Facultad de Ingeniería y Ciencias Hídricas, Ciudad Universitaria, Paraje "El Pozo",  
S3000, Santa Fe, Argentina.





UNIVERSIDAD NACIONAL DEL LITORAL  
Facultad de Ingeniería y Ciencias Hídricas  
Instituto de Desarrollo Tecnológico para la Industria Química

**MODELADO DE EPIDEMIAS  
UTILIZANDO  
SISTEMAS AUTO-ORGANIZADOS**

**Leonardo López**

**Lugar de Trabajo:**

sinc(i)-UNL-CONICET

Instituto de Investigación de Señales, Sistemas e Inteligencia Computacional

Facultad de Ingeniería y Ciencias Hídricas

Universidad Nacional del Litoral

**Director:**

Dr. Leonardo Giovanini

sinc(i)-UNL-CONICET

**Co-director:**

Dr. Pablo Beldoménico

ICIVET-UNL-CONICET

**Jurado Evaluador:**

Dr. Marcelo Kuperman

Instituto Balseiro-UNCUYO

Dra. Nara Guisoni

IFLySiB - UNLP

Dr. Gustavo Sibona

IFEG-CONICET-UNC

**2017**






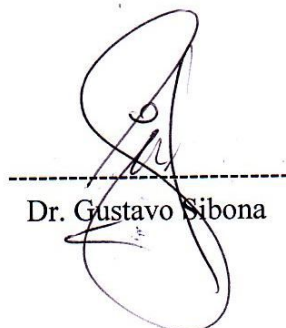
**UNIVERSIDAD NACIONAL DEL LITORAL**  
**Facultad de Ingeniería y Ciencias Hídricas**


Santa Fe, 20 de marzo de 2017.

Como miembros del Jurado Evaluador de la Tesis de Doctorado en Ingeniería titulada **“Modelado de epidemias utilizando sistemas auto-organizados”**, desarrollada por el Lic. Leonardo Rafael LOPEZ, en el marco de la Mención “Inteligencia Computacional, Señales y Sistemas”, certificamos que hemos evaluado la Tesis y recomendamos que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

La aprobación final de esta disertación estará condicionada a la presentación de dos copias encuadernadas de la versión final de la Tesis ante el Comité Académico del Doctorado en Ingeniería.

  
Dra. Nara Guisoni

  
Dr. Gustavo Sibona

  
Dr. JOSÉ LUIS MACOR  
SECRETARIO DE POSGRADO  
Facultad de Ingeniería y Cs. Hídricas  
Dr. Marcelo Kuperman (\*)

Santa Fe, 20 de marzo de 2017.

Certifico haber leído la Tesis, preparada bajo mi dirección en el marco de la Mención “Inteligencia Computacional, Señales y Sistemas” y recomiendo que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

.....  
Dr. Pablo Beldoménico  
Codirector de Tesis

.....  
Dr. Leonardo Giovanini  
Director de Tesis

(\*) participó por videoconferencia

  
Dr. JOSÉ LUIS MACOR  
SECRETARIO DE POSGRADO  
Facultad de Ingeniería y Cs. Hídricas

Universidad Nacional del Litoral  
Facultad de Ingeniería y  
Ciencias Hídricas

Secretaría de Posgrado

Ciudad Universitaria  
C.C. 217  
Ruta Nacional N° 168 - Km. 472,4  
(3000) Santa Fe  
Tel: (54) (0342) 4575 229  
Fax: (54) (0342) 4575 224  
E-mail: posgrado@fich.unl.edu.ar





## DECLARACION DEL AUTOR

Esta Tesis ha sido remitida como parte de los requisitos para la obtención del grado académico Doctorado en Ingeniería Mención Señales, Sistemas e Inteligencia Computacional ante la Universidad Nacional del Litoral y ha sido depositada en la Biblioteca de la Facultad de Ingeniería y Ciencias Hídricas para que esté a disposición de sus lectores bajo las condiciones estipuladas por el reglamento de la mencionada Biblioteca.

Citaciones breves de esta Tesis son permitidas sin la necesidad de un permiso especial, en la suposición de que la fuente sea correctamente citada. Solicitudes de permiso para la citación extendida o para la reproducción parcial o total de ese manuscrito serán concebidos por el portador legal del derecho de propiedad intelectual de la obra.



## **Dedicatoria**

A mi familia.



## Agradecimientos

Al escribir estas líneas siento que por mas memoria que haga voy a olvidar a alguien, a esas personas les pido mil disculpas. Quiero comenzar agradeciendo a mis compañeros de trabajo, todas las personas que integran este maravilloso elenco denominado *sinc(i)*. El tiempo apremia así que espero que sepan disculpar que no los nombre uno a uno. Me gustaría extender una mención especial a mis compañeros de oficina, a estos si los voy a nombrar porque son menos. Juan, Sebastián, José y Leandro. Gracias por hacer ameno cada día de trabajo. Me gustaría hacer mención a quienes iniciaron esta cruzada conmigo. Milton, Iván, Guido, David y Guille. Un especial agradecimiento a quienes conforman la cabeza del instituto, por llevar adelante la gestión y hacer posible que funcione todo esto. Quiero agradecer especialmente a mi director, Leonardo Giovanini y a mi co-director Pablo Beldoménico, no solo me han enseñado mucho a lo largo de estos años, si no que demostraron ser dos grandes personas incluso fuera del ámbito académico. Aprovecho para extender este agradecimiento a Dario Manzoli y la gente del ICIVET, por brindarnos valiosa información para nuestros modelos. Quiero agradecer a mi familia, a mis padres, Noemí y José, por acompañarme a lo largo de mi vida. A mis hermanos Benito y Mario, mis primeros amigos. A mis amigos Fabricio, Martín, Julián, Mauro, Carolina, Emilia, Inti, Cristian, Marcia, Gonzalo y perdón si me olvidé de alguno. A mis padres políticos Edgardo, Mabel e Irene, me acompañaron e hicieron posible esto. El agradecimiento mas importante sin dudas es para vos, Florencia, mi compañera, la madre de mi hija, la que me aguantó absolutamente todo. A vos te quiero agradecer especialmente, no hay palabras que puedan expresar lo que siento. Finalmente, como dicen los anglo parlantes, last but not least, Lucía, mi hija, un norte, a riesgo de sonar trillado y meloso, la luz de mis días.

Gracias a todos.



# Contenido

<b>Agradecimientos</b>	<b>XIII</b>
<b>1. Introducción</b>	<b>1</b>
1.1. Introducción . . . . .	1
1.2. Estado del arte . . . . .	7
1.2.1. Modelos clásicos . . . . .	7
1.2.2. Modelos basados en individuos . . . . .	23
1.3. Objetivos . . . . .	28
1.4. Organización . . . . .	28
1.5. Conclusiones . . . . .	29
<b>2. Modelo</b>	<b>31</b>
2.1. Introducción . . . . .	31
2.2. Estado del arte . . . . .	32
2.3. Modelo propuesto . . . . .	35
2.3.1. Comportamiento individual . . . . .	37
2.3.2. Comportamiento de individuos en epidemias . . . . .	40
2.4. Comportamiento social . . . . .	46
2.5. Estado epidémico . . . . .	46
2.6. Conclusión . . . . .	47
<b>3. Modelos basados en redes de contactos</b>	<b>48</b>
3.1. Introducción . . . . .	48
3.2. Maco teórico . . . . .	49
3.2.1. Redes de contacto . . . . .	49
3.2.2. Modelo basado en autómatas . . . . .	51
3.2.3. Autómatas celulares como redes de contactos . . . . .	55
3.2.4. Estado del arte . . . . .	58
3.3. Modelo . . . . .	62
3.4. Estimación de parámetros . . . . .	69
3.5. Validación del modelo . . . . .	72
3.5.1. Validación visual . . . . .	72
3.5.2. Validación numérica . . . . .	75
3.6. Otros Resultados . . . . .	78

3.7. Conclusiones . . . . .	82
<b>4. Modelos espacialmente explícitos</b>	<b>83</b>
4.1. Introducción . . . . .	83
4.2. Marco teórico . . . . .	84
4.2.1. Agentes . . . . .	84
4.2.2. Ambientes . . . . .	85
4.2.3. Sistemas multi-agentes . . . . .	86
4.3. Estado del arte . . . . .	88
4.4. Modelo . . . . .	91
4.5. Estimación de parámetros . . . . .	99
4.6. Validación . . . . .	100
4.6.1. Validación visual . . . . .	100
4.6.2. Validación numérica . . . . .	103
4.7. Otros Resultados . . . . .	106
4.8. Conclusiones . . . . .	110
<b>5. Múltiples cepas y estrategias de control</b>	<b>112</b>
5.1. Introducción . . . . .	112
5.2. Estado del arte . . . . .	113
5.3. Implementación . . . . .	116
5.3.1. Múltiples cepas . . . . .	117
5.3.2. Vacunación . . . . .	120
5.3.3. Cuarentena . . . . .	123
5.4. Conclusiones . . . . .	125
<b>6. Conclusiones</b>	<b>126</b>
6.1. Conclusiones . . . . .	126
6.2. Publicaciones resultantes . . . . .	129
6.3. Trabajo Futuro . . . . .	130
<b>A. Rutinas implementadas del modelo de redes de contactos</b>	<b>133</b>
<b>B. Rutinas del modelo espacialmente explícito</b>	<b>136</b>
<b>C. Estimación de la probabilidad de error de los modelos propuestos</b>	<b>138</b>
<b>Bibliografía</b>	<b>139</b>





# Lista de Figuras

1-1.	Flujo de población en el modelo SIR estático . . . . .	8
1-2.	Dinámica temporal de modelo SIR . . . . .	9
1-3.	Dinámica temporal de modelo SIR . . . . .	11
1-4.	Flujo de población en el modelo SEIR dinámico . . . . .	12
1-5.	Diagrama de un modelo poblacional basado en ondas . . . . .	13
1-6.	Diagrama de modelo sub-poblacional . . . . .	15
1-7.	Diagrama de modelo sub-poblacional móvil . . . . .	16
1-8.	Diagrama de red de contactos de 10 individuos . . . . .	18
1-9.	Modelo basado en individuos espacialmente implícito . . . . .	24
1-10.	Modelo basado en individuos móviles . . . . .	25
2-1.	Diagrama esquemático del marco de trabajo propuesto . . . . .	36
2-2.	Emociones y sus relaciones a diferentes niveles. . . . .	38
2-3.	Diagrama de un FCM de 6 conceptos . . . . .	39
2-4.	Mapa cognitivo difuso implementado para emular el comportamiento de los individuos en la epidemia . . . . .	44
2-5.	(a) Histograma de los valores $C_{10}$ obtenidos (b) Gráfico de cajas . . . . .	45
3-1.	Diagrama de estados de (a) una máquina de Moore (b) una máquina de Mealy . . . . .	52
3-2.	Topologías . . . . .	53
3-3.	Tipos de vecindades . . . . .	54
3-4.	Visualización de una grilla de autómatas como una red de contactos . . . . .	56
3-5.	Proceso de reescritura de la red de contactos . . . . .	57
3-6.	Diagrama del modelo basado en redes de contacto . . . . .	63
3-7.	Grafo de transición de estados del autómata que modela el estado epidémico. . . . .	68
3-8.	Diferentes representaciones de la red de contactos como (a) grilla, (b) red booleana y (c) red de propagación . . . . .	68
3-9.	Evolución de la epidemia para una población homogénea, promedio de 500 realizaciones . . . . .	71
3-10.	Evolución de la epidemia para vecindad de Moore, promedio de 500 realizaciones. . . . .	71
3-11.	Respuestas promedio del modelo con y sin conducta social . . . . .	73
3-12.	(a) Gráfico de residuos (b) Distribución del error . . . . .	74

3-13. Evolución de la epidemia para una distribución normal y bimodal de la tasa de infección. Resultados de una realización por caso. . . . .	78
3-14. (a)Residuos para el modelo con distribución normal y bimodal. (b) Distribución del error para el modelo normal y bimodal . . . . .	80
3-15. Dinámicas temporales para diferentes distribuciones espaciales iniciales de individuos infectados asumiendo movimiento aleatorio. . . . .	81
4-1. Vecindad de un agente . . . . .	88
4-2. (a) $\alpha$ -lattice (b)cuasi $\alpha$ -lattice . . . . .	94
4-3. Diagrama del modelo basado en individuos con espacio explícito . . . . .	98
4-4. Ajuste del modelo, respuesta promedio de 500 corridas . . . . .	99
4-5. Ajuste del modelo. Respuesta promedio sobre 500 corridas. . . . .	100
4-6. (a) Residuos del modelo con espacio explícito. (b) Residuos del modelo clásico. . . . .	102
4-7. Evolución de la epidemia para diferentes distribuciones espaciales de los individuos infecciosos (a) sin conducta social, (b) con conducta social. . . . .	107
4-8. Dinámica temporal para distribución espacial localizada para (a) $t = 0$ (b) $t = 10$ (c) $t = 30$ (d) $t = 50$ . . . . .	108
4-9. Dinámica temporal para distribución espacial uniforme para (a) $t = 0$ (b) $t = 10$ (c) $t = 30$ (d) $t = 50$ . . . . .	109
5-1. Diagrama de flujo del modelo de Samsuzzoha, Singh, y Lucy (2012). . . . .	114
5-2. Diagrama de flujo del modelo con estado de cuarentena $SIQS$ . . . . .	114
5-3. Diagrama de flujo del modelo con estado de cuarentena $SIQR$ que contempla población recuperada . . . . .	115
5-4. Diagrama de flujo del modelo propuesto en (Nuno, Feng, Martcheva, & Castillo-Chavez, 2005), . . . . .	116
5-5. Diagrama de flujo del modelo $SIR$ que contempla múltiples cepas e inmunidad cruzada (Casagrandi, Bolzoni, Levin, & Andreasen, 2006) . . . . .	116
5-6. Diagrama esquemático del marco de trabajo propuesto . . . . .	117
5-7. Grafo de estados del modelo de infección con múltiples cepas . . . . .	119
5-8. Ajuste del modelo para múltiples cepas . . . . .	120
5-9. (a) Gráfico de residuos (b) Distribución del error . . . . .	121
5-10. Grafo de estados del modelo con vacunación . . . . .	122
5-11. Evolución de la epidemia para diferentes porcentajes de individuos vacunados. Respuesta promedio para 500 realizaciones por caso. . . . .	123
5-12. Evolución de la epidemia para diferentes períodos de duración de cuarentena. Respuesta promedio para 500 . . . . .	124
C-1. Probabilidad de que el error del modelo implementado sea menor al error del modelo clásico . . . . .	139



# Lista de Tablas

2-1. Cuadro de conceptos del FCM . . . . .	41
2-2. Pesos entre los diferentes conceptos del FCM. . . . .	43
3-1. Matriz de transición . . . . .	64
3-2. Matriz de transición por contacto . . . . .	65
3-3. Vector de estado inicial . . . . .	66
3-4. Parámetros minimizados del modelo clásico . . . . .	70
3-5. Estadísticos del error . . . . .	73
3-6. Significancia de los errores para el modelo sin comportamiento social . . . .	77
3-7. Significancia de los errores para el modelo con comportamiento social . . .	77
3-8. Estadísticos del error . . . . .	79
4-1. Parámetros minimizados del modelo clásico . . . . .	100
4-2. Estadísticos del error . . . . .	101
4-3. Significancia de los errores para el modelo sin comportamiento social . . . .	105
4-4. Significancia de los errores para el modelo con comportamiento social . . .	105
5-1. Estadísticos del error . . . . .	119
C-1. Probabilidad de la significancia estadística del error . . . . .	138



## Resumen

Hasta el momento, gran parte del desarrollo de modelos epidemiológicos se ha centrado principalmente en el uso de modelos compartimentales. Estos modelos se basan en una representación del sistema bajo estudio en términos de un conjunto de ecuaciones diferenciales ordinarias acopladas, las cuales representan las poblaciones involucradas en el sistema. Estos tienen limitaciones como las hipótesis de homogeneidad de contactos, comportamiento y distribución, entre otros. En los IBM, el comportamiento de los individuos está definido para cada una de ellos. Esto permite una gran flexibilidad, haciéndolos especialmente útiles en problemas donde la red de contactos es muy importante, ya que los individuos son considerados como entidades simples que interactúan de forma dinámica en espacios multidimensionales. Son sistemas dinámicos formados por un número finito de elementos que interactúan en un medio discreto. La dinámica de los individuos es descrita por funciones dinámicas y el estado de cada individuo es determinado por un conjunto de reglas locales. El estudio de la dinámica de los fenómenos de salud y enfermedad en sistemas naturales permite un mejor entendimiento de estos procesos biológicos, ya que no están influenciados por la medicina preventiva, la instauración de tratamientos y las respuestas a brotes. Sólo mediante el estudio de los determinantes de la distribución y dinámica de enfermedades se logrará contar con el conocimiento necesario para mitigar y controlar el impacto que éstas pueden tener sobre la salud. En esta tesis se presenta un marco de trabajo que consiste en un modelo modular construido a partir de unidades básicas que definen su dinámica a partir del comportamiento. El comportamiento de estas unidades está constituido por tres bloques principales: el comportamiento individual, el comportamiento social y el comportamiento epidemiológico. Por un lado el comportamiento individual, el cual puede ser visto como el resultado de aplicar un conjunto de reglas y sistemas basados en lógica difusa. En segundo lugar, el comportamiento social se refiere a las acciones que toman los individuos como resultado de las interacciones con otros individuos, estas interacciones pueden ser vistas desde el punto de vista de las redes de contacto que se forman en las poblaciones naturales, donde el espacio físico no es relevante o pueden ser vistas desde el punto de vista del espacio físico en donde los individuos se desenvuelven, siendo el entorno una variable epidemiológica más a tener en cuenta. Finalmente el comportamiento epidemiológico que es resultado del estado de salud del individuo es el tercer factor que determina el comportamiento global. El marco propuesto resulta entonces muy atractivo para modelar diferentes tipos de epidemias en donde el comportamiento de los individuos, sus relaciones y su entorno juegan un rol crucial. La construcción del comportamiento de los agentes a partir de tres bloques otorga al modelo una amplia gama de aplicaciones, ya que se pueden modelar diferentes situaciones simplemente realizando cambios en alguno de sus bloques. El modelo fue validado utilizando datos epidemiológicos de gripe y además se proponen diferentes escenarios de control que permiten evaluar la capacidad adaptativa del enfoque.

## Abstract

To date, much of the development of epidemiological models has focused mainly on the use of compartmental models. These models are based on a representation of the system under study in terms of a set of coupled ordinary differential equations, which represent the populations involved in the system. These have limitations as the hypotheses of homogeneity of contacts, behavior and distribution, among others. In IBM models, the behavior of individuals is defined for each of them. This allows for great flexibility, making them especially useful in problems where the network of contacts is very important, since individuals are considered as simple entities that interact dynamically in multidimensional spaces. IBM models has been used as an alternative method to simulate ecological systems. They are dynamic systems formed by a finite number of elements that interact in a discrete medium. The dynamics of individuals are described by dynamic functions and the state of each individual is determined by a set of local rules. The study of the dynamics of health and disease phenomena in natural systems allows a better understanding of these biological processes, since they are not influenced by preventive medicine, the establishment of treatments and the responses to outbreaks. Only through the study of the determinants of the distribution and dynamics of diseases will the knowledge necessary to mitigate and control the impact they have on human and animal health be achieved. In this thesis is presented a framework that consists of a modular model constructed from basic units that define their dynamics from the behavior. The behavior of these units is constituted by three main blocks: individual behavior, social behavior and epidemiological behavior. On the one hand the individual behavior, which can be seen as the result of applying a set of rules and systems based on fuzzy logic. Secondly, social behavior refers to the actions that individuals take as a result of interactions with other individuals, these interactions can be seen from the point of view of the contact networks that are formed in natural populations, where Physical space is not relevant or can be seen from the point of view of the physical space where individuals are developed, the environment being an epidemiological variable to be taken into account. Finally, the epidemiological behavior that results from the health status of the individual is the third factor that determines the overall behavior. The proposed framework is then very attractive to model different types of epidemics where the behavior of individuals, their relationships and their environment play a crucial role. The construction of the behavior of the agents from three blocks gives the model a wide range of applications, since different situations can be modeled simply by making changes in some of its blocks. The model was validated using influenza epidemiological data and in addition different control scenarios are proposed that allow to evaluate the adaptive capacity of the approach.



# 1. Introducción

## 1.1. Introducción

La epidemiología es una disciplina científica que estudia la distribución, frecuencia, determinantes, relaciones, predicciones y control de los factores relacionados con la salud y enfermedad en poblaciones humanas y animales. Ocupa un lugar especial en la intersección entre las ciencias biomédicas, las ciencias sociales y las matemáticas aplicadas, empleando métodos y principios de estas disciplinas para el estudio de la salud y la enfermedad en poblaciones. En este capítulo se introducen los principios básicos de la epidemiología, su paradigma central y la relación con la medicina clínica. Se aborda también la problemática en torno a las variables epidemiológicas y a la definición de enfermedad. Se introduce el concepto epidemiológico de población y la relación entre individuo y población a nivel de patrones de enfermedad, heterogeneidad y variables epidemiológicas. Además, se trata sobre la naturaleza inherentemente dinámica de las poblaciones humanas y sus variaciones a distintos niveles. Por otro lado, se realiza un análisis del estado del arte sobre el modelado epidemiológico. Se introducen los principales conceptos matemáticos que definen a los modelos clásicos y las principales ventajas y desventajas de este enfoque. A continuación se realiza un repaso sobre los aspectos más relevantes que hacen al análisis de redes de contactos y su aplicación directa con la problemática estudiada reinterpretando los modelos clásicos desde esta perspectiva. También se hace una revisión minuciosa de los modelos basados en individuos de forma de preparar al lector para una comprensión más adecuada del trabajo realizado. Finalmente se cierra este capítulo con una breve exposición de los objetivos planteados de esta tesis y se culmina con la conclusión del capítulo en donde se intenta realizar una recapitulación de los principales conceptos desarrollados.

### ¿Qué es la epidemiología?

Los individuos son únicos, sus características y comportamientos aportan a las causas de sus enfermedades. No obstante, es difícil hacer predicciones a nivel individual. A nivel poblacional, en cambio, se presentan patrones distintivos causados por las variaciones en la exposición de los individuos a las causas de enfermedad. Si poblaciones similares se exponen a las mismas causas de enfermedad, exhibirán un patrón similar. Pero los patrones dependen no sólo de las características de los individuos, sino de los tipos de interacciones con otros individuos y con el medio ambiente, y de la estructura social en

la que se encuentran inmersos. Si la sociedad o el ambiente cambian, también cambiarán los riesgos individuales y los patrones asociados (Szklo & Nieto, 2014). En este contexto, la epidemiología procura:

- Entender las causas de enfermedad en una población;
- Prevenir y controlar enfermedades en poblaciones;
- Proveer guías de salud, políticas de salud y planeamientos; y
- Asistir en la dirección y cuidado de la salud de los individuos saludables y enfermos.

Entre las hipótesis sobre las que se sustenta la epidemiología (Bhopal, 2002; Bonita, Beaglehole, & Kjellström, 2006), podemos destacar:

- La enfermedad en la población es más que la suma de las enfermedades en los individuos;
- Las poblaciones difieren en su experiencia ante la enfermedad;
- Dentro de una población, difieren entre subgrupos;
- Las variaciones en la enfermedad pueden describirse y pueden explorarse sus causas, evaluando si las variables se asocian con patrones de enfermedad; y
- El conocimiento de las enfermedades en las poblaciones puede aplicarse a individuos y viceversa.

La epidemiología es la ciencia y práctica que describe y explica los patrones de enfermedad en las poblaciones, y usa este conocimiento para prevenir y controlar las enfermedades y mejorar la salud. Al ocuparse más de poblaciones que de individuos, entran en juego la sociedad, comportamientos, actitudes, interacciones, leyes, tradiciones, cultura, constituyéndose en una ciencia a la vez biológica, matemática y social (Berkman, Kawachi, & Glymour, 2014). El paradigma central de la epidemiología supone que los patrones de enfermedad en las poblaciones pueden analizarse sistemáticamente para proporcionar un entendimiento de las causas y controlar la enfermedad. La estrategia clave es, entonces, analizar las diferencias y similitudes en los patrones de enfermedad de las poblaciones. La medición válida de la frecuencia de enfermedad y los factores que influyen es crucial para los fines epidemiológicos. Sin embargo, la medición es un medio, no un fin. Excelencia en las medidas no da por sí misma excelencia en epidemiología; la calidad de la epidemiología debe juzgarse por la contribución a lograr los fines, lo que también ocurre con el diseño de estudios epidemiológicos (Bhopal, 2002).

Las ideas comprobadas en el estudio sobre las enfermedades están usándose cada vez más para estudiar la salud y su cuidado (Rothman, 2012). La epidemiología también ha

sido útil en el laboratorio para entender procesos biológicos y definir estándares y rangos de valores biológicos y sociales normales. Los valores normales se derivan de la distribución en la población o demostrando los efectos adversos asociados con valores particulares. La epidemiología también está siendo clave en la medicina clínica y profesiones afines.

Para tratar con fenómenos complejos es de crucial importancia el uso de variables epidemiológicas. La variación en los patrones de enfermedad es la base de la epidemiología, pero normalmente no se aplica a enfermedades (variables de salida) sino a factores que ayuden a describir y entender el patrón de enfermedad (variables de exposición). Las variables ayudan en el análisis de patrones dentro y entre poblaciones. Se efectúan análisis por edad, sexo, clase social, ocupación, país de residencia, país de nacimiento, etnia, etc. La mayoría de las variables son marcadores para fenómenos complejos que no pueden medirse fácilmente (Woodward, 2013). Por ejemplo, la clase social es un indicador para ingreso, ocupación, educación, hábitos de consumo. El sexo puede relacionarse con características hormonales, psicológicas, genéticas o sociales. Por eso es importante discernir y separar los componentes de cada variable. Una buena variable epidemiológica debería tener impacto en la salud individual y poblacional, ser precisamente medible, diferenciar poblaciones en alguna característica subyacente relevante para la salud (herencia, ingresos, hábitos de salud, etc.) y generar hipótesis etiológicas testeables hasta que no se hayan resuelto temas clínicos y patológicos, como la definición de enfermedad (Porta, Greenland, Hernán, dos Santos Silva, & Last, 2014).

La epidemiología se encuentra en una relación interdependiente con la medicina clínica. La investigación y comprensión de las causas de una enfermedad forma parte de la epidemiología, mientras que la toma de medidas de prevención y control forma parte de la salud pública (Haynes, 2012). Se evidencia la dependencia de la epidemiología al ver que ésta debe presentar evidencia científica de forma que guíe a acciones de salud pública. Por otro lado, una enfermedad desconocida debe definirse con base clínica. Si no hay forma de definirla, no puede diagnosticarse bien y la epidemiología es ineficaz.

Por eso, la primera pregunta para el epidemiólogo incumbe a la validez y a la definición de la enfermedad o problema bajo investigación. La precisión de los diagnósticos depende de la claridad de la definición, pudiendo darse un sesgo por exceso o por defecto. La decisión sobre qué se considera como enfermedad suele ser pragmática (Bhopal, 2002). Para la investigación científica, puede ser adecuada una definición que incluya personas con alta probabilidad de tener la enfermedad. Para salud pública, tal vez no, dado que si se subestiman los casos, se pierden los pacientes con síntomas tempranos, los que más podrían beneficiarse. Por lo tanto, para propósitos epidemiológicos, un caso de enfermedad es:

- Una enfermedad diagnosticada por un médico;
- Un caso dentro de un grupo o brote de casos;

- Un caso que ocurra en una persona que tiene, al menos, dos de tres problemas (por ejemplo: desórdenes gastrointestinales, sarpullido, desorden mental); y
- Un caso sin diagnóstico claro.

El efecto del error en la definición puede ser muy grande, y comparar diferentes poblaciones puede ser engañoso. Cambios en la definición son comunes, cambiando el número de casos, y mostrando cómo muchas veces la definición es pragmática.

### **El concepto epidemiológico de población**

Las poblaciones incluyen individuos aislados, pero una situación así es muy rara. La interacción en sociedad es la piedra basal de la epidemiología como ciencia médica. No pueden hacerse estudios epidemiológicos en una sola persona; para comparar patrones y estados, se necesitan al menos dos individuos. Pueden ser miles, millones, aunque usualmente pueden ser muy pocos. Los casos se analizan para ver patrones en el tiempo, lugares y sub-poblaciones (triada tiempo-espacio-persona), para ver las causas de los patrones, probabilidades de enfermedad y sus causas. Las conclusiones pueden aplicarse directamente al grupo estudiado, pero sólo indirectamente a los individuos, y sólo a individuos razonablemente "típicos" de la población.

La información de características demográficas y socio-económicas dice mucho del estado y riesgo probable de los individuos en la población. Por ejemplo, saber que el promedio de una población es rico ayuda a predecir un patrón de muerte por cáncer o enfermedades del corazón, y no por infecciones o desnutrición. Esto aplicará a nivel poblacional, pero no siempre individual. Estas generalizaciones son posibles porque las características poblacionales influyen los riesgos individuales (Porta y col., 2014).

Los factores que determinan los patrones de una enfermedad suelen ser interacciones complejas entre individuos, su ambiente y su sociedad. Por lo tanto, se necesita un entendimiento detallado de cómo vive la población. Un estudio epidemiológico debería estar basado en poblaciones definidas en términos de edad, localidad, tamaño, sexo y exposición ambiental, entre otras. La idea es definir y utilizar la heterogeneidad inherente de la población y medir características que son explicaciones potenciales de la variación.

Por la gran influencia de la sociedad, también se efectúa una distinción entre individuos enfermos y poblaciones enfermas (Szklo & Nieto, 2014; Berkman y col., 2014). Los atributos de una población se describen tanto por la distribución de frecuencias como por parámetros estadísticos como la media, la desviación estándar, el sesgo, entre otros. La distribución de población de la mayoría de las mediciones biológicas sigue la distribución normal, pero esto debe verificarse en cada caso. A veces, las medidas individuales no tienen utilidad tomadas en forma conjunta (por ejemplo, las huellas dactilares, la personalidad, el color de ojos), porque no siguen un patrón o no tienen una interpretación útil. Las medidas de la población o el medio ambiente también pueden carecer de sentido cuando se aplican a un individuo (densidad de población, PBI, temperatura ambien-

te). Estas medidas sólo pueden aplicarse en su propia escala (Woodward, 2013; Berkman y col., 2014). Algo paradójico es que algunos datos individuales (por ejemplo, el número de personas que viven en un hogar), usados a nivel global (por ejemplo, como densidad de población de una zona), deja de ser significativo a nivel individual (Bhopal, 2002).

Las variables sociales suelen medirse en individuos, las variables ambientales, no. Los desafíos contemporáneos en epidemiología incluyen la medición precisa de la exposición ambiental en las personas, y la medición de características sociales como un conjunto. Las características sociales, incluida la cohesión social, el trabajo en equipo, y el estado de la transición económica, probablemente tengan efectos profundos sobre la salud, pero aún están descritos en forma incompleta (Haynes, 2012). Los avances en epidemiología están siendo impulsados por el re-análisis crítico del concepto de población.

En realidad, no existen poblaciones estáticas como para efectuar una generalización amplia. Las poblaciones de seres vivos son dinámicas en forma inherente. En las comunidades de seres vivos, el tamaño y composición de las poblaciones está cambiando constantemente debido a las muertes, nacimientos y migraciones. Aun si el tamaño de la población permanece fija, la población sigue cambiando. Además, aunque tomemos un grupo fijo, las personas envejecen, por lo que la edad promedio aumenta, con el consiguiente efecto en el patrón de la enfermedad. En forma más gradual, lo mismo puede ocurrir en sociedades enteras (aumento del promedio de vida, aumento de la edad promedio, etc.), lo que se denomina transición demográfica (Bhopal, 2002). Yendo al extremo, consideremos un lugar donde el número de personas es fijo, donde cuando un individuo muere o se va, es sustituido por otro y donde la edad, sexo y composición étnica no varía en el tiempo. Un lugar de este tipo podría ser una cárcel o una institución como un hogar de ancianos. Incluso así, los patrones de salud no son fijos, porque los comportamientos humanos evolucionan constantemente, y su entorno cambia. En resumen, a los fines epidemiológicos, no existen poblaciones humanas estáticas (Woodward, 2013).

Si bien todas las poblaciones son dinámicas, el concepto de población fija es útil en algunos casos (por ejemplo, al pensar en formas de medir la ocurrencia de una enfermedad). Los estudios de cohorte son un intento de crear una población "cerrada". En poblaciones cerradas, las personas que mueren o migran no son reemplazadas (Porta y col., 2014). El cambio de tamaño de la población se puede predecir si sabemos la tasa de muerte y de emigración. La predicción del tamaño en poblaciones "abiertas" (las poblaciones naturales) es mucho más difícil y requiere, además de lo anterior, una estimación de la inmigración y de patrones de fecundidad (Bhopal, 2002). La predicción del tamaño y composición de poblaciones abiertas es un componente clave de la demografía. Si el número de personas entrando está equilibrado por el número de personas saliendo, la población se encuentra en un estado estable. Los cambios en el medio ambiente y en el comportamiento conducen a influencias temporales sobre las pautas de enfermedad. Por ejemplo, las mujeres eran menos propensas a ser fumadoras en los años 20 que en los 50. La combinación de la transición demográfica y los efectos de influencia más amplias (medio ambiente,

comportamiento, etc.), conduce a un profundo cambio en el patrón de enfermedad, y el predominio de ciertas enfermedades crónicas. En muchos países, con el aumento de la riqueza, están predominando enfermedades relacionadas con el estilo de vida opulento y con los grupos de mayor edad. Para la epidemiología, estas transiciones son una ventaja, porque han llevado a una mayor variación en las pautas de morbilidad y exposición a las causas, proporcionando el combustible para el avance epidemiológico (Rothman, 2012; Galea, 2013).

### **Introducción al modelado epidemiológico**

Un reto fundamental en la epidemiología matemática es determinar cómo la estructura de una población determina la transmisión de una enfermedad. Dentro de esta estructura, la componente espacial es fundamental. Si se quiere incluir explícitamente esta componente en el modelo matemático, éste debe formularse entonces en términos de *ecuaciones diferenciales en derivadas parciales* (PDE). Por ejemplo, la ecuación de Bailey

$$\frac{\partial n}{\partial t} = r(n + D_{DC}\nabla^2 n) + D_{DI}\nabla^2 n - \mu n, \quad (1-1)$$

la cual determina un gradiente tanto de individuos como de interacción entre ellos. Este tipo de ecuaciones permite capturar la dinámica en forma precisa, pero el tratamiento matemático es usualmente muy complicado y en muchos casos innecesario. Efectuando una serie de hipótesis simplificadoras es posible modelar el sistema de forma satisfactoria por medio de modelos compartimentales basados en ecuaciones diferenciales ordinarias (Ordinary Differential Equation - ODE) en donde se supone una mezcla homogénea entre los diferentes compartimentos. En este tipo de modelos el *número reproductivo*  $R_0$  es el número promedio de casos nuevos que genera un caso dado lo largo de un período infeccioso y es determinado por la infectividad intrínseca de un solo caso y del ambiente en donde se produce suponiendo que los individuos son idénticos con respecto a la infectividad. Estos modelos también incorporan la hipótesis de una mezcla homogénea, a veces llamada la *ley de acción de masas*, en el que la tasa de aumento en la incidencia de la epidemia es proporcional al producto del número de personas infecciosas y susceptibles. Este supuesto ha sido relajado, pero no se elimina de los modelos clásicos porque esta hipótesis es robusta en el sentido de que es compatible con numerosos escenarios para la transmisión de persona a persona de la enfermedad. Esta hipótesis equivale a un modelo en el que todos los individuos de una población están en contacto entre si y tienen probabilidades idénticas de transmisión en contacto con un individuo infeccioso. Aunque esta hipótesis no es realista, facilita el análisis matemático y, en algunos casos, ofrece una aproximación razonable. Sin embargo, las poblaciones pueden ser muy heterogéneas en cuanto a la susceptibilidad, la infectividad, tasas de contacto, y los modelos sencillos que suponen mezcla homogénea no permiten la variación de estos parámetros. La heterogeneidad en la susceptibilidad y la infectividad es una característica importante de muchas enfermedades infecciosas y ha sido considerado para mejorar la precisión de los modelos

epidemiológicos. En el análisis de estos modelos, el enfoque ha sido puesto en el impacto de la heterogeneidad en el tamaño final de las epidemias, así como sobre sus consecuencias para el control de enfermedades, y la interpretación de datos. En el contexto de los modelos epidemiológicos, se ha demostrado que el tamaño final de la epidemia se reduce cuando el riesgo de infección es heterogéneo en la población, tanto para las formulaciones deterministas como para las estocásticas. Los nuevos modelos capturan algunas, pero no todas estas características. Un modelo epidemiológico debería incorporar todas estas características:

- Un individuo dado puede tener contacto con un número limitado de otros individuos de la población en un instante determinado y los contactos que resultan en la transmisión de la enfermedad son eventos usualmente cortos y repetitivos;
- El número y frecuencia de contacto entre individuos puede ser variable;
- El número de contactos puede cambiar con el tiempo; y
- Los individuos infectados tienen el potencial de transmitir el patógeno a los individuos susceptibles con los que entre en contacto.

Los modelos basados en individuos tienen la capacidad de poder modelar todos estos supuestos y es por ello que están siendo utilizados por muchos investigadores como una alternativa razonable a la hora de modelar epidemias.

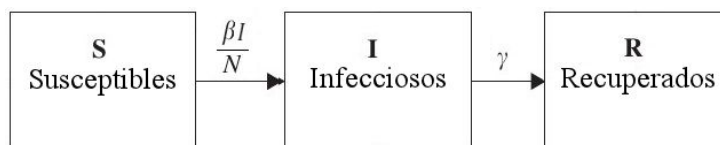
## 1.2. Estado del arte

### 1.2.1. Modelos clásicos

Los modelos epidemiológicos basados en ecuaciones diferenciales ordinarias suelen llamarse en la bibliografía *modelos compartimentales*, entendiendo por compartimiento a una región con un grupo de individuos distribuidos uniformemente. Los modelos compartimentales realizan con las siguientes hipótesis (Galvani & May, 2005a):

- La distribución de los individuos es uniforme en cada compartimiento;
- La distribución de individuos al ingresar a un compartimiento es instantánea;
- La tasa de paso de un material de un compartimiento a otro se relaciona con la cantidad de material existente; y
- Existe una ley de conservación, pero no es estrictamente necesario.

Los modelos clásicos son aproximaciones no espaciales que se concentran en la dinámica temporal de una epidemia. Estos modelos están basados en poblaciones, con la población dividida en un número limitado de compartimentos o grupos. La división mas simple consiste en tres compartimentos excluyentes. Un segmento de la población incluye a aquellos individuos que son susceptibles a la enfermedad ( $S$ ). Un segundo segmento de la población incluye a todos los individuos infectados con capacidad de transmitir la enfermedad a los individuos susceptibles ( $I$ ) y el segmento restante se incluye a los individuos recuperados ( $R$ ) (Figura 1-1).



**Figura 1-1.:** Flujo de población en el modelo SIR estático

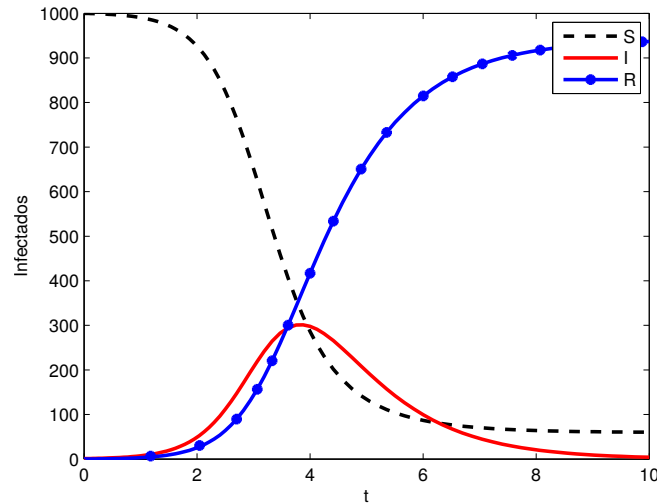
Este modelo de tres compartimentos, donde la unidad de modelado es cada segmento de la población, se conoce comúnmente como modelo  $SIR$ . Para cualquier instante de tiempo  $t$  los individuos se mueven del compartimento  $S$  al  $I$  a  $R$  a una tasa constante. Suponiendo que es una dinámica continua, el sistema puede ser modelado a través de un conjunto de ecuaciones diferenciales de la forma:

$$\begin{aligned}
 \frac{dS}{dt} &= -\beta S(t)I(t), & S(0) &= S_0 \\
 \frac{dI}{dt} &= \beta S(t)I(t) - \gamma I(t), & I(0) &= I_0 \\
 \frac{dR}{dt} &= \gamma I(t).
 \end{aligned}
 \tag{1-2}$$

En este modelo  $S$ ,  $I$  y  $R$  denotan las poblaciones susceptible, infectada y recuperada respectivamente,  $\beta$  es la tasa de infección de los individuos susceptibles y  $\gamma$  es la tasa de recuperación de individuos infectados. El tamaño de las poblaciones  $S$ ,  $I$  o  $R$  pueden estar disponibles o bien pueden ser estimados a partir de alguna observación disponible. Los parámetros  $\beta$  y  $\gamma$  son las incógnitas en este sistema de ecuaciones y deben ser estimados a partir de las mediciones y datos disponibles

Durante el modelado, el número de nuevos casos de infección observados por período de tiempo se puede visualizar como una curva que depende del tiempo. El tamaño de la población infectada  $I$  tiende a incrementarse luego del inicio de la epidemia y luego tiende a disminuir una vez que esta ha alcanzado su pico. Ajustando el valor de los parámetros del modelo la dinámica temporal del mismo puede ajustarse con la dinámica real observada en la epidemia (Figura 1-2). Los valores de estos parámetros son derivados para análisis subsecuentes, analizando sus implicaciones en el proceso epidemiológico. Como el compartimento poblacional es la unidad de modelado en este tipo de enfoque,





**Figura 1-2.:** Dinámica temporal de modelo SIR

los valores que toman los parámetros del modelo son usualmente estimados como un promedio sobre el total de la población.

Los modelos *SIR* son los más simples de una familia de modelos basados en poblaciones. Modelos más complejos pueden tener en cuenta consideraciones adicionales (Satsuma, Willox, Ramani, Grammaticos, & Carstea, 2004; Casagrandi, Bolzoni, Levin, & Andreasen, 2006). Aparte de ello, las bases de este tipo de enfoque han permanecido sin cambios a lo largo de la epidemiología moderna. La idea central de este tipo de modelos es utilizar población agregada en forma de compartimentos como unidad de modelado, concentrándose de esta manera en los aspectos temporales del fenómeno y dejando de lado los factores espaciales. Los modelos poblacionales no consideran las interacciones que se puedan dar entre individuos.

Una serie de suposiciones con respecto al comportamiento de los individuos están implícitas en este tipo de modelos. Por un lado, se supone que los individuos que pertenecen a un compartimento en particular son idénticos entre sí. Por otra parte, todos los individuos interactúan entre sí, esta hipótesis se conoce como *mezcla homogénea*. Las hipótesis relacionadas con la distribución espacial no son explícitamente abordadas por este tipo de modelos pero pueden ser inferidas.

El enfoque determinista y el reducido número de parámetros que utiliza hace a estos modelos simples y de fácil implementación. Los mismos pueden reproducir de manera razonable la dinámica temporal de una epidemia y reflejar el estado de salud de promedio de la población de manera acertada. Estos modelos han sido utilizados ampliamente durante los últimos 100 años y han servido como piedra fundamental de la epidemiología moderna. De estas simplificaciones surgen una serie de modelos compartimentales aplicados a epidemias. Entre los más utilizados se encuentran los modelos *SEIR*. En este

tipo de modelos, los individuos se encuentran divididos en cuatro grupos (Li & Jin, 2005):

- **Susceptibles ( $S$ ):** Individuos que están sanos pero que pueden contraer la enfermedad. Esto implica que aún no han contraído la enfermedad.
- **Expuestos ( $E$ ):** Individuos que portan la enfermedad pero que, al hallarse en su período de incubación, no muestran síntomas y no infectan a otros.
- **Infeciosos ( $I$ ):** Individuos que tienen la enfermedad y pueden transmitirla.
- **Recuperados ( $R$ ):** Se agrupa con esta denominación a los individuos que se han curado y están inmunizados, y a los individuos que están enfermos pero en cuarentena (por lo que no infectan a otros).

Las hipótesis básicas realizadas para construir estos modelos son (Li, Wang, & Jin, 2006):

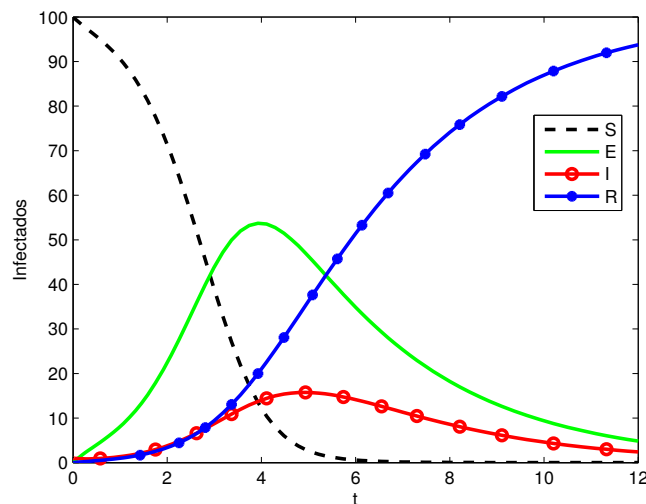
- La enfermedad que se modela confiere inmunidad a las personas que se recuperan.
- La población total  $N$  es constante,  $S(t) + E(t) + I(t) + R(t) = N \forall t$ , Esto implica que los nacimientos y muertes en el período son despreciables, y que la enfermedad no es mortal;
- La población se encuentra uniformemente distribuida en la región considerada;
- La tasa de incremento de  $E$  es proporcional al número de contactos entre  $S$  y  $I$  (dado que el flujo es unidireccional, todo nuevo expuesto proviene de la población susceptible); y
- Las tasas de disminución de  $S$ ,  $E$  e  $I$  son proporcionales a  $S$ ,  $E$  e  $I$ , respectivamente (ya que a mayor población, mayor cantidad de individuos se ven expuestos, infectados o recuperados en cada caso). Asimismo, el aumento de  $R$  es igual a la disminución de  $I$ .

Las ecuaciones que describen un modelo de tales características son (Zhang & Teng, 2008):

$$\begin{aligned}
 \frac{dS}{dt} &= \frac{\beta}{N} S(t)I(t), & S(0) &= S_0 \\
 \frac{dE}{dt} &= \frac{\beta}{N} S(t)I(t) - \varepsilon E(t), & E(0) &= E_0 \\
 \frac{dI}{dt} &= \varepsilon E(t) - \gamma I(t), & I(0) &= I_0 \\
 \frac{dR}{dt} &= \gamma I(t).
 \end{aligned}
 \tag{1-3}$$

Es decir, a partir del momento de inicio de la infección,  $S$  sólo puede disminuir, porque una vez recuperado, el individuo no puede infectarse nuevamente. La tasa de disminución es proporcional al producto entre el tamaño de la población sana y la fracción

infectada en cada momento ( $I/N$ ), pues esto determina la fracción de contactos potencialmente infecciosos y, por lo tanto, la velocidad de infección. El aumento del número de individuos expuestos es, como previamente se explicó, igual a la disminución de individuos sanos. Al mismo tiempo, existe una tasa de disminución de individuos expuestos, determinada por la duración del período de latencia y el ingreso de los individuos expuestos a la población de individuos infectados. Esta tasa es proporcional a la población misma de expuestos, ya que cuantos más individuos expuestos haya, mayor cantidad pasará al estado infeccioso (Figura 1-3). La tasa de disminución de individuos expuestos



**Figura 1-3.:** Dinámica temporal de modelo SIR

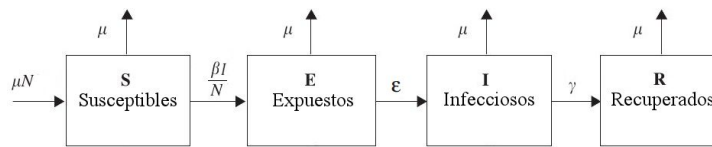
corresponde a la de aumento en individuos infecciosos. Los infecciosos pasan al compartimento de recuperados también a su propia tasa, proporcional al número de infecciosos. La tasa de individuos que se recuperan es proporcional al número de infectados en cada momento, por lo que el término es idéntico al de disminución de infecciosos, pero con signo opuesto. Los parámetros que afectan a cada término de la ecuación pueden tomarse como constantes o variantes en el tiempo, y dependen de la naturaleza de la enfermedad, de la población, su distribución, hábitos y zona geográfica.

Si la duración de la enfermedad es lo suficientemente prolongada, deben tenerse en cuenta las tasas de nacimiento y defunción, porque afectan a la evolución temporal de la epidemia. Durante el transcurso de la epidemia, los nacimientos se incorporan directamente a la población susceptible. Al mismo tiempo, las defunciones pueden ocurrir en todas las poblaciones, por lo que un término relacionado con la tasa de defunción debe ser incluido en cada ecuación. Como aún se considera a la población constante, la tasa de nacimientos debe estar balanceada con la tasa de defunciones. Estas hipótesis conducen al siguiente

modelo (Li & Muldowney, 1995):

$$\begin{aligned}
 \frac{dS}{dt} &= \mu(N(t) - S(t)) - \frac{\beta}{N}S(t)I(t), & S(0) &= S_0 \\
 \frac{dE}{dt} &= \frac{\beta}{N(t)}S(t)I(t) - (\varepsilon + \mu)E(t), & E(0) &= E_0 \\
 \frac{dI}{dt} &= \varepsilon E(t) - (\gamma + \mu)I, & I(0) &= I_0 \\
 \frac{dR}{dt} &= \gamma I(t) - \mu R(t).
 \end{aligned}
 \tag{1-4}$$

En este modelo se ha incorporado un término de nacimientos, proporcional a la población  $N$ . Se considera que todos los nacimientos se incorporan a la población susceptible, por lo que este término se incorpora a la primer ecuación del sistema de ecuaciones (1-3). Se incorpora también un término de decesos en cada ecuación del sistema (1-4), proporcional a la población en cada compartimiento. Se considera que todos están afectados por la misma tasa de defunciones. De este modo, en cada compartimiento ya no se resta sólo la población que pasa el siguiente compartimiento sino la población que fallece.



**Figura 1-4.:** Flujo de población en el modelo SEIR dinámico

La tasa se considera igual para los nacimientos y para las defunciones, por lo que el parámetro usado es el mismo. En total, las defunciones en cada compartimiento dan cuenta de la tasa en toda la población. El flujo del modelo así formulado queda representado en la figura 1-4. Si fuera necesario considerar distintas tasas para los nacimientos y las defunciones, o incluso si varían en el tiempo, el modelo se modifica simplemente reemplazando el parámetro  $\mu$  por el parámetro de natalidad en el primer término de la ecuación correspondiente a la población susceptible (Li, Graef, Wang, & Karsai, 1999):

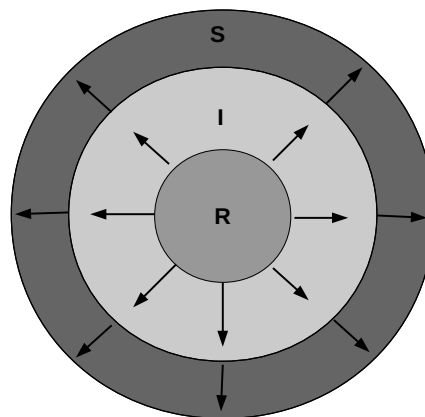
$$\frac{dS}{dt} = \nu N(t) - \frac{\beta}{N(t)}S(t)I(t) - \mu S(t).
 \tag{1-5}$$

Es este caso ya no se cumple el requerimiento de población constante de los modelos compartimentales ( $S + E + I + R = N \forall t$ ).

### **Modelos de ondas**

En la década de 1980, desde la geografía se propuso un nuevo marco espacial para los modelos epidemiológicos, este marco debía considerar explícitamente la dispersión espacial en los procesos infecciosos. En su forma mas simple, estos modelos espaciales pueden

describirse como un sistema de tres anillos concéntricos también conocido como ola de tres anillos. La primera infección ocurre en el anillo central y desde allí la epidemia se extiende a los anillos exteriores en todas las direcciones en forma de ola. El segmento de población infecciosa se encuentra en la cresta de la ola, la población susceptible se sitúa en el frente de la cresta, y el segmento de población recuperada se sitúa en la parte posterior de la cresta. A un instante de tiempo determinado, los tres segmentos forman un patrón de tres anillos concéntricos. En el anillo del medio se encuentra la porción de población infecciosa, el anillo exterior contiene a la porción de población susceptible y la población de recuperados se encuentra en el espacio central alrededor del punto donde comienza el fenómeno. La ubicación de los tres segmentos poblacionales, o los tres anillos, cambia de manera dinámica durante el proceso infeccioso a medida que este se extiende a lo largo del espacio (Rhodes & Anderson, 1997). El modelo de tres anillos proyecta el modelo po-



**Figura 1-5.:** Diagrama de un modelo poblacional basado en ondas

blacional clásico en el espacio, extendiendo su enfoque temporal no espacial a un nuevo enfoque que considera el espacio de manera explícita. Excepto por las consideraciones referidas a las características espaciales de una epidemia, los modelos de olas poseen todos los principios de diseño de los modelos poblacionales clásicos en términos del uso de compartimentos como unidad de modelados, la imposibilidad de interacción entre individuos, y su capacidad de modelar la dinámica temporal. Todas las hipótesis que los modelos clásicos hacen con respecto a los individuos son adoptados por estos modelos. Las enfermedades, como fenómeno, se mueven a lo largo del espacio, pero cada individuo que se infecta y subsecuentemente transmite la enfermedad permanece inmóvil. La movilidad de las enfermedades es representada como cambios en el estado (susceptible, infeccioso o recuperado) de estos individuos inmóviles formando un patrón espacial similar al de las ondas mecánicas que se forman en la superficie de un estanque cuando se arroja una piedra, donde las partículas de agua vibran verticalmente con movimiento horizontal nulo. Las ondas de enfermedad se pueden modelar de manera similar (Figura 1-5).

Los modelos poblacionales basados en olas marcaron el inicio de los modelos epidemiológicos con representación explícita del espacio. Estos modelos tienen un marco conceptual que contempla una orientación espacial explícita que puede ser implementado a través de métodos de análisis espacial, por ejemplo autómatas celulares (Tobler, 1970) y de microsimulación (Amrhein & MacKinnon, 1988).

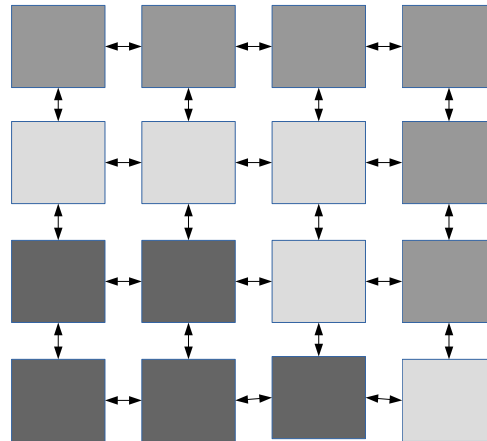
En la pasada década, los modelos basados en poblaciones han cosechado un número creciente de críticos. El argumento en el que se basan las críticas apunta a la hipótesis de homogeneidad, lo que acota su uso a un número reducido de problemas, ya que no explican la heterogeneidad espacio temporal observada en la transmisión de enfermedades infecciosas (Holmes, 1997; Fuks & Lawniczak, 2001; Kikuchi, Tominaga, Arita, Takahashi, & Tomita, 2003; Dye & Gay, 2003; Malcovati y col., 2003; Kothapalli, Kirkness, Natale, & Meyers-Wallen, 2003; Galvani, Ferrini, & Ottani, 2004; Cordero, Koopman, & van der Helm, 2004; Galvani & May, 2005b). La fortaleza de estos modelos proviene de su habilidad para predecir el estado de salud a nivel poblacional. Una aplicación para estos modelos son epidemias en donde la enfermedad se esparce a través del espacio de la misma manera que una ola, como por ejemplo la pandemia de gripe española de 1918 – 1920.

### **Modelos sub-poblacionales**

Los modelos sub-poblacionales dividen a la población en un número substancial de sub-poblaciones, donde una sub-población es la unidad de modelado básica. Estos modelos surgen hacia finales de la década de 1990 y tienen como objetivo aumentar la heterogeneidad en una población de manera de producir resultados mas realistas que los obtenidos mediante modelos poblacionales clásicos. Una población puede ser dividida en poblaciones mas pequeñas usando diferentes criterios. Aquellos modelos que dividen la población mediante un criterio basado en consideraciones espaciales se denominan *modelos estructurados espacialmente* (Keeling, 2000).

Gran parte de estos modelos estructurados dividen el espacio en celdas que contienen una sub-población. Los modelos de sub-población representan la evolución natural de los modelos clásicos pero con una unidad de modelado mas pequeña, haciendo uso de los mismos principios de diseño que los modelos poblaciones basados en olas, mientras que aparentan ser desarrollados de forma independiente de estos. Con un número de unidades de modelado mayor, las interacciones entre las mismas son modeladas como entradas a estos modelos, sin embargo el mecanismo de interacción usualmente no es explícitamente explicado. Las enfermedades pasan entre las sub-poblaciones a través de interacciones entre las celdas de la grilla y de esta manera se mueve a través del espacio, mientras que las celdas permanecen inmóviles. En este sentido, los modelos sub-poblacionales heredan las hipótesis de los modelos clásicos pero a nivel sub-poblacional, esto quiere decir que los individuos en una misma celda de la grilla se suponen idénticos, mezclados de manera homogénea, distribuidos de manera homogénea e inmóviles. En la Figura 1-6 se puede observar un diagrama de un modelo sub-poblacional en don-

de cada celda representa una porción de la población total en un estado epidemiológico determinado y las flechas representan las relaciones establecidas entre las unidades.



**Figura 1-6.:** Diagrama de modelo sub-poblacional

La partición de una población en una escala mas pequeña de modelado y las interacción entre estas unidades distinguen a los modelos sub-poblacionales de los modelos poblacionales basados en ondas, aunque las hipótesis de mezcla homogénea y la aproximación determinista se conservan en cada celda. El cambio en la unidad de modelado añade al modelo heterogeneidad espacial a la estimación del estado de salud de la población y tiene un gran impacto en muchos aspectos del modelado epidemiológico. Con ciertas limitaciones, los modelos sub-poblacionales brindan resultados mas realistas (Grenfell & Harwood, 1997; Keeling, 2000). Debido a la hipótesis de heterogeneidad entre sub-poblaciones, los modelos sub-poblacionales son muy buenos para modelar escenarios en donde un patógeno se dispersa entre comunidades inmóviles con una alta densidad de población. De manera alternativa, estos modelos son buenos para modelar la dinámica eco-epidemiológica de enfermedades que afectan animales de corral, por ejemplo, se han utilizado para modelar fiebre aftosa en vacas y ovejas. En estas aproximaciones el ganado se supone perfectamente mezclado y cada celda o unidad de modelado representa un granja, de esta forma la enfermedad se transmite entre granjas lindantes (Keeling y col., 2001; Doran & Laffan, 2005).

Casi una década después, especialmente luego de toda la discusión que surgió sobre las limitaciones de los modelos clásicos, una variedad de nuevos diseños han sido propuestos para los modelos sub-poblacionales. Entre los cambios propuestos se plantea la modificación de la forma de transmisión entre células vecinas, permitiendo explícitamente el intercambio de individuos entre celdas (Doran & Laffan, 2005; Watts, Muhamad, Medina, & Dodds, 2005; Colizza, Barrat, Barthelemy, Valleron, & Vespignani, 2007; Mao & Bian, 2010). La transmisión de la enfermedad entre celdas es modelada utilizando aproximaciones probabilísticas basadas en reglas.

Otra revisión realizada sobre estos modelos que surgió con el propósito de añadir heterogeneidad al mismo es la de fijar ciertas características de los individuos de cada unidad de modelado siguiendo alguna distribución de densidad de probabilidad particular (tamaño de la población, número de individuos por hogar, estructura etaria, etc). Los individuos dentro de una celda están colectivamente representados de acuerdo a estas estadísticas (Ferguson y col., 2006). Las interacciones entre sub-poblaciones y la movilidad de los individuos también está representada por funciones probabilísticas que siguen ciertas propiedades estadísticas. Estas modificaciones, las cuales surgieron en su mayoría a mediados de la década pasada, han incorporado en varios niveles procesos de interacción estocásticos y de esta forma han alterado las hipótesis de homogeneidad realizadas por los primeros modelos sub-poblacionales.

### Modelos sub-poblacionales móviles

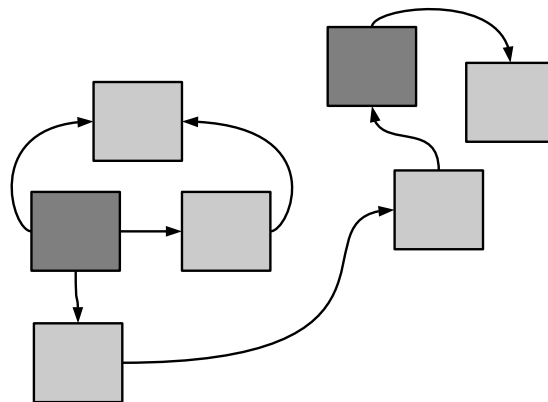


Figura 1-7.: Diagrama de modelo sub-poblacional móvil

Estos modelos identifican una o más sub-poblaciones como la fuente de infección. Con el tiempo, estas sub-poblaciones infectadas pueden moverse a diferentes ubicaciones. En las nuevas ubicaciones, una sub-población infectada puede unirse a una sub-población no infectada o dividirse en varias sub-poblaciones infectadas que también se mueven, se unen a otras sub-poblaciones o se vuelven a dividir. De esta manera, las enfermedades son transmitidas a través del espacio y el tiempo. Debido a que estos modelos tienden a enfocarse en donde las unidades se unen o dividen, la ubicación de una unidad siempre es representada de forma explícita. A pesar de estas diferencias, los modelos sub-poblacionales móviles hacen uso de hipótesis idénticas a las de los modelos sub-poblacionales y los modelos basados en olas a nivel sub-poblacional, con excepción del movimiento de las unidades. De manera más específica, los individuos de una sub-población son idénticos, se encuentran homogéneamente mezclados y homogéneamente distribuidos. La hipótesis de movilidad le brinda a este tipo de modelos sub-poblacionales la flexibilidad para modelar tanto la dinámica temporal como la dinámica espacial de un proceso epidémico. Dado



que la hipótesis de homogeneidad es mantenida, estos modelos son más efectivos cuando son aplicados en situaciones en donde se modelan sub-unidades con alta densidad poblacional y móviles, como por ejemplo, campos de refugiados. Las unidades militares infectadas, por ejemplo, pueden verse como sub-poblaciones móviles que se transfieren a una locación central donde se unen a otras unidades militares. De manera subsecuente algunas de las nuevas unidades infectadas en la locación central pueden transferirse a otras locaciones para unirse a otras unidades y de manera subsecuente difundir la enfermedad por el espacio (Smallman-Raynor & Cliff, 2001).

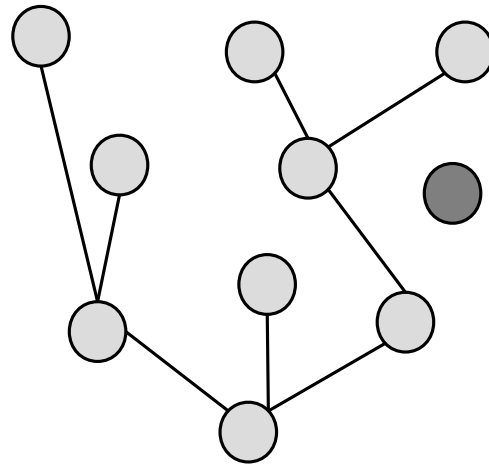
### **Modelos de redes de contacto**

El concepto de *red de contactos* es cada vez más popular en áreas relacionadas a las ciencias de la salud. El concepto se utiliza a menudo como sinónimo de *parentesco*, *colaboración*, o incluso *grupo*. Pero otras veces, se utilizan con la intención más específica de describir las relaciones que existen entre grupos de individuos u organismos, y los recursos a los que miembros de estos grupos tienen acceso. El análisis de redes se ha vuelto popular en la epidemiología, especialmente en HIV y tuberculosis. La estructura es derivada de las regularidades en los patrones de relación entre las diferentes entidades que componen la población, que pueden ser personas, grupos u organizaciones. El análisis de redes sociales es cuantitativo y ha sido ampliamente utilizado en campos como la sociología y la matemática.

El concepto de red social refiere a una estructura social compuesta por un conjunto de actores (tales como individuos u organizaciones) que están relacionados de acuerdo a algún criterio (relación profesional, amistad, parentesco, etc.). Normalmente se representan simbolizando los actores como nodos y las relaciones como líneas que los unen. El tipo de conexión representable en una red social es una relación diádica o lazo interpersonal. Los *actores* son miembros distintos de la red, por ejemplo, personas, clientes, vecinos, ciudadanos, etc; o bien, unidades colectivas, por ejemplo, organizaciones, unidades de un ejército, escuelas, etc. Los *lazos relacionales* unen actores en la red. Estos lazos pueden ser informales (por ejemplo, cuando personas de un grupo conocen a personas de otro grupo) o formales (cuando por ejemplo un grupo se encuentra con otro grupo, una persona se encuentra con otra persona). En la Figura 1-8 se puede observar un ejemplo de una red compuesta por 10 actores en la que todos menos uno que se encuentran relacionados de forma directa o indirecta.

El análisis de redes sociales estudia estructuras sociales aplicando la *teoría de grafos* e identificando las entidades o actores como *nodos* o *vértices* y las relaciones o lazos como *enlaces* o *aristas*. De esta forma se pueden aplicar todas las técnicas existentes para el estudio de grafos al análisis de sistemas sociales complejos. La estructura del grafo resultante es a menudo una red compleja. Como se ha dicho, en su forma más simple una red social es un mapa de todos los lazos relevantes entre todos los nodos estudiados.

- Un lazo puede ser llamado *conector* si su eliminación causa que los puntos que conecta se transformen en componentes distintos de un grafo;



**Figura 1-8.:** Diagrama de red de contactos de 10 individuos

- La *centralidad* representa una de las medidas de la importancia que un nodo tiene dentro de una red, dependiendo de la ubicación dentro de ésta;
- La *centralización* mide la diferencia entre el número de enlaces para cada nodo, dividido la cantidad máxima posible de diferencias. Dicho de otra forma, una red centralizada tendrá muchos de sus vínculos dispersos alrededor de uno o unos cuantos nodos, mientras que una red descentralizada es aquella en la que hay poca variación entre el número de enlaces de cada nodo posee;
- El *coeficiente de agrupamiento* de una red mide la probabilidad de que dos actores vinculados a un nodo se asocien a sí mismos. Un coeficiente de agrupación más alto indica un mayor *exclusivismo*;
- La *cohesión* es el grado en que los actores se conectan directamente entre sí por vínculos cohesivos. Los grupos se identifican como *cliques* si cada individuo está vinculado directamente con cada uno de los otros, *círculos sociales* si hay menos rigor en el contacto directo y este es impreciso, o bloques de cohesión estructural si se requiere la precisión;
- La *densidad* mide el grado de relaciones de un actor de la red. La densidad de la red, o densidad global, es la proporción de vínculos en una red en relación con el total de vínculos posibles (redes escasas contra densas);
- El *flujo de centralidad de intermediación* mide el grado en que un nodo contribuye a la suma del flujo máximo entre todos los pares de nodos (excluyendo ese nodo).

### Modelos epidemiológicos como redes de contactos

Las epidemias causadas por la transmisión de agentes infecciosos están marcadas por los efectos de la heterogeneidad. La heterogeneidad en los patógenos, la población afectada y las interacciones entre los individuos afectan profundamente las dinámicas de la infección. Por ejemplo, en 1984, en un estudio sobre la propagación de gonorrea en los Estados Unidos, mostró que un 60 % del total de las infecciones fueron causadas por un grupo pequeño de individuos que solo componían el 2 % de la población total (Hethcote & Yorke, 1984). En 1999, un grupo de cinco casos de sarampión en una pequeña escuela primaria en Los Países Bajos dio lugar a un brote de aproximadamente 3000 casos, a pesar del hecho de que el 96 % de la población en los Países Bajos había sido vacunada contra el sarampión hasta ese momento (Control & (US), 2000). En la primavera de 2003, dos personas que viajan desde Vancouver y Toronto, respectivamente, se infectaron casi simultáneamente con el SARS por una fuente en Hong Kong. No hubo casos secundarios reportados desde el individuo de Vancouver; Por el contrario, el individuo infectado de Toronto llegó a infectar a otras cinco personas, lo que dio lugar a un brote de alrededor de 200 casos en Toronto (Poutanen y col., 2003).

Estos son algunos de los casos que muestran como las enfermedades se propagan a través de las poblaciones humanas a través de estrechas interacciones físicas. Los patrones de contacto interpersonal que subyacen a la transmisión de enfermedades de forma natural pueden ser considerados para formar una red, donde los enlaces unen a los individuos que interactúan entre sí. Durante un brote, la enfermedad se extiende a lo largo de estos enlaces. Todos los modelos epidemiológicos hacen suposiciones acerca de la red subyacente de interacciones, a menudo sin indicarlas explícitamente. Los modelos de redes de contactos, sin embargo, permiten formalizar matemáticamente este concepto intuitivo de modo que los cálculos epidemiológicos puedan considerar explícitamente los patrones complejos producto de las interacciones.

Formalmente, modelo de red de contactos (o red social) representa explícitamente las interacciones entre huéspedes que median la propagación de una enfermedad. Un nodo en una red de contacto representa un individuo, y un enlace entre dos nodos representa una interacción que puede permitir la transmisión de enfermedades. El grado de nodo es el número de nodos unidos a este (es decir, su número de contactos) y el grado de distribución de una red es la frecuencia de distribución de grados a lo largo de toda la población. Un modelo de red de contactos exacta requiere el conocimiento de cada individuo de una población y de todos los contactos que causan enfermedades entre los individuos (por ejemplo, estornudos en el caso de las enfermedades transportadas por el aire o el contacto sexual en el caso de las enfermedades de transmisión sexual). Incluso para poblaciones pequeñas, que suele ser inviable, y por lo tanto los investigadores suelen trabajar con redes aproximadas. Existen varias técnicas para recopilar la información necesaria para construir modelos de redes de contactos realistas. Entre ellas se incluye por ejemplo el seguimiento de todos los individuos infectados y sus contactos durante o después de un brote (Klovdahl y col., 1977), consultar a los individuos de una población (Eubank y col.,

2004) por medio del censo con el objetivo de extraer las principales características sociales (Gornick & Meyers, 2003; Crompton, 2008).

La caracterización de la estructura de red se ha convertido en una industria artesanal multidisciplinaria, con investigadores del área de epidemiología, sociología, biología, informática y física en busca patrones en los grandes volúmenes de datos. Los investigadores buscan a menudo propiedades estadísticas globales de la red, y han prestado especial atención a las denominadas redes *small-world*, caracterizadas por altos niveles tanto de agrupamiento local y conectividad global (Watts & Strogatz, 1998) y las redes *free-scale* que se caracterizan por presentar distribuciones de grado que siguen alguna ley y suelen presentar pequeñas fracciones de los grupos altamente conectados (Barabási & Albert, 1999).

Las redes *free scale* se han utilizado en muchos sistemas tecnológicos (por ejemplo, Internet, la World Wide Web) y sistemas biológicos (por ejemplo, cascadas metabólicas, interacciones de proteínas, regulación de la transcripción y dominio de proteínas) (Barabási & Albert, 1999; Wang, Birol, & Hatzimanikatis, 2004; Luscombe y col., 2004). Estas redes altamente estructuradas a menudo contrastan con tres clases de redes *nulas*: (i) redes en la que todos los nodos tienen el mismo grado, y cualquier nodo dado está conectado a los nodos físicamente próximos, (ii) las redes aleatorias regulares en los que todos los nodos tienen el mismo grado, pero cualquier nodo seleccionado está conectado a nodos elegidos al azar a lo largo de la red, y (iii) las redes aleatoria de Poisson (también llamados grafos aleatorios Erdos-Renyi) en las que se asignan un número total específico de conexiones es asignado a los nodos completamente al azar, produciendo un grado de distribución de Poisson en toda la red.

Las estructuras de las redes de contactos humanas, sin duda, juegan un papel crucial en la transmisión de enfermedades (Pastor-Satorras, Vázquez, & Vespignani, 2001; Barrat, Barthelemy, Pastor-Satorras, & Vespignani, 2004; Ferrari y col., 2006). Por ejemplo, los epidemiólogos han dado cuenta hace tiempo que el umbral de epidemia, el valor crítico para la tasa de infección por encima del cual una enfermedad puede propagarse y persistir, disminuye a medida que la desviación estándar del grado de la distribución aumenta (Hethcote, 2000). La pregunta obvia es, por tanto: ¿cuáles son las estructuras de las redes de contacto en el mundo real? El campo se ha centrado sobre todo en las redes aleatorias de escala libre (Moreno, Pastor-Satorras, & Vespignani, 2002; Lloyd-Smith, Schreiber, Kopp, & Getz, 2005), basado en la aparente ubicuidad de este tipo de redes en los sistemas naturales y sistemas intervenidos por el ser humano y un limitado conjunto de estudios de patrones de contacto epidemiológico relevantes (Liljeros, Edling, Amaral, Stanley, & Åberg, 2001; Albert & Barabási, 2002). Estas redes se caracterizan por la presencia un número anómalamente alto de potenciales contactos causantes de enfermedades, llamados **súper spreaders**, que tienen importantes implicaciones epidemiológicas (Lloyd-Smith y col., 2005; Stein, 2011). Con varianza grande o infinita en grado, las redes de escala libre pueden tener umbrales epidémicos muy bajos o inexistentes. Esto signifi-

ca que incluso las redes más dispersas son muy vulnerables a las epidemias. A pesar de su reciente popularidad, sin embargo, no está claro que las redes epidemiológicas reales sean generalmente de escala libre.

### **Los modelos clásicos desde la óptica de las redes**

El modelo *SIR* como una red de contactos puede reescribirse como

$$\begin{aligned}\frac{dS}{dt} &= -\lambda S(t), & S(0) &= S_0 \\ \frac{dI}{dt} &= \lambda S(t) - \gamma I(t), & I(0) &= I_0 \\ \frac{dR}{dt} &= \gamma I(t),\end{aligned}\tag{1-6}$$

donde  $\gamma$  es la tasa de recuperación de los individuos infectados que se obtiene una distribución exponencial de los períodos infecciosos,  $\lambda$  puede ser pensado como un parámetro compuesto de tres factores:

$$\lambda = \alpha \iota_t \tau,\tag{1-7}$$

donde  $\alpha$  es el número de personas con las que un individuo susceptible tiene un contacto eficaz (transmite la enfermedad);  $\iota_t = I(t)/N$  es la proporción de contactos que son infecciosos; y  $\tau$  es la tasa por contacto en la que se transmite la enfermedad entre un individuo infeccioso e individuos susceptibles (Keeling y col., 2001). Este modelo supone mezcla homogénea entre los individuos. Tomado literalmente, el supuesto significa que en cualquier punto en el tiempo, cada individuo susceptible tiene la misma probabilidad de ponerse en contacto con otro individuo de la población. Tomado de manera más flexible, es lo mismo que decir que todas las personas susceptibles tienen patrones de contacto comparables ( $\alpha$ ) y se encuentran con personas infectadas a una velocidad que corresponde a la prevalencia general  $\iota_t = I(t)/N$ . Las ecuaciones estándar dadas anteriormente pueden ser aproximadas a un modelo de red en la que todos los individuos tienen un número idéntico de contactos (red de contactos aleatoria) (Bansal, Grenfell, & Meyers, 2007). Considerando una red aleatoria regular con grado homogéneo  $k$ . Supongamos que la enfermedad se propaga desde los nodos infectados a los contactos susceptibles con una tasa  $\tau'$ . Como los contactos son aleatorio, podemos asumir que la fracción de contactos que es infectada es igual a la fracción de individuos infectados en la red completa, o  $\iota_t = I(t)/N$ . De esta forma, la fuerza de infección sobre cualquier individuo susceptible dado será  $\lambda' = k \iota_t \tau'$ , que es igual a la fuerza de infección asumida por el modelo clásico bajo la hipótesis de mezcla homogénea con  $\alpha = k$  y  $\tau = \tau'$ . Esta interpretación es solo aproximada porque los dos modelos ven al contacto efectivo de manera diferente. En los modelos compartimentales, la identidad de un contacto se determina al azar y de forma instantánea para cada evento de transmisión; en los modelos de red, por otra parte, la identidades de los contactos se eligen al azar, pero permanecen fijas para la duración del

período infeccioso de un individuo. Así, los modelos conllevan diferentes valores para el ratio reproductivo básico y los valores de equilibrio esperados (Keeling, 2000). Algunos autores han argumentado que la hipótesis de mezcla homogénea se traduce en una red en la que cada individuo está conectado a todos los demás individuos en la población (una red completa) (Aparicio & Pascual, 2007). De hecho, una red completa es sólo una red aleatoria regular en el que cada individuo de la red esta conectado con los otros  $N - 1$  individuos. La fuerza de infección en una red completa es  $\lambda \approx (N - 1)\iota_t\tau'$ . Para la mayoría de las poblaciones, sin embargo, no es realista imaginar que las personas están en contacto constante con todos los demás individuos. Por lo tanto, es más realista para interpretar los arcos de la red como contactos *posibles* y el término de transmisión  $\tau'$  como una tasa de contacto y la transmisión efectiva combinada.

### **Modelos con estructura etaria**

Hasta este punto, hemos considerado enfermedades en las cuales el período de desarrollo de las mismas es despreciable cuando se lo compara con el desarrollo de los individuos. Sin embargo, existen casos (mayormente en la ecología) en donde el desarrollo de las enfermedades es comparable al período de desarrollo de los individuos. Ejemplo de estas situaciones son el modelo de Cao y Zhou (2012), donde se estudia la importancia que juega la estructura etaria de la población en la transmisión de la tuberculosis debido a la permanencia del patógeno en las poblaciones afectadas. El modelo es implementado a través de un sistemas de ecuaciones en diferencias y se realizan para diferentes análisis con el propósito de calcular  $R_0$  en los diferentes grupos que componen la población. En el trabajo publicado por Yang, Qiu, y Li (2014), los autores formulan un modelo estructurando la población según diferentes rangos de edad para describir la dinámica de transmisión del cólera. El modelo está basado en ecuaciones en derivadas parciales que modelan mecanismos de contagio directos e indirectos, infectividad dependiente de la edad y períodos de infectividad de la población.

Los ejemplos antes mencionados constituyen un argumento que pone de manifiesto las limitaciones que tienen los modelos basados en ecuaciones diferenciales ordinarias, que como se mencionó con anterioridad, si bien simplifican el problema de modelado, dejan afuera muchas variables que pueden resultar importantes a la hora de explicar la dinámica de una enfermedad, por otro lado no resultan útiles para modelar el problema ya que no pueden reflejar estos períodos de desarrollo de los individuos y los efectos de la enfermedad en dichos estados. Una alternativa es utilizar *ecuaciones diferenciales ordinarias con retraso de tiempo* (DDE). Las DDE son un tipo de ecuación diferencial en las cuales la derivada de una cierta función en un instante de tiempo  $t$  está dada en términos del valor de dicha función en un instante de tiempo  $t - \tau$ . Este tipo de modelos permiten explorar las propiedades de la estructuración etaria de una población de manera mas simple que los sistemas basados en derivadas parciales. En el trabajo de López, Izquierdo, Manzoli, Beldoménico, y Giovanini (2016), los autores introducen un modelo basado en DDE con

el propósito de modelar un proceso de miasis en una población de aves silvestres. La separación de la población parasitada en tres grupos etarios prueba que el daño producido por las larvas de la mosca depende del estado de desarrollo del pichón.

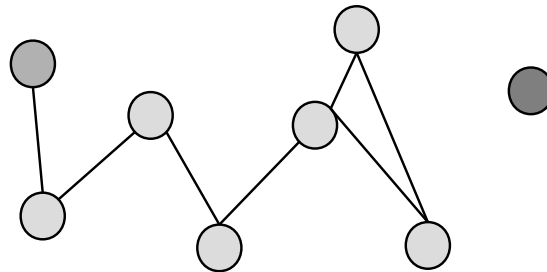
### 1.2.2. Modelos basados en individuos

Los modelos basados en individuos espacialmente implícitos representan un marco de modelado significativamente diferente al propuesto por los modelos clásicos, aunque ambos fueron ideados por la misma comunidad de investigadores. El ejemplo más común de este tipo de este método es el modelo discreto de transmisión individual desarrollado hacia finales de la década 1990 (Ghani, Swinton, & Garnett, 1997; Keeling, 1999; Koopman & Lynch, 1999). Los conceptos básicos que subyacen en este tipo de modelos son dos.

1. Los individuos son diferentes unos de otros, este concepto básico debería ser una hipótesis básica en estudios epidemiológicos.
2. Las interacciones e infecciones individualizadas juegan un rol vital en la dispersión de la enfermedad a través de toda la población (Cordero y col., 2004).

El diseño básico de estos modelos hace énfasis en individuos únicos y las interacciones entre ellos. El uso de individuo como unidad de modelado hace a estos modelos similares a los autómatas celulares, pero los mismos se desarrollan independientemente. Cada individuo interactúa con un número limitado de individuos conformando una red social (ver Figura 1-9), los individuos con los que interactúa no son necesariamente espacialmente adyacentes. Estas interacciones individualizadas finitas a través de redes sociales proveen un mecanismo de transmisión de la enfermedad a través del tiempo y del espacio. Las dinámicas de la epidemia dependen de la estructura de la red social. Comúnmente se pueden incluir diferentes factores, como por ejemplo, el número de individuos con los que un individuo en particular interactúa, la frecuencia con la que lo hace, el tiempo que duran dichas interacciones y las relaciones que existen dentro de un grupo de individuos (Ghani y col., 1997). La estructura de las redes sociales determina de manera consecuente la forma en que las enfermedades se distribuyen en la población. Debido a que el enfoque inicial de este tipo de modelos se centra sobre las conexiones sociales, la dinámica espacial del proceso epidémico no siempre es contemplado de manera explícita.

Las interacciones finitas no solo rompen la hipótesis de mezcla homogénea, sino que indirectamente rompen la hipótesis de distribución homogénea, debido a que los individuos solo interactúan con un número limitado de individuos conformando una red social, limitando de esta forma el espacio de influencia de los mismos. Además, estos individuos pueden desplazarse a diferentes locaciones con el objetivo de interactuar con otros. La heterogeneidad expresada explícitamente en los individuos y sus interacciones hacen efectivo y necesario el uso de aproximaciones estocásticas para modelar la transmisión



**Figura 1-9.:** Modelo basado en individuos espacialmente implícito

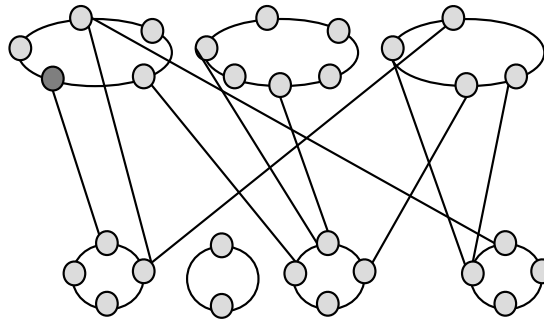
de la enfermedad. En otras palabras, el patrón de dispersión a nivel poblacional emerge de procesos dinámicos locales. La mayoría de estos modelos han sido desarrollados principalmente para modelar enfermedades de transmisión sexual y usualmente involucran un número pequeño de individuos (Kretzschmar & Morris, 1996; Ghani y col., 1997; Adams, Barth-Jones, Chick, & Koopman, 1998; Van der Ploeg y col., 1998; Welch, Chick, & Koopman, 1998; Koopman & Lynch, 1999). Aunque las consideraciones espaciales están incluidas de forma indirecta, este tipo de modelo posee un rol crítico en la evolución del diseño espacial de los modelos epidemiológicos por su inequívoco enfoque en los individuos como entidades únicas y comportamiento único (interacciones e infecciones). Estos modelos también sirvieron como piedra fundamental para los modelos basados en individuos móviles.

### **Modelos basados en individuos móviles**

Los modelos basados en individuos móviles surgieron como metodología para modelar fenómenos epidemiológicos hacia mediados de la pasada década (Bian, 2004; Dibble & Feldman, 2004; Eubank y col., 2004; Huang, Sun, Hsieh, & Lin, 2004; Dunham, 2005; Ferguson y col., 2005; Longini y col., 2005; Cooley y col., 2008). Este tipo de modelos representan una extensión espacial de los modelos basados en individuos espacialmente implícitos. Los mismos contemplan la heterogeneidad teniendo en cuenta los tres principios de diseños: *i)* los individuos como la unidad de modelado, *ii)* las interacciones entre ellos, *iii)* el proceso espacial y temporal. Para modelar la transmisión de la enfermedad estos modelos contemplan el concepto de geografía espacial y la teoría de redes (Keeling, 1999; Albert & Barabási, 2002) para representar la dinámica temporal y espacial.

Las trayectorias de los individuos en un período de tiempo (días, horas, etc) son representadas como trayectorias espacio-temporales. Estas trayectorias se intersecan en diferentes locaciones, como casas, lugares de trabajo y oficinas de servicios en donde los individuos interactúan entre ellos, conformando de esta forma una red social (ver Figura 1-10). La enfermedad se distribuye a través de la red por medio de dos mecanismos, las interacciones entre los individuos en un lugar y los traslados de los individuos entre locaciones para interactuar con otros individuos (Bian, 2004; Mikler, Venkatachalam, & Abbas, 2005).





**Figura 1-10.:** Modelo basado en individuos móviles

Debido a que la ubicación y el tiempo tanto de las interacciones y los movimientos entre ellos están representados de manera explícita, la dinámica espacial de transmisión de la enfermedad se puede modelar de forma explícita. Estos modelos pueden revelar un rango de heterogeneidades temporales y espaciales en enfermedades de transmisión, dependiendo de la estructura de la red social (Bian, 2004). Todas las hipótesis que subyacen en los modelos basados en individuos móviles son opuestas a aquellas realizadas por los modelos clásicos (incluyendo los modelos poblacionales basados en olas). Específicamente, los individuos son únicos, interactúan con un número finito de otros individuos, están heterogeneamente distribuidos, y se mueven. Este tipo de modelo utiliza de forma casi exclusiva un enfoque estocástico que surge naturalmente como consecuencia de las interacciones y transmisiones localizadas que de forma colectiva contribuyen al patrón poblacional de la dinámica espacial.

Dado de que la localización y la movilidad de los individuos está registrada de forma explícita, este tipo de modelos resultan ser una herramienta poderosa para estimar la velocidad temporal y la extensión espacial de una epidemia. El enfoque distintivo de los individuos únicos hace de los modelos basados en individuos móviles una aproximación aplicable a la hora de modelar vulnerabilidades relacionadas a la salud a nivel individual. El alcance de la población puede variar de cientos de personas en un lugar de trabajo, cientos de miles de personas en una comunidad, a millones de personas en una región (Eubank y col., 2004; Cooley y col., 2008; Mao & Bian, 2010) Aún en su infancia, los modelos basados en individuos móviles enfrentan muchos desafíos. La hipótesis individual, por ejemplo requiere de una gran cantidad de información detallada sobre los individuos que suele no estar disponible. EL estado de salud y otros datos a nivel individual son recolectados bajo directrices restringidas debido a cuestiones de privacidad. Para compensar la escasez de esta información, se puede utilizar otro tipo de información que sustituya y permita estimar los aspectos faltantes (atributos y comportamientos). En estudios más recientes, este tipo de modelos se han utilizado para investigar el comportamiento de conexiones individuales en medios de alta resolución temporal y espacial usando información de dispositivos inalámbricos y redes sociales.

Los hallazgos de estos estudios son invaluable para el diseño de modelos basados en individuos móviles. La disponibilidad de información en tiempo real y métodos para recolectar esta información agrega utilidad a este tipo de modelos. Por ejemplo, analizando el estado de salud de forma en tiempo real de los individuos durante una epidemia, la misma puede ser seguida por períodos de tiempo mas cortos que utilizando los canales convencionales con un nivel de exactitud razonable. A pesar de todos los retos, los modelos basados en individuos móviles poseen gran potencial para pronosticar la dinámica temporal y espacial de una epidemia en tiempo real mediante la incorporación de dicha información. La hipótesis de movilidad utilizada en los modelos aporta flexibilidad, aunque en diferentes grados, a la modelización de la heterogeneidad y a la dinámica de la transmisión de enfermedades. A diferencia de los tres primeros tipos de modelos en los que la dispersión de las enfermedades está representada por el cambio en el estado de salud de las unidades de modelado inmóviles, en los últimos tres modelos el cambio de ubicación de las unidades de modelado se realiza de forma explícita. A diferencia de modelar la propagación de la enfermedad en forma de ondas, el modelado de la propagación puede considerarse similar al que se utiliza para el movimiento de objetos discretos. Una sub-población o un individuo se pueden representar fácilmente como un objeto. Cuando el objeto se mueve, sus partes (por ejemplo, todos los individuos en una sub-población móvil), sus atributos (por ejemplo, el estado de salud), y sus comportamientos de interacción (por ejemplo, infección) y se mueven con él. Esta representación permite que la unidad de modelado se mueva libremente en el espacio, representando así la naturaleza dinámica de las fuentes de infección, las interacciones y la transmisión de enfermedades.

### **Modelos basados en autómatas celulares**

Los modelos basados en autómatas celulares pueden ser considerados una extensión de los modelos clásicos (Holmes, 1997). Fueron desarrollados hacia finales de la década de 1990. Al igual que los modelos sub-poblacionales, los modelos basados en autómatas celulares dividen el espacio en celdas que representan individuos discretos en lugar de una agregación de muchos individuos idénticos. Los principios de diseño de estos modelos son diferentes a los principios que gobiernan a los modelos clásicos (incluyendo a los modelos poblacionales basados en olas). Además de usar al individuo como unidad de modelado, las interacciones entre las unidades son representadas de forma explícita. Una célula interactúa con un número finito de células adyacentes. La transmisión espacial y temporal de la enfermedad es representado por medio de transmisiones localizadas que comienzan desde una célula y se expanden hacia células vecinas susceptibles. Este tipo de modelos también facilita las dispersiones de larga distancia estableciendo nuevos focos de transmisión localizados a cierta distancia de las células que ya están infectadas. En los modelos basados en autómatas celulares los individuos difieren unos de otros y cada uno de ellos tiene una localización explícita en la grilla. De esta forma, un individuo en particular solo interactúa con un número limitado de individuos cercanos.

Aunque las hipótesis de distribución homogénea e inmovilidad son también adoptadas, debido a las interacciones locales finitas entre las unidades de modelado, los modelos basados en autómatas celulares son una alternativa atractiva para modelar los aspectos que los modelos clásicos no contemplan. Como resultado, la dinámica del modelo puede diferir (Holmes, 1997). Para la representación espacial de la transmisión de una enfermedad, el modelo basado en autómatas celulares adopta una aproximación heurística basada en reglas. En lugar de utilizar una sola ecuación para determinar el comportamiento de toda la población o la sub-población, los autómatas celulares se enfocan en la transmisión a nivel local y como dichas interacciones contribuyen de forma colectiva al patrón de dispersión a nivel población.

Por la hipótesis de inmovilidad y la regla de transmisión entre individuos adyacentes, estos modelos son adecuados para modelar enfermedades que se transmiten entre unidades inmóviles, como plantas (Holmes, 1997). Desde los modelos basados en olas, los modelos de sub-poblaciones, a los individuos basados en autómatas, la unidad de modelado cambia de un segmento de población a un grupo de individuos idénticos a un grupo de individuos únicos, respectivamente. Pese a los cambios introducidos en el modelo basado en autómatas, las unidades de modelado permanecen inmóviles, al igual que en los modelos descritos anteriormente. El movimiento puede ser contemplado como el cambio de salud entre individuos adyacentes más que un cambio explícito de sus locaciones en la grilla. De esta forma la enfermedad se mueve por el espacio.

### **Modelos basados en agentes**

Los agentes son entidades que actúan en forma autónoma en un ambiente determinado, para poder actuar los agentes deben poder percibir dicho ambiente y todo lo que hay en él, por lo son capaces de interactuar entre sí; para ello los agentes poseen un conjunto de funciones que le permiten desenvolverse como tal. Al igual que los autómatas los agentes modifican su comportamiento de acuerdo al estado en el que se encuentren gracias a una estructura de datos interna que le permite almacenar cierto tipo de información y actuar en base a las percepciones actuales y a la historia del sistema. Los sistemas multi-agente permiten utilizar múltiples agentes unidos en una red para la resolución de un determinado problema. En el contexto de las simulaciones puede verse a un sistema multi-agente como un conjunto de agentes que interactúan en un ambiente determinado, permitiendo simular en forma directa a los individuos, sus comportamientos e interacciones. Al momento de simular una epidemia, puede verse a la población como a un conjunto de agentes que interactúan entre sí, implementando así un agente por individuo. Los ambientes pueden ser ciudades, pueblos, zonas rurales, todo depende de lo que se intente modelar. Dentro de estos ambientes los agentes se comportarán de determinada manera, dicho ambiente puede ser una grilla en donde los agentes se muevan libremente e interactúen con otros que se encuentre dentro de un radio de vecindad. Todos estos factores hacen a este tipo de modelado una buena alternativa a la hora de pensar en una forma

de simular la pandemia de gripe española, por otro lado, dicha técnica ya ha sido utilizada con anterioridad para modelar otro tipo de enfermedades en poblaciones humanas, lo cual nos brinda antecedentes lo suficientemente amplios a la hora de implementar un modelo propio. A lo largo de una epidemia la infectividad de los individuos depende de la carga viral, por otro lado, esta no permanece constante, varía a través del tiempo  $t$  dependiendo de diversos factores; la susceptibilidad es otra característica que no permanece constante a lo largo de la epidemia. Es posible modelar la respuesta inmune de cada individuo y la carga viral total con sistemas de ecuaciones al igual que en los modelos clásicos pero a nivel individuo; utilizando sistemas de ecuaciones de orden elevado es posible introducir retardos al modelo y de esta forma simular los períodos entre las diferentes etapas de la enfermedad. De esta forma se puede modelar el proceso global por medio de la interacción de estos individuos que son en si un sistema único.

### 1.3. Objetivos

Los objetivos propuestos para esta tesis son:

1. Desarrollar una metodología para modelar y simular el comportamiento de sistemas con múltiples poblaciones en un ambiente;
2. Estudiar el problema de heterogeneidad y comportamiento social. Considerando herramientas de control epidemiológico como vacunación y cuarentena;
3. Desarrollar herramientas matemático-computacionales para modelar y simular el comportamiento dinámico de los sistemas epidemiológicos;
4. Evaluar y validar el modelo propuesto utilizando un caso de estudio.

Estos aspectos tendrán impacto significativo en el estudio de los problemas epidemiológicos ya que hasta el momento estos aspectos no han sido abordados integralmente, permitiendo una mejor y más precisa descripción de los problemas y una mayor y mejor capacidad de análisis de los mismos.

### 1.4. Organización

La presente tesis doctoral está organizada de la siguiente manera. En este primer capítulo se han introducido algunos de los principios básicos de la epidemiología, hemos revisado las principales características de los modelos epidemiológicos más importantes de las últimas décadas en el área. En el capítulo 2 se describe el modelo basado en individuos implementado en esta tesis. En el capítulo 3 se hará una revisión del estado del arte

en cuanto al modelado basado en autómatas celulares y los modelos de redes de contactos. Se ejemplifica su uso mediante una breve revisión del estado del arte revisando los modelos epidemiológicos que abordan la problemática desde distintas perspectivas. En este capítulo se presenta uno de los enfoques que utilizamos, el cual permite una descripción de los sistemas epidemiológicos desde el punto de vista de los contactos entre los actores. En el capítulo 4 se introducen los conceptos elementales que constituyen el enfoque de modelado espacialmente explícito. En primer lugar se introduce un marco teórico apropiado que permite definir de forma clara el modelo que luego es presentado. A continuación se explican las principales características que incorpora, como auto organización y comportamiento social. Finalmente se presentan los resultados arrojados por el mismo. En el capítulo 5 se presentarán otros problemas de modelado que pueden ser encarados mediante los modelos aquí presentados. Entre estos problemas pueden mencionarse por ejemplo, desarrollo de epidemias cuando conviven múltiples cepas de un virus, la aplicación de políticas de control, como influye el comportamiento de los individuos en la dinámica epidémica, entre otros. En el capítulo 6 se presentan las conclusiones finales del trabajo y las líneas de trabajo futuras que se pretende abordar.

Al final de esta tesis se incorporan dos apéndices en donde se describen los algoritmos desarrollados. El apéndice A contiene el detalle de los algoritmos implementados en el modelo descrito en el capítulo 3. Finalmente el apéndice C describe las rutinas implementadas en el modelo implementado en el capítulo 4

## 1.5. Conclusiones

La epidemiología estudia los patrones de enfermedades en las poblaciones. La hipótesis central es que estos patrones pueden analizarse sistemáticamente para entender sus causas y controlar la enfermedad. El uso de variables epidemiológicas tales como edad o sexo como marcadores de fenómenos más complejos permite entender mejor estos patrones. Al tratar con poblaciones, la interacción social se transforma en un factor importante de la epidemiología. Así como se habla de *individuos enfermos o sanos*, en epidemiología se consideran *poblaciones enfermas o sanas*, y las causas de enfermedad difieren en cada caso. Información a nivel individual puede no ser relevante a nivel poblacional y viceversa.

Existen diversos modelos epidemiológicos basados en ODE. En estos modelos, el número de reproducción  $R_0$  determina la permanencia en forma endémica de la epidemia o su extinción. Este parámetro depende de la tasa de infección, la tasa de paso de expuesto a infeccioso, la tasa de recuperación y la tasa de mortandad, y determina un equilibrio endémico. Para que el modelo funcione deben cumplirse razonablemente las hipótesis de distribución uniforme e instantánea, así como la uniformidad de parámetros en todos los individuos. Esto puede cumplirse razonablemente bien en poblaciones grandes, pero como veremos en capítulos posteriores, no lo hace en poblaciones más pequeñas.

Los modelos basados en individuos representan un marco de modelado significativamente diferente del propuesto por los modelos clásicos. Los conceptos que subyacen en ellos son dos. Por un lado, los individuos son diferentes unos de otros, este concepto debería ser una hipótesis básica en estudios epidemiológicos. Por otro lado, las interacciones e infecciones individualizadas juegan un rol vital en la dispersión de la enfermedad a través de toda la población. El uso de individuo como unidad de modelado hace a los autómatas celulares y a los modelos basados en agentes una herramienta casi obligatoria para implementar esta lógica de modelado.

Sin embargo, los diferentes enfoques de modelado varían considerablemente y no existe un marco teórico unificado para la implementación de modelos basados en individuos. Existe un gran número de herramientas de software, plataformas y librerías que permiten implementar todo tipo de modelos, pero no ofrecen un marco unificado si mas bien un conjunto de marcos diferentes.



## 2. Modelo

### 2.1. Introducción

En la década de los noventa apareció un nuevo enfoque de modelado con una filosofía y perspectiva diferentes a la modelización clásica utilizada hasta entonces, la llamada *modelización basada en el individuo* (IBM). Los IBM son modelos en los que los individuos que forman el sistema son tratados como entidades autónomas y discretas (Grimm & Railsback, 2013). El esfuerzo de modelado entonces se centra en la caracterización de estas partes discretas mediante reglas de comportamiento, que permiten que interaccionen entre ellas y con el entorno en el que se hallan. Se tratan de modelos computacionales que requieren de simplificaciones distintas a las realizadas en los modelos clásicos, y son una muy buena alternativa de modelización para abordar el estudio de determinados aspectos relativos a los sistemas complejos. La mayoría de los ejemplos de utilización de estos modelos se encuentran en el área de ecología, en distintos campos como el estudio de poblaciones de insectos, peces, mamíferos, aves, plantas o árboles, entre otras opciones (Gornick & Meyers, 2003; Epstein, 2006; Grimm & Railsback, 2013). En la actualidad, los IBM ya se han consolidado como una herramienta válida para describir y tratar con sistemas complejos formados por entidades autónomas. Al mismo tiempo, este tipo de modelo computacional facilita la experimentación sobre los sistemas modelizados, mediante la observación y análisis del conjunto de simulaciones que resultan de su implementación en un entorno informático adecuado.

El presente capítulo se organiza de la siguiente manera. En la sección *estado del arte* se realiza una revisión bibliográfica sobre los trabajos más importantes de los últimos años en el modelado basado en individuos, se hace un repaso de las principales herramientas introducidas en el campo, analizando sus principales ventajas y desventajas. Posteriormente introducimos un nuevo marco de trabajo para los modelos basados en individuo. En este capítulo se describen las principales características del modelo y sus ventajas con respecto a los enfoques ya existentes. Finalmente en el final de capítulo realizaremos una breve recapitulación e introduciremos las configuraciones que adopta el modelo propuesto en los capítulos 3, 4 y 5.



## 2.2. Estado del arte

Los modelos basados en individuos simulan poblaciones y comunidades teniendo en cuenta a los individuos y sus propiedades. Se han utilizado en el modelado ecológico por más de cuatro décadas y su uso y ubicuidad en el campo de la epidemiología ha crecido rápidamente en las últimas dos décadas. Los modelos basados en individuos se han utilizado en muchos problemas aplicados, tales como proteccionismo y manejo de poblaciones particulares en lugares específicos, pero su uso para abordar problemas teóricos también se ha incrementado rápidamente, ayudándonos a entender cómo los conjuntos de rasgos de organismos individuales influyen en el ensamblaje de comunidades y redes alimentarias. Los modelos basados en individuos desempeñan un papel cada vez más importante en las preguntas que surgen del estudio de sistemas biológicos complejos.

Los modelos basados en individuos, representan una forma de modelar poblaciones y comunidades, permitiendo un alto grado de complejidad que surge como producto de usar individuos como unidad básica de modelado y las interacciones entre ellos (Keeling, 1999; Koopman & Lynch, 1999). Los modelos basados en individuos simulan poblaciones o sistemas de poblaciones como organismos compuestos de individuos discretos. Cada individuo tiene un conjunto de estados o atributos y comportamientos. Las variables de estado pueden incluir localización espacial, rasgos fisiológicos y conductuales. Estos atributos varían entre los individuos y pueden cambiar con el tiempo. Los comportamientos pueden incluir crecimiento, reproducción, selección de hábitat y dispersión. A diferencia de los modelos poblacionales basados en ecuaciones diferenciales, que se describen en términos de parámetros de la población (como las tasas de natalidad y mortalidad), los modelos basados en individuos son modelos ascendentes en los que los comportamientos poblacionales emergen de las interacciones entre individuos y su entorno. Una ventaja de los modelos basados en individuos sobre los modelos tradicionales es que pueden incorporar cualquier número de mecanismos a nivel individual. Estos modelos han sido exitosamente utilizados para modelar entre otras cosas enfermedades de transmisión sexual o tuberculosis (Adams y col., 1998; Ghani y col., 1997; Koopman & Lynch, 1999; Kretzschmar & Morris, 1996; Van der Ploeg y col., 1998; Welch y col., 1998).

El primer uso importante que se le dio a los modelos basados en individuos vino por parte del modelo *JABOWA* de Botkin (1993). *JABOWA* es modelo de *sustitución de fase de hueco*, que describía de que manera se formaban comunidades nuevas de árboles en porciones de terreno que fueron previamente ocupadas por comunidades previas (típicamente de  $1000m^2$ ), teniendo en cuenta los efectos del sombreado en árboles de diferentes especies y alturas en relación con sus vecinos. Una característica importante de *JABOWA* y de todos los *modelos de hueco* posteriores es que un árbol se caracteriza por una sola variable de estado, el diámetro de su tallo. Todas las demás variables del individuo, como la altura y el perfil vertical de las hojas, entre otras, se deducen del diámetro del tallo a través de relaciones alométricas. Diferentes especies se representan a través de diferentes

parámetros de crecimiento y alométricos, no a través de diferentes estructuras de modelo. Modelos forestales posteriores (Worm & Myers, 2003), ampliaron el espacio modelado más allá de  $1000m^2$  incluyendo las ubicaciones espaciales de cada árbol. Por otra parte añaden mayor detalle fisiológico y visualizaciones más sofisticadas.

La extensión de los modelos basados en individuos a otras áreas de la ecología fue inicialmente lenta. Un área pionera en el uso de modelos basados en individuos fue la biología marina en la cual se implementaron estos modelos para el relevamiento en poblaciones de peces y de esta forma evaluar el impacto que los seres humanos tienen en la mortalidad de los peces jóvenes. Esto requiere de una comprensión de las complejas interacciones que surgen del tamaño, la velocidad y la vulnerabilidad a la depredación de bancos de peces, lo cual podría ser mejor abordado modelando las interacciones de peces individuales (Ellemers, Wilke, & Van Knippenberg, 1993). Los efectos de la variabilidad en el ambiente físico (flujo, temperatura, salinidad, turbulencia y luz) se consideran explícitamente en el estudio de la dispersión, el crecimiento y la mortalidad de los individuos. No es sorprendente que los modelos basados en individuos se apliquen a muchas otras poblaciones de animales más allá del ámbito de la biología marítima. Son adecuados para incorporar paisajes espaciales de manera realista y simular comportamientos complejos en las poblaciones. Estos modelos son ideales para especies que resultan de particular interés, para las cuales a menudo hay información empírica detallada. Modelos espacialmente explícitos de este tipo incluyen pastoreo (Turner, Wu, Romme, & Wallace, 1993), colonias de aves (Wolff, 1994), estudio del comportamiento de osos pardos (Wiegand, Naves, Stephan, & Fernandez, 1998) y estudio de diferentes poblaciones de aves (Mooij, Bennetts, Kitchens, & DeAngelis, 2002).

Recientemente, muchos sistemas de simulación basados en agentes se han desarrollado para analizar la difusión epidémica en comunidades o ciudades, entre los cuales podemos mencionar: *MASON* (Luke, Cioffi-Revilla, Panait, Sullivan, & Balan, 2005), *Geograph* (Dibble & Feldman, 2004), *EpiSims* (Mniszewski, Del Valle, Stroud, Riese, & Sydoriak, 2008), y *FluTE* (Chao, Halloran, Obenchain, & Longini Jr, 2010).

*MASON* (Luke y col., 2005) provee un conjunto de librerías en JAVA para simulación basada en agentes. Es compatible con el desarrollo de sistemas de simulación de epidemias a gran escala. Está constituido por un núcleo de simulación de eventos discretos y una librería de visualización escrita en Java, diseñada para ser lo suficientemente flexible como para ser utilizada para una amplia gama de simulaciones simples, pero con un énfasis especial en simulaciones de enjambre. *MASON* fue construido pensando en librerías de fácil integración, de forma tal que un programador experimentado de Java pueda agregar fácilmente características a los modelos diseñados, en lugar de usar un número reducido de librerías con muchas características relacionadas con el dominio, entrelazadas que son difíciles de eliminar o modificar.

*Geograph* brinda un conjunto de herramientas computacionales diseñados para realizar experimentos de sistemas geográficos. En Dibble y Feldman (2004), Pires y Crooks (2017)

se describe el desarrollo de un conjunto de herramientas en *Geograph* que se pueden utilizar para explorar el comportamiento espacialmente explícito de cualquier modelo epidémico *SEIR* por medio agentes móviles heterogéneos espacialmente. Los modelos implementados con esta herramienta tienen las siguientes características:

- Permite formalizar un enfoque basado en agentes que integra sistemas de información geográfica y análisis de redes sociales para modelar fenómenos sociales;
- El ambiente físico y los atributos de los agentes son desarrollados usando datos empíricos;
- El marco cognitivo del agente permite modelar sus actividades en el espacio físico y las interacciones en el espacio social; y
- A través de la integración de las tres técnicas, se puede captar la compleja naturaleza no lineal de diferentes prácticas sociales.

*EpiSims* es un sistema de simulación de epidemias que está desarrollado en base al sistema de simulación de transporte urbano TRANSIMS. En *EpiSims*, la teoría de grafos se aplica a las redes de tráfico que modelan los patrones de movilidad humana con el fin de analizar el patrón de transmisión de enfermedades infecciosas. *EpiSims* está diseñado como un simulador de eventos discretos distribuidos. Los modelos se implementan como un código de paso de mensajes basado en MPI para su uso en una arquitectura de clúster donde es posible escalar el número de procesadores utilizados con el tamaño de problema. La mayor parte del cálculo en *EpiSims* se asocia con dos tipos de objetos: personas y ubicaciones. La gente se mueve de un lugar a otro, interactuando con su ubicación actual y cualquier otra persona presente en ese lugar. *EpiSims* implementa esto manteniendo contenedores para objetos persona en cada objeto de ubicación. La transmisión de la enfermedad es mediada a través de subgrupos de personas en una localización específica. Este sistema ha sido utilizado de forma exitosa para modelar epidemias de gripe (Halleran y col., 2008), viruela (Barrett, Eubank, & Smith, 2005) y brotes epidémicos en entornos urbanos (Eubank y col., 2004).

*Biowar* es un sistema basado simulación de brotes epidémicos y ataques biológicos en ciudades (Carley y col., 2006). Este sistema integra un conjunto de modelos computacionales de redes sociales, medios de comunicación, transmisión de enfermedades, espacialidad urbana y poblacionales. *Biowar* está diseñado como un sistema modular. Un componente, o módulo, en el sistema interactúa con otros módulos a través de una interfaz métodos simple. Un módulo corresponde conceptualmente a una capacidad de simulación *Biowar*, como la progresión de la enfermedad y el diagnóstico (el *módulo de la enfermedad*). Esta correspondencia no es necesariamente uno a uno. Esta flexibilidad hace posible un rápido desarrollo de modelos de simulación más fiables y más realistas.

*FluTE* es un modelo basado en individuos capaz de simular la propagación de una epidemia de gripe en las principales áreas metropolitanas de los Estados Unidos continental.

Construye el modelo a partir de datos provenientes de censos. En cada comunidad, la población se organiza como un conjunto de hogares organizados en términos de la distribución del tamaño de la familia que lo habita. Estos sistemas basados en agentes a gran escala por lo general dependen de la disponibilidad de datos. Mientras tanto, las técnicas que se utilizan en estos sistemas para visualizar brotes epidémicos en espacios geográficos requieren una alta potencia de cálculo.

Hasta aquí se han descrito y se han repasado numerosas herramientas de modelado que han sido propuestas por otros autores. Estos trabajos brindan diferentes enfoques para resolver una amplia gama de problemas, pero se encuentran restringidos a sistemas que cumplen las propiedades para las cuales fueron diseñadas. Por otro lado, son utilidades, herramientas de software o librerías que reúnen un amplio número de modelos, pero no representan un marco teórico o una nueva forma de concebir el modelado basado en individuos. En la siguiente sección introduciremos un nuevo marco de trabajo para los modelos basados en individuos

## 2.3. Modelo propuesto

En esta sección proponemos un marco conceptual para el desarrollo de los modelos basados en individuos. El individuo constituye la unidad básica de modelado, es decir, que la dinámica global del sistema va a surgir de la agregación de las dinámicas individuales de cada una de estas unidades. El comportamiento de los individuos es un factor determinante para comprender la dinámica del mismo, el comportamiento es lo que determina como se desenvuelva la unidad en el entorno y por lo tanto, define su dinámica. El comportamiento o conducta es la manera en que se comportan un individuo o un conjunto de individuos, en relación con su entorno y los estímulos que reciben y perciben del mismo. existen diferentes formas de definir el comportamiento y la conducta de los individuos teniendo en cuenta diferentes factores según la ciencia que lo investigue.

En el marco propuesto en esta tesis el comportamiento de los individuos se define como la combinación de tres unidades mas simples que definen la dinámica del comportamiento global del individuo. En la Figura 2-1 puede observarse un diagrama del enfoque propuesto. En la parte superior izquierda puede observarse una fracción de la grilla con individuos, si nos centramos en uno de estos elementos puede observarse que el comportamiento está conformado por tres módulos (parte inferior izquierda). De esta forma, la respuesta de los individuos se construye de forma modular y cada uno de esos módulos pueden ser modelados de la forma mas conveniente. El *comportamiento social* es el comportamiento o conducta dirigido hacia la sociedad que tiene lugar entre miembros del mismo grupo, viene dado por relaciones intraespecíficas. Muchos comportamientos sociales intraespecíficos son parte de la comunicación y provocan una respuesta, o cambio de comportamiento del receptor, sin actuar directamente sobre él. La comunicación

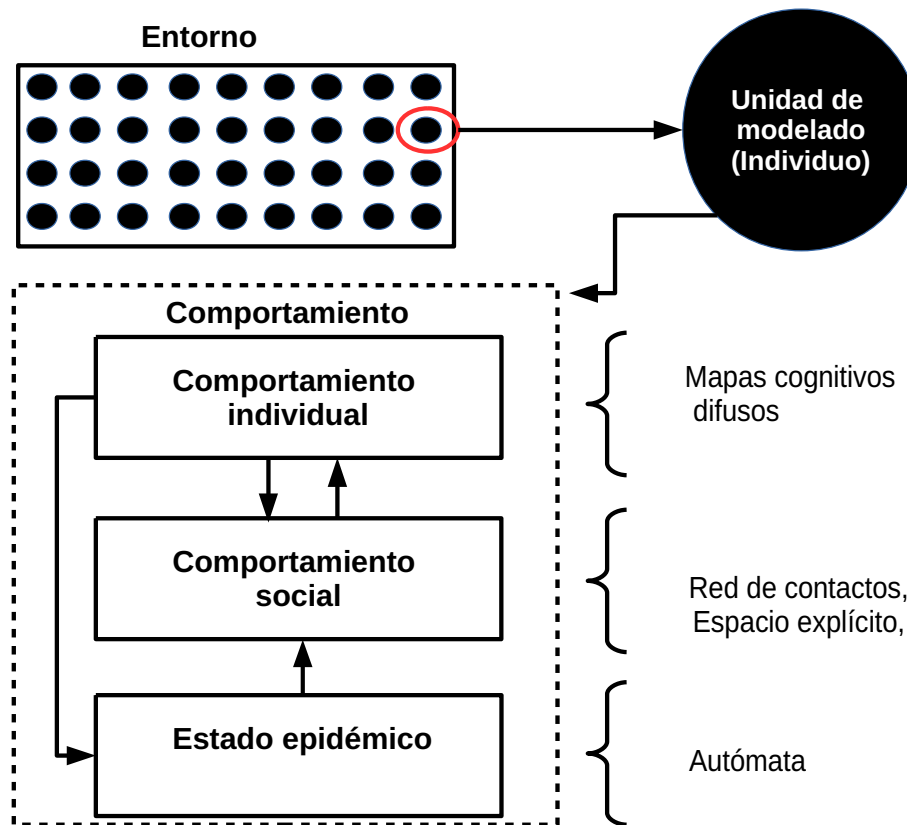


Figura 2-1.: Diagrama esquemático del marco de trabajo propuesto

entre miembros de diferentes grupos no se considera comportamiento social. La forma más original del comportamiento social humano viene dada por el lenguaje humano pero también por diferentes prácticas sociales.

El comportamiento social establece de que forma un individuo se relaciona con otros individuos. Está relacionado con el comportamiento individual. El comportamiento individual determina en parte el comportamiento social y por otro lado el comportamiento social define ciertos aspectos del comportamiento individual en una relación de retroalimentación.

El módulo correspondiente al comportamiento social puede ser modelado de diferentes formas. Por un lado puede utilizarse un enfoque basado en *redes de contacto* para modelar epidemias en donde la topología del ambiente no es relevante pero si lo son las relaciones intrapersonales establecidas entre los individuos que conforman la población, sería natural en este caso optar dicho enfoque. Por otro lado el comportamiento social puede ser modelado a través de un modelo de *espacio explícito* si se intenta modelar un fenómeno en donde las características del espacio físico juegan un rol preponderante, por ejemplo en fenómenos parasitarios en donde el espacio y la topología del ambiente ya que determinan las concentraciones de huéspedes y parásitos de acuerdo a la disponibilidad

de recursos y geografía del ambiente. Incorporar al modelo un bloque que contemple al espacio de forma explícita parece un enfoque más natural.

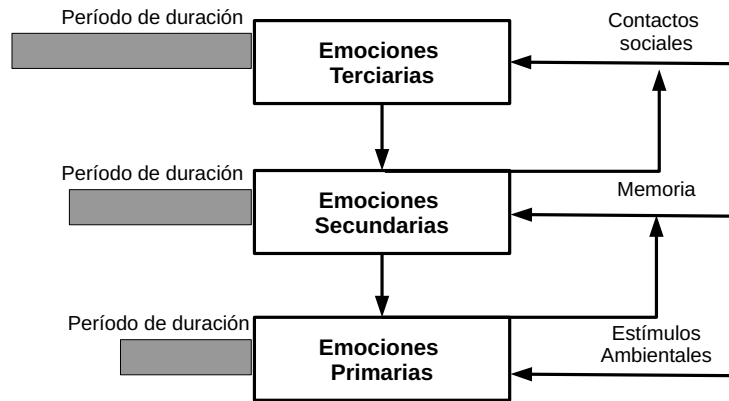
El *estado epidémico* es lo que determina de qué forma se modela el contagio del patógeno. Para modelar este aspecto se puede optar por incorporar estados epidemiológicos y funciones deterministas o estocásticas que determinen la transición entre los diferentes estados, o bien se pueden usar enfoques más realistas e incorporar modelos de respuesta inmune. En el análisis epidemiológico se consideran la estructura, las características epidemiológicas y el estado epidémico de una enfermedad. La estructura epidemiológica es la forma en que interactúan los factores inherentes al agente causal, ya sea físico, químico o biológico, el medio ambiente y el huésped en una población y en un período de tiempo determinado. Las características epidemiológicas son el resultado de la estructura epidemiológica y se expresan por la frecuencia y distribución de la enfermedad en la población. De acuerdo a las características epidemiológicas, la estructura epidemiológica es dinámica, modificándose según el comportamiento de la enfermedad en la comunidad.

El *comportamiento individual* se puede modelar aplicando un conjunto de reglas a un conjunto de datos de entrada, a través de un autómata de estados finitos o bien a través de algoritmos que simulen diferentes conductas a nivel individual. También se pueden utilizar herramientas de la lógica difusa que mapeen los datos de entrada para determinar la salida con el objetivo de modelar procesos causales como los que describen el comportamiento humano, o bien a través de un procedimiento mixto que emplee ambas técnicas. El modelo utilizado en esta tesis se basa en un enfoque mixto en donde se utilizan modelos basados en lógica difusa y un conjunto de reglas que determinan como se mueven los individuos ante determinadas situaciones epidemiológicas.

### 2.3.1. Comportamiento individual

El *comportamiento individual* está fuertemente determinado por las emociones del individuo. Las emociones humanas pueden clasificarse psicológicamente en tres niveles: *i*) emociones primarias, *ii*) secundarias y *iii*) terciarias (Xiao-Juan & Wei-Ren, 2007). Se consideraba las emociones primarias como respuestas intrínsecas de un individuo a estímulos externos. Las emociones secundarias se activan cuando las emociones primarias se conectan con la percepción actual y pasada; Las emociones terciarias se producen por el curso de los contactos sociales a largo plazo en un ambiente dado. En la Figura 2-2 se muestra un diagrama de los diferentes niveles que conforman al sistema emocional de un individuo y la relación entre dichos niveles (Damasio, 1994).

Las emociones instantáneas son intensas y de corta duración, debido a causas específicas y la naturaleza de la percepción, y se transforman fácilmente en otros estados. Por lo tanto, las emociones instantáneas son emociones primarias. Los estados de ánimo son más leves, difusos y duraderos, los cuales no tienen contenido específico de percepción. Los estados de ánimo aparecen y desaparecen lentamente sin causas obvias, y duran poco.



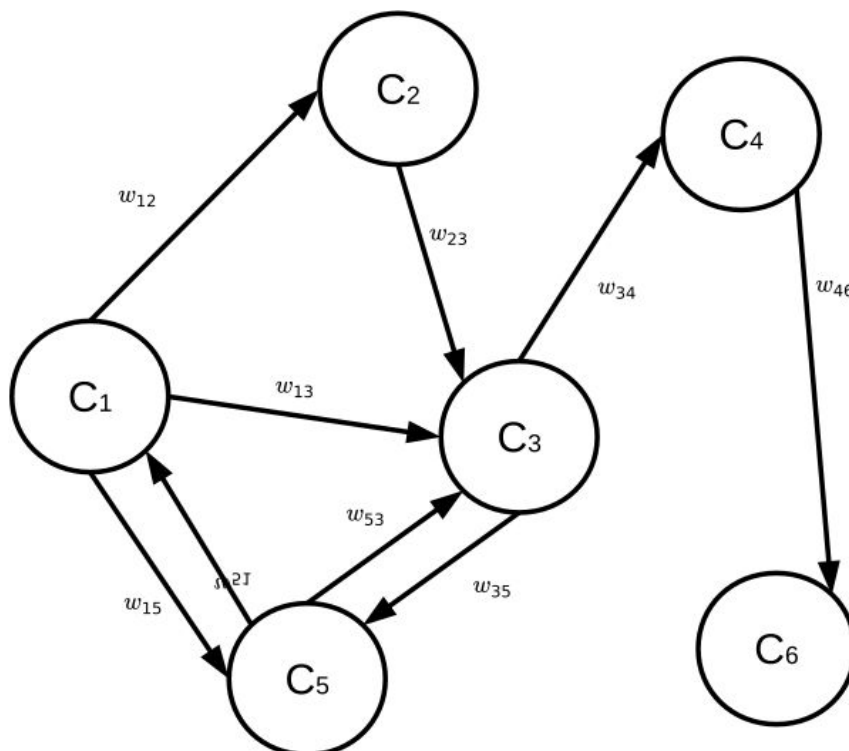
**Figura 2-2.:** Emociones y sus relaciones a diferentes niveles.

De esta forma, los estados de ánimo son emociones secundarias. La competencia de la sociedad es relevante en cada entorno social y cultural, que constituye más potencial y contexto emocional subyacente para la cognición y el comportamiento de las personas. La aparición y desaparición de la competencia social es más lenta cuando se la compara con los estados de ánimo, lo que sugiere que la competencia social es una emoción superior. Las emociones primarias están parcialmente dominadas por emociones terciarias, aunque están principalmente influenciadas por estímulos ambientales y por la auto percepción. La dominación se vuelve aún más evidente cuando las emociones primarias entran en conflicto con las emociones terciarias. Por ejemplo, una persona optimista (con respecto a las emociones terciarias) tiene emociones instantáneas menos intensas (con respecto a las emociones primarias) cuando se enfrenta a frustraciones (con respecto a estímulos ambientales). Además, existen ajustes entre emociones al mismo nivel, que llevan las emociones a cambiar hacia una dirección ventajosa. (Picard & Picard, 1997)

En la siguiente sección se introduce un marco teórico que permite modelar este aspecto del comportamiento de los individuos utilizando herramientas que combina elementos de la lógica difusa y de la psicología cognitiva. Los *Mapas Cognitivos Difusos* son estructuras de grafos que se utilizan para representar razonamiento causal. La estructura que brindan estos grafos permite la propagación sistemática causa, particularmente el avance hacia atrás y hacia adelante.

Los mapas cognitivos difusos o *FCM* (Fuzzy Cognitive Maps) proveen de una metodología para la representación y modelado de sistemas complejos. Describen diferentes aspectos del comportamiento de estos sistemas en términos de conceptos; cada concepto representa un estado o una característica del sistema y estos conceptos interactúan entre sí mostrando las dinámicas del sistema. Brindan una ilustración del sistema completo mediante un grafo mostrando la causa y efecto entre los conceptos y son una forma simple de describir el comportamiento del sistema de manera simbólica, explotando todo el conocimiento generado del sistema. Un mapa cognitivo integra experiencia y conoci-

miento acumulado sobre la forma en la que opera un sistema determinado, como resultado del método por el cual el mismo fue generado, esto es: usando expertos humanos que conocen cómo funciona el sistema bajo estudio junto con todos sus componentes bajo diferentes circunstancias. Lo que es más, las técnicas de entrenamiento de un FCM, las cuales implementan técnicas de la teoría de redes neuronales, permiten seleccionar un conjunto de pesos adecuados para las conexiones entre los diferentes conceptos dentro del mapa. Los FCM son grafos signados retro alimentados. Consisten de nodos conceptos  $C_i$  e interconexiones  $e_{ij}$  entre conceptos  $C_i$  y  $C_j$ . Un mapa modela la dinámica de un sistema complejo como una colección de conceptos y relaciones causa-efecto entre dichos conceptos (Figura 2-3).



**Figura 2-3.:** Diagrama de un FCM de 6 conceptos

Las interconexiones  $e_{ij}$  entre los diferentes conceptos están caracterizadas por un peso  $w_{ij}$  que describe el grado de causalidad entre dos conceptos. Los pesos pueden tomar valores en el intervalo cerrado  $[-1, 1]$ . El signo del peso indica el tipo de causalidad indica causalidad positiva  $w_{ij} > 0$  entre el concepto  $C_i$  y  $C_j$ , lo cual significa que a medida que se incrementa el concepto  $C_i$  también lo hará el concepto  $C_j$  y un decremento en el concepto  $C_i$  causará un decremento en el concepto  $C_j$ . Cuando existe una causalidad negativa entre dos conceptos, entonces  $w_{ij} < 0$ , un incremento en el concepto  $C_i$  causará un decremento en el concepto  $C_j$  y un decremento en el concepto  $C_i$  causará un incremento en  $C_j$ . Cuando no existe relación entre dos conceptos, entonces  $w_{ij} = 0$ . El valor de un peso  $w_{ij}$  indica



el grado de influencia que existe entre los conceptos  $C_i$  y  $C_j$ . Generalmente, el valor de cada concepto es calculado, calculando la influencia que ejercen otros conceptos por sobre el concepto específico, esto se puede hacer aplicando la fórmula de la ecuación que se muestra a continuación:

$$x_i(t) = f\left(\sum_{j=1, j \neq i}^n x_j(t-1)w_{ji}\right) \quad (2-1)$$

Donde  $x_i(t)$  es el valor del concepto  $C_i$  en el tiempo  $t$ ,  $x_j(t-1)$  es el valor del concepto  $C_j$  en el tiempo  $t-1$ ,  $w_{ji}$  es el peso de interconexión entre los conceptos  $C_i$  y  $C_j$  y  $f$  es una función de umbral que puede ser por ejemplo la función sigmoidea o el arco tangente por ejemplo. Esta función tiene como objetivo convertir el resultado de la ecuación (2-1) al intervalo difuso  $[0, 1]$  o  $[-1, 1]$ . En cada instante de tiempo, los valores del FCM cambian y se recalculan de acuerdo a la ecuación (2-1). La regla de cálculo para cada paso de la simulación incluye el cálculo de cada uno de los conceptos.

Otra regla para calcular el valor de los conceptos consiste en considerar el valor previo de cada concepto. En este caso el FCM tiene lo que se llama capacidad de memoria en un instante de tiempo, el último valor de cada concepto está involucrado en la determinación del nuevo valor del concepto actual, de esta manera los valores de los conceptos tendrán una leve varianza durante cada paso de la simulación. La fórmula que se propone para este método puede verse en la ecuación (2-2), donde  $x_i(t)$  es el valor del concepto  $C_i$  en el tiempo  $t$ ,  $x_i(t-1)$  es el valor del concepto  $C_i$  en el tiempo  $t-1$ ,  $x_j(t-1)$  es el valor del concepto  $C_j$  en el tiempo  $t-1$ ,  $w_{ji}$  es el peso de interconexión de  $C_j$  a  $C_i$  y  $f$  es la función de umbral. El parámetro  $k_2$  representa la proporción de contribución del valor previo del concepto en el cálculo del nuevo concepto y  $k_1$  expresa la influencia de los conceptos interconectados en la configuración del nuevo valor del concepto  $x_i$ . Los dos parámetros  $k_1$  y  $k_2$  satisfacen  $0 < k_1, k_2 < 1$

$$x_i(t) = f\left(k_1 \sum_{j=1, j \neq i}^n x_j(t-1)w_{ji} + k_2 x_i(t-1)\right) \quad (2-2)$$

### 2.3.2. Comportamiento de individuos en epidemias

La toma de decisiones en humanos está influenciada tanto por las emociones y la cognición. Las emociones juegan un papel importante e indispensable en que afecta a la memoria, la atención y el razonamiento. La inteligencia humana permite no sólo el pensamiento racional y el razonamiento lógico, sino que también tiene fuertes capacidades emocionales. Psicológicamente hablando, las emociones representan la reacción de un individuo para y la evaluación de sus propios estados internos y las interacciones entre ellos y su medio ambiente. Las emociones humanas pueden ser psicológicamente clasificadas en

**Tabla 2-1.:** Cuadro de conceptos del FCM

Concepto	Significado
$C_1$	Nro. de infectados que interactúan con el agente
$C_2$	Nro. de recuperados que interactúan con el agente
$C_3$	Estado del sistema inmune del agente.
$C_4$	Conocimiento de la situación epidemiológica local
$C_5$	Conocimiento de la situación epidemiológica global
$C_6$	Conocimiento de la situación epidemiológica local y global
$C_7$	Optimismo del agente.
$C_8$	Memoria de emociones secundarias.
$C_9$	Emociones primarias.
$C_{10}$	Decisión del agente.

tres niveles: las emociones primarias, secundarias, las emociones y las emociones de alto nivel. Se considera emociones primarias como a las respuestas intrínsecas de un ser humano a estímulos externos; las emociones secundarias se activan cuando las emociones primarias se conectan con la percepción actual y el pasado; las emociones de alto nivel surgen como resultado de los contactos sociales a largo plazo en un entorno determinado. El mapa utilizado en esta tesis esta basado en el modelo de Mei y col. (2014). En este modelo los conceptos  $C_i$  se dividen en tres grupos: *i*) conceptos de entrada, que representan las percepciones que el agente recibe del entorno (emociones primarias); *ii*) conceptos internos, que representan las emociones y las sensaciones del individuo (emociones secundarias) y *iii*) conceptos de salida (emociones terciarias), que representa las acciones que toma el agente como resultado de un proceso de decisión. En la Tabla (2-1) se detalla el significado de cada concepto.

Los conceptos  $C_1$ ,  $C_2$  y  $C_5$  son conceptos de entrada, es decir que si valor viene dado por las percepciones que el agente recibe del entorno.  $C_1$  y  $C_2$  pueden tomar valores dentro del intervalo cerrado  $[0, 1]$ , mientras que  $C_5$ , corresponde al nivel de alerta que emiten las autoridades de salud. En nuestro caso suponemos que los niveles de alerta son emitidos por el ministerio de salud y que dichos niveles son los mismos utilizados por la Organización Mundial de la salud. Como se modela un brote epidemiológico local  $C_5 = 0,5$  durante toda la simulación. Los conceptos  $C_3$ ,  $C_4$ ,  $C_6$ ,  $C_7$ ,  $C_8$  y  $C_9$  son conceptos internos de cada agente, es decir, los mismos no se ven influenciados de manera directa por el entorno. Finalmente  $C_{10}$  el concepto de salida y representa la percepción de la situación epidemiológica global por parte del individuo. En la Figura 2-4 se puede ver el esquema del mapa utilizado. El mapa fue entrenado siguiendo el algoritmo detallado en Mei y col., 2014. En el mismo se tienen en cuenta los valores esperados para el concepto de salida del mapa ( $DOC_{10}$ ), suponiendo que  $DOC_{10} \in [DOC_{10}^{min}, DOC_{10}^{max}] \subseteq [0, 1]$ . La de-

terminación de la matriz  $W$  esta sujeta la relación de causalidad entre conceptos del mapa. En este caso, se establece que la relación entre conceptos puede ser *negativa muy fuerte* cuando  $w_{ij} \in [-1, -0,75]$ , *negativa fuerte* cuando  $w_{ij} \in [-1, -0,5]$ , *negativa media* cuando  $w_{ij} \in [-0,75, -0,25]$ , *negativa débil* cuando  $w_{ij} \in [-0,5, 0]$ , *cero* cuando  $w_{ij} \in [-0,25, 0,25]$ , *positiva débil* cuando  $w_{ij} \in [0, 0,5]$ , *positiva media* cuando  $w_{ij} \in [0,25, 0,75]$ , *positiva fuerte* cuando  $w_{ij} \in [0,5, 1]$  y *positiva muy fuerte* cuando  $w_{ij} \in [0,75, 1]$ . De esta forma los valores de la matriz  $W$  se asignan de manera aleatoria siguiendo una distribución uniforme con valores dentro de los intervalos especificados según el tipo de relación entre los conceptos del mapa. Por ejemplo, si el efecto causal entre el concepto  $C_i$  sobre  $C_j$  se presume como positivamente medio, un número aleatorio dentro del intervalo  $[0,25, 0,75]$  es fijado para el elemento  $w_{ij}$  correspondiente. Si no existe relación de causalidad entre ambos conceptos el valor  $w_{ij} = 0$ . En la tabla 2-2 se resumen las relaciones causales entre los conceptos del mapa y los rangos dentro de los cuales pueden tomar valores los pesos  $w_{ij}$ . La regla que se utiliza para actualizar los valores de la matriz  $W$  se detalla en la ecuación 2-3.

$$w_{ji}^t = \begin{cases} 0 & , \text{ si } w_{ji}^{t-1} = 0 \\ \xi w_{ji}^{t-1} + \eta x_i^{t-1} (x_j^{t-1} - |w_{ji}^{t-1}| x_i^{t-1}) & , \text{ si } w_{ji}^{t-1} \neq 0, \end{cases} \quad (2-3)$$

donde  $\eta$  es el coeficiente de aprendizaje y toma valores muy pequeños cercanos a cero y  $\xi$  es la decaída del coeficiente de aprendizaje de los pesos  $w_{ji}$ . Otro aspecto a tener en cuenta es el criterio de terminación del algoritmo. Mei y col., 2014 proponen dos criterios diferentes denominados  $F1$  y  $F2$ . El primero de ellos establece que debe minimizarse la distancia euclídea entre el valor actual del concepto de salida y el valor esperado. De esta forma  $F1 = \sqrt[2]{DOC_{10}^t - DOC_{10}^{esperado}}$ , como se requiere que  $DOC_{10} \in [DOC_{10}^{min}, DOC_{10}^{max}]$ , el valor de  $DOC_{10}^{esperado}$  se fija como:  $DOC_{10}^{esperado} = (DOC_{10}^{min} + DOC_{10}^{max})/2$ . El criterio  $F2$  asegura que el entrenamiento termine luego de un número determinado de iteraciones. El segundo criterio toma la forma  $F2 = |DOC_{10}^{t+1} - DOC_{10}^t| < \epsilon$ , siendo  $\epsilon$  un valor muy pequeño cercano a 0 que se resume en el siguiente algoritmo 1.

---

#### Algoritmo 1 Entrenamiento del FCM

---

{Determinar la Matriz inicial  $W_0$  que determina la relación causal entre los conceptos}  
 {Asignar los valores iniciales de los conceptos  $C_i^0 = x_i^0$ }  
 {Calcular  $x_i^t$  de acuerdo a la ecuación 2-1}  
 {Calcular  $w_{ji}^t$  de acuerdo a la ecuación 2-3}  
 {Si  $w_{ji}^t > 1$ , entonces  $w_{ji}^t = 1$ . Si  $w_{ji}^t < -1$ , entonces  $w_{ji}^t = -1$  para que  $w_{ji}^t \in [-1, 1]$ }  
 {Comprobar los criterios de terminación  $F1$  y  $F2$ , si ninguno de ellos se cumple, volver al paso 2 }

---

El algoritmo 1 asegura la convergencia de los conceptos del mapa luego de un número limitado de iteraciones. Los intervalos utilizados para establecer el valor de cada entrada

**Tabla 2-2.:** Pesos entre los diferentes conceptos del FCM.

$w_{ij}$	Concepto causa	Concepto efecto	intervalo	Relación causal
$w_{14}$	$C_1$	$C_4$	[0,25, 0,75]	<i>positiva media</i>
$w_{24}$	$C_2$	$C_4$	[-1, -0,75]	<i>negativa muy fuerte</i>
$w_{210}$	$C_2$	$C_{10}$	[-1, -0,5]	<i>negativa fuerte</i>
$w_{34}$	$C_3$	$C_4$	[0, 0,5]	<i>positiva débil</i>
$w_{46}$	$C_4$	$C_6$	[0,25, 0,75]	<i>positiva media</i>
$w_{56}$	$C_5$	$C_6$	[0,25, 0,75]	<i>positiva media</i>
$w_{69}$	$C_6$	$C_9$	[0,75, 1]	<i>positiva muy fuerte</i>
$w_{610}$	$C_6$	$C_{10}$	[0, 0,5]	<i>positiva débil</i>
$w_{74}$	$C_7$	$C_4$	[-1, -0,75]	<i>negativa muy fuerte</i>
$w_{75}$	$C_7$	$C_5$	[-1, -0,75]	<i>negativa muy fuerte</i>
$w_{78}$	$C_7$	$C_8$	[-0,5, 0]	<i>negativa débil</i>
$w_{79}$	$C_7$	$C_9$	[-1, -0,75]	<i>negativa muy fuerte</i>
$w_{87}$	$C_8$	$C_7$	[-0,25, 0,25]	<i>cero</i>
$w_{89}$	$C_8$	$C_9$	[-0,25, 0,25]	<i>cero</i>
$w_{96}$	$C_9$	$C_6$	[0, 0,5]	<i>positiva débil</i>
$w_{98}$	$C_9$	$C_8$	[0, 0,5]	<i>positiva débil</i>

de la matriz  $W$  es resultado del conocimiento generado por expertos en el sistema bajo estudio.

Otro aspecto a tener en cuenta es la función de umbral  $f$  en la ecuación 2-1. En el paper al que se hace referencia los autores usan una función sigmoidea de la forma:

$$f(x) = \frac{1}{1 + e^{-0,3x}}, \quad (2-4)$$

donde  $x$  es el valor del concepto  $C_i$  que se está calculando. El problema con esta función es que acota los valores de salida al intervalo  $[0,1, 0,6]$ . Una de las características de la función sigmoidea es que su dominio es el intervalo  $(-\infty, \infty)$ , mientras que el rango de la misma es  $(0, 1)$ , es decir que no se puede obtener 0 o 1 como valor para ningún concepto. Lo que es más, una función sigmoidea con  $\lambda = 5$  es un buen normalizador en  $[0, 1]$ , la pendiente de la función en  $x = 0$  es muy diferente a la pendiente en  $x = 1$ , por lo que dicha función no es la más indicada para usar como normalizador. Por este motivo se optó por usar una función lineal como función de umbral. Una función lineal, no resultaría apropiada como función de umbral para el proceso de inferencia de un FCM, porque es monótonamente creciente para en el intervalo  $(-\infty, \infty)$ . Sin embargo, si se conoce el el rango del dominio en donde pueden tomar valores  $x$  durante el proceso de inferencia, se puede utilizar una función de umbral que sea lineal (Lee, Kim, & Cho, 2012). Por ese motivo se propone una

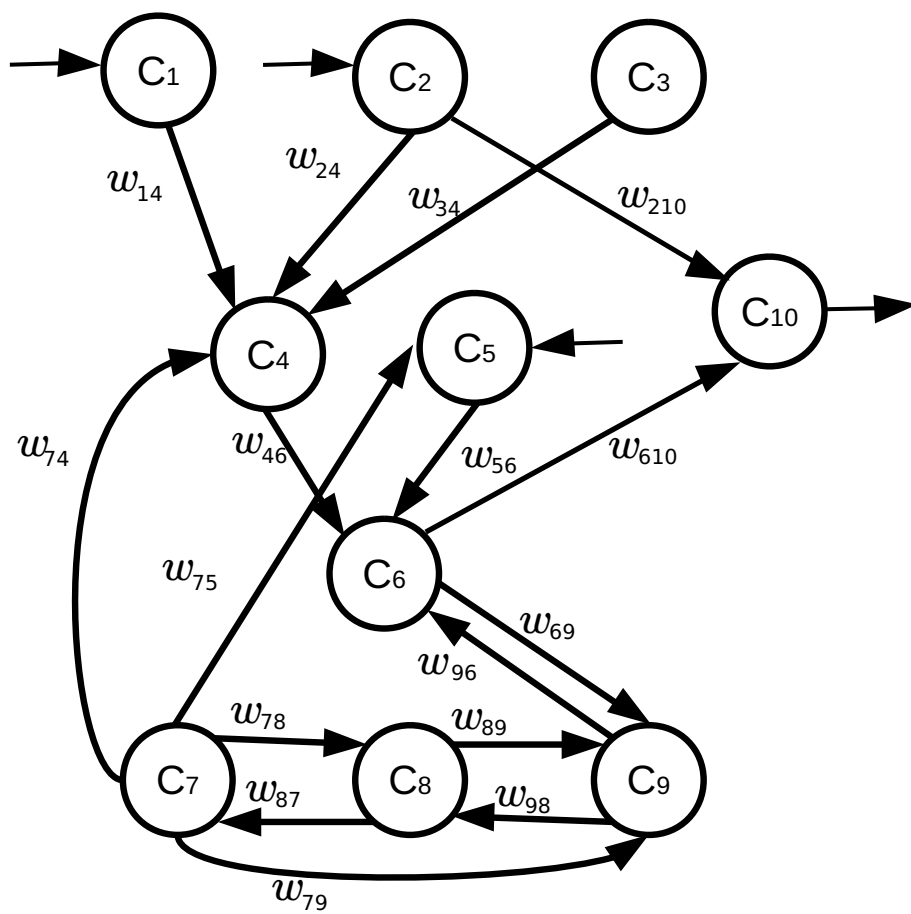


Figura 2-4.: Mapa cognitivo difuso implementado para emular el comportamiento de los individuos en la epidemia

función de la forma:

$$f(x) = \frac{1}{2}(\alpha x + 1), \quad (2-5)$$

donde  $\alpha$  toma valor según:

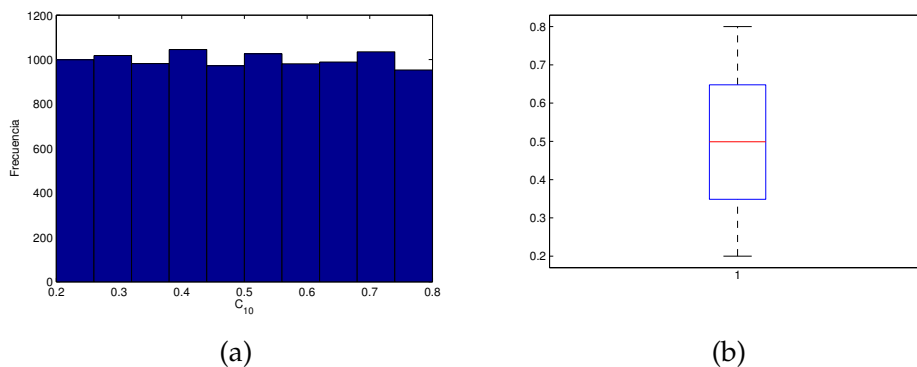
$$\alpha = \frac{0,5}{(\rho_1 + \rho_2 \|W\|)n^{1/2}}, \quad (2-6)$$

dónde  $\rho_1$  y  $\rho_2$  son valores dentro de un intervalo  $[0, 1]$ ,  $W$  es la matriz con los pesos de conexión entre conceptos y  $n$  es el número de nodos (Lee y col., 2012).

La matriz  $W$  obtenida para modelar el comportamiento individual se detalla a continuación.

$$W = \begin{pmatrix} 0 & 0 & 0 & 0,343 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0,1413 & 0 & 0 & 0 & 0 & 0 & 0 & -0,3451 \\ 0 & 0 & 0 & 0,4448 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,5255 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0,0535 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,8536 & 0,3705 \\ 0 & 0 & 0 & -0,1370 & -0,2732 & 0 & 0 & -0,0335 & -0,2507 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -0,2142 & 0 & -0,0743 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0,1426 & 0 & 0,0960 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Para validar la matriz final obtenida se generaron 10000 vectores de conceptos  $X_0$ , donde el valor de cada elemento  $x_i$ ,  $i = 1, \dots, 10$  fue generado al azar dentro del intervalo  $[0, 1]$  usando una distribución uniforme y si calculó el valor de  $X_1$  de acuerdo a la ecuación 2-1. Los resultados muestran que los valores del concepto  $C_{10}$  están siempre dentro del intervalo  $[0,2, 0,8]$ , los cuales son los valores esperados. En las figuras 2-5a y 2-5b pueden observarse los resultados de dicha validación.



**Figura 2-5.:** (a) Histograma de los valores  $C_{10}$  obtenidos (b) Gráfico de cajas

## 2.4. Comportamiento social

El *comportamiento social* determina de que forma un individuo se relaciona con otros individuos. Está relacionado con el comportamiento individual estableciendo una relación de retroalimentación. El módulo correspondiente al comportamiento social puede ser modelado de diferentes formas. A lo largo de la tesis se irán proponiendo diferentes modificaciones para evaluar sus resultados. En el capítulo 3 se utiliza un enfoque basado en *redes de contacto*. Las redes de contactos son utilizadas generalmente para la simulación y validación de los sistemas sociales reales. Las redes de contacto nos brindan un marco en el cual topología del ambiente en donde se desenvuelven los actores no es relevante. La grilla puede ser vista como un contenedor y las relaciones entre individuos se establecen en un espacio de contactos.

En el capítulo 4 el comportamiento social es modelado a través de un modelo de *espacio explícito*. En este enfoque el ambiente no solo actúa como contenedor de los individuos, el modelo puede ser visto como un sistema multi-agente en donde el ambiente es un agente más del modelo. Contemplar el espacio de manera explícita permite tener en cuenta variables como la topología de los ambientes reales, obstáculos naturales, centros urbanos, centros que actúan como repulsores. De esta forma la conducta social de los agentes es fuertemente determinada por su ambiente y por los estímulos que perciben del mismo.

## 2.5. Estado epidémico

En el análisis epidemiológico se consideran la estructura, las características epidemiológicas y el estado epidémico de una enfermedad. La estructura epidemiológica es la forma en que interactúan los factores inherentes al agente causal, ya sea físico, químico o biológico, el medio ambiente y el huésped en una población y en un período de tiempo determinado. Las características epidemiológicas son el resultado de la estructura epidemiológica y se expresan por la frecuencia y distribución de la enfermedad en la población. De acuerdo a las características epidemiológicas, la estructura epidemiológica es dinámica, modificándose según el comportamiento de la enfermedad en la comunidad.

El estado epidémico determina de que forma se modela el contagio. Este aspecto puede ser modelado de forma similar a la utilizada en los modelos clásicos *SIR* o *SEIR*. De esta forma los estados epidemiológicos pueden ser modelados como autómatas estocásticos de estados finitos. En el capítulo 3 se utiliza este enfoque para reproducir este comportamiento. La principal ventaja de este enfoque es que permite modelar diferentes situaciones epidemiológicas como cuarentena, vacunación o infecciones por múltiples cepas, simplemente modificando el autómata y ajustando las transiciones entre los estados a partir de las tasas que brindan los modelos clásicos. En el capítulo 5 se ejemplifica esto. Por otro lado, el estado epidémico se puede aproximar utilizando un enfoque más realista en cuanto al estado inmune del individuo, basado en ecuaciones diferenciales ordinarias

que modele el estado de la respuesta inmune para cada individuo en el tiempo  $t$ . De forma similar al enfoque anterior, se establecen estados epidémicos, solo que la transición ya no es determinada de manera estocástica, si no que se realiza de forma determinista a través de la integración de la respuesta inmune a nivel individual y la comparación de esta con los umbrales de transición establecidos.

## 2.6. Conclusión

Los modelos basados en individuos, poseen ventajas con respecto a los modelos clásicos a la hora de modelar sistemas biológicos ya que permiten un alto grado de complejidad que surge como producto de las interacciones entre los individuos. Estos modelos simulan poblaciones como un conjunto discreto de individuos. A diferencia de los modelos poblacionales basados en ecuaciones diferenciales ordinarias, que se describen en términos de parámetros impositivos de la población de arriba hacia abajo (como las tasas de natalidad y mortalidad), los modelos basados en individuos son modelos ascendentes en los que los comportamientos poblacionales emergen de las interacciones entre individuos autónomos entre sí y su entorno. Permiten incorporar gran variedad de mecanismos a nivel individual. Recientemente, muchos sistemas de simulación basados en agentes han sido propuestos para modelar y analizar la difusión epidémica en comunidades o ciudades. Estos trabajos brindan diferentes enfoques para resolver una amplia gama de problemas, pero se encuentran restringidos a sistemas que cumplen un conjunto acotado de propiedades. Por otro lado, son utilidades, herramientas de software o librerías que reúnen un amplio número de modelos, pero no representan un marco teórico o una nueva forma de concebir el modelado basado en individuos. El marco de trabajo propuesto en este capítulo permite gran flexibilidad y adaptabilidad por parte del modelo ya que define la dinámica de los individuos a partir del comportamiento y la construcción de dicho comportamiento a través de tres módulos acoplados hace de este enfoque un marco flexible y adaptable. En las siguientes secciones se introducirán diferentes cambios en algunos de los módulos que determinan el comportamiento de los individuos con el objetivo de modelar diferentes situaciones epidemiológicas.





# 3. Modelos basados en redes de contactos

## 3.1. Introducción

Las redes de sociales o de contactos sirven como medio fundamental para la transmisión de información, la difusión de epidemias, y la propagación de comportamiento a través de actividades de intercambio entre los individuos. Recientemente ha surgido un gran interés en el estudio de la influencia social y la dinámica de difusión en las redes de contactos. Los modelos de redes sociales son utilizados generalmente para la simulación y validación de los sistemas sociales reales mediante el uso de datos. Pero la mayoría de estos modelos de redes sociales se construyen en para estudiar las características estadísticas de la población y del modelo, como el grado de distribución, coeficiente de agrupamiento, y la estructura de la comunidad, entre otras, teniendo poca consideración de la dinámica evolutiva de los actores involucrados, que son la base de las redes sociales y también un importante tema de investigación.

Modelar el comportamiento humano para reflejar la dinámica social es una tarea difícil. En este capítulo, modelaremos el comportamiento de las redes sociales usando autómatas celulares (CA) para entender los aspectos dinámicos de un individuo, así como de un grupo de individuos en términos de las relaciones entre ellos. Los CA constituyen una técnica poderosa para modelar las interacciones entre individuos, los cuales pueden ser vistos como los nodos de una red, con estados que modelan el estado epidemiológico del individuo.

En el presente capítulo se introduce una aproximación al modelado del comportamiento social desde las perspectiva de los modelos basados en redes de contacto y su efecto en la epidemia. La implementación del modelo se hace a través del uso de autómatas celulares. En la sección 3.2 se realiza un revisión de las implementaciones mas relevantes en los últimos años de este tipo de modelos, haciendo énfasis en sus principales ventajas así como sus limitaciones y se describen las ventajas de implementar una red social por medio de autómatas celulares. En la sección 3.3 se realiza una descripción detallada de la herramienta utilizada y se describe en detalle el modelo implementado así como el ajuste de los parámetros. En la sección 3.6 se procede a presentar los resultados del modelo y finalmente en la sección 3.7 se presentan las conclusiones de este capítulo.

## 3.2. Maco teórico

### 3.2.1. Redes de contacto

Las redes sociales son grafos de individuos y sus relaciones. Ejemplos de ellos son: la comunicación por correo electrónico, colaboración científica, difusión de la innovación, la propagación de rumores y también la transmisión de enfermedades infecciosas (Moreno, Nekovee, & Pacheco, 2004; Firestone, Christley, Ward, & Dhand, 2012; Hong, Yang, Zhao, & Ma, 2016). En estos grafos, los individuos y sus relaciones se representan por medio de nodos y enlaces de entre los mismos. De este modo podemos entender la estructura y la evolución dinámica de los sistemas sociales mediante el análisis de estos grafos. La influencia social y la difusión de información en los sistemas sociales se reproducen fácilmente en este tipo de redes, muchos estudios se han centrado recientemente en este dominio (Liljeros y col., 2001; Chierichetti, Lattanzi, & Panconesi, 2011). Sobre la base de estos conjuntos de datos a gran escala, los investigadores han descubierto algunos fenómenos sociales, incluyendo el efecto *small-world* y la propiedad *free-scale* (Watts & Strogatz, 1998; Zheng, Li, & Sun, 2008; Wang, Zeng, Zheng, & Wang, 2008), que son diferentes a las hipótesis clásicas. Al mismo tiempo, con el desarrollo de las teorías de grafos, los investigadores han desarrollado algunas medidas esenciales de las redes para describir cuantitativamente las propiedades inherentes de las redes de contactos entre individuos en la naturaleza. Ejemplos de estas medidas son: grado de distribución, coeficiente de agrupamiento, la longitud del camino más corto, y la estructura de clúster, todos los cuales se miden sobre la base de las relaciones en estos diagramas.

Las mediciones de las redes del mundo real ha puesto de manifiesto la existencia de diferentes tipos de redes, que se caracterizan por una gran variabilidad en las métricas básicas y propiedades estadísticas. Esta clasificación, a su vez ha alimentado un esfuerzo de investigación teórica intensa dedicada al estudio de los diferentes modelos para intentar reproducir dichas redes. La utilidad de estos modelos es que sirven como generadores de redes sintéticas, con propiedades topológicas controlados, en los que el comportamiento de los procesos dinámicos tales como epidemias puede ser estudiado en detalle.

#### Redes homogéneas uniformes

El primer modelo teórico de redes aleatorias es el modelo de grafo aleatorio clásico (Solomonoff & Rapoport, 1951; Erdős & Rényi, 1959). En su formulación más simple, el grafo  $G_p(N)$  se construye a partir de un conjunto de  $N$  nodos en el que cada uno de los  $N(N-1)/2$  enlaces posibles está presente con una probabilidad  $p$ . El grado de distribución está dado por una función binomial, que en el límite de grado promedio  $\langle k \rangle = pN$  y  $N$  grande puede ser aproximado por una distribución de Poisson  $P(k) = e^{-\langle k \rangle} \langle k \rangle^k / k!$ . El coeficiente de agrupamiento se da simplemente por  $\langle c \rangle = p$ , y la longitud del camino más corto medio es  $\langle \ell \rangle \approx \log N / \log \langle k \rangle$  (Dorogovtsev, 2010). Este modelo es, por tanto, adecuado en

el caso de las redes que se rigen únicamente por la estocasticidad, aunque  $G_p(N)$  tiende a un grafo regular para valores grandes  $N$  y  $p$  constante. El grado de distribución alcanza su punto máximo alrededor del valor medio, denotando así homogeneidad de los nodos que lo conforman. Curiosamente, las características del modelo para  $\langle k \rangle > 1$  mas pequeño el diámetro observado en la mayoría de las redes del mundo real. Sin embargo, otras propiedades estructurales, incluyendo el generalmente alto coeficiente de agrupamiento observado en las redes del mundo real, no pueden ser reproducidos por este modelo.

### **Modelos small-world**

El modelo de Watts y Strogatz (1998) representa un primer intento de obtener una red con diámetro pequeño  $\langle \ell \rangle$  y un gran coeficiente de agrupamiento. Este modelo considera un enrejado ordenado, tal como un anillo de  $N$  vértices, cada uno de los cuales está conectado simétricamente a sus  $2m$  vecinos más cercanos. Esta configuración inicial tiene un gran coeficiente de agrupamiento y la longitud del camino más corto general promedio también es grande. A partir de ella, se reconecta una fracción  $p$  de los arcos de la red, revisando los  $m$  arcos en sentido horario de cada vértice o nodo y la conexión entre ellos propician, con una probabilidad  $p$ , a otro nodo elegido al azar. En otra versión del modelo (Monasson, 1999), una fracción  $p$  de arcos entre las parejas elegidas al azar de nodos es agregada. El efecto global de los procesos de reconexión es agregar atajos de largo alcance, que, incluso para valores pequeños de  $p \sim N^{-1}$ , reducen en gran medida la ponderación media de camino más corto, preservando al mismo tiempo una gran coeficiente de agrupación para valores no muy grandes de  $p$ . Este modelo, aunque es más adecuado para redes sociales con alto coeficiente de agrupamiento, tiene un grado de distribución y medidas de centralidad que decaen de forma exponencial rápidamente desde los valores promedio. Los modelos small world, por lo tanto, generan redes homogéneas donde el promedio de cada métrica es un valor típico compartido, con pocas variaciones, para todos los nodos de la red.

### **Redes con distribución heavy-tailed**

La evidencia empírica proveniente desde diferentes áreas de investigación ha demostrado que muchas redes del mundo real exhiben niveles de heterogeneidad no anticipada hace algunos años. Las distribuciones estadísticas que caracterizan a las redes heterogéneas son por lo general sesgadas y varían sobre varios órdenes de magnitud. Por lo tanto, las redes del mundo real son estructuradas siguiendo una jerarquía de nodos teniendo algunos de ellos mayor pesos que otros, es decir, un mayor grado de conectividad, mientras que la mayoría de los nodos tienen menor grado. De manera mas precisa, en contraste con las redes homogéneas o las redes regulares que se caracterizan por un grado típico  $k$ , cercano al promedio  $\langle k \rangle$ , las redes heterogéneas exhiben distribuciones con alto grado de variabilidad entre los pesos de los nodos que pueden ser aproximadas por medio de una distribución que siga una ley de potencias de la forma  $P(k) \sim k^{-\gamma}$ , lo que implica

una probabilidad no despreciable de vértices o nodos con alto grado. El exponente  $\gamma$  de muchas redes reales toma valores de entre 2 y 3. En tales casos, las redes se denominan *heavy-tailed*, ya que el segundo momento del grado de distribución diverge en el límite infinito de la red ( $N \rightarrow \infty$ ). Se entiende que en las redes del mundo real el tamaño finito de  $N$  y la presencia de factores biológicos y las limitaciones físicas imponen un límite superior al grado del segundo momento. Sin embargo, el segundo momento de la distribución es en muchos casos abrumadoramente grande, lo que refleja enormes fluctuaciones de conectividad. La presencia de fluctuaciones a gran escala asociados a distribuciones de colas pesadas a menudo es verdadera no sólo para el grado de los nodos, sino que también se observa para la magnitud de los valores de los enlaces de conexión, los flujos de transporte, y otras cantidades de base.

Las redes sociales son sistemas complejos que son adecuados para ser analizada por los autómatas celulares, debido a su capacidad de adaptación, auto-organización y un modelo dinámico. Existen muchos trabajos que utilizan autómatas celulares para modelar las redes sociales.

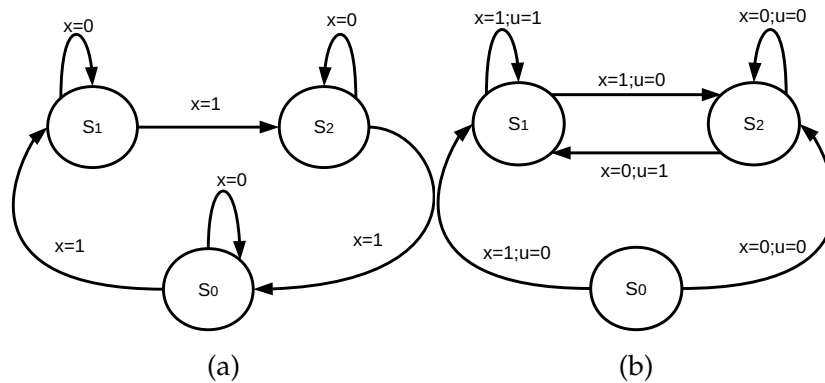
### 3.2.2. Modelo basado en autómatas

Un autómata se define como un modelo matemático que realiza cálculos en forma automática sobre una entrada para producir una salida. El mismo consta de un alfabeto, un conjunto de estados y un conjunto de transiciones entre estados. El funcionamiento básico de un autómata está dado por: (i) se parte de un estado inicial; (ii) se lee una cadena de caracteres de la entrada; (iii) a medida que se desplaza de un estado a otro por medio de una función de transición; (iv) finalmente se detiene en un estado final o de aceptación, que representa la salida del autómata. Los estados pueden ser finitos o infinitos.

Formalmente, un *autómata de estados finitos determinista* puede definirse con un quintuplete de la forma  $\mathcal{A} = \mathcal{F}(\mathcal{X}, \mathcal{U}, \mathcal{X}_0, \mathcal{E}, \mathcal{R})$ , donde:

- $\mathcal{X}$  es un conjunto finito estados.
- $\mathcal{U}$  es el alfabeto de entrada.
- $\mathcal{X}_0 \in \mathcal{X}$  es el estado inicial.
- $\mathcal{E} \subseteq \mathcal{X}$  es un conjunto de estados finales o de aceptación.
- $\mathcal{R}$  son las reglas o función transición entre estados.

La función de transición  $\mathcal{R}$  indica a qué estado se va a pasar sabiendo cuál es el estado actual y el símbolo que se está leyendo. En un autómata determinista la transición posible es sólo una, y el estado actual es una representación de la actividad pasada del sistema, que es suficientemente detallada como para servir de base para determinar qué salida producirá y cómo será actualizado por la entrada.



**Figura 3-1.:** Diagrama de estados de (a) una máquina de Moore (b) una máquina de Mealy

La salida puede depender del estado en que se encuentra el autómata, en cuyo caso recibe el nombre de *máquina de Moore* (Figura 3-1a), o depender del estado en el que se encuentra así como de la entrada que recibe, denominándose entonces *máquina de Mealy* (Figura 3-1b). En este último caso, es equivalente a que la salida dependa del estado y la entrada.

Un *autómata celular probabilístico* es aquel en el que las reglas de transición son probabilísticas, generalmente representadas como una matriz de transición. La probabilidad de pasar a cierto estado a otro puede depender del estado actual o de la entrada. También puede que el valor de salida dependa de una función de distribución de probabilidad propia de cada estado o entrada. Las transiciones mediante entradas vacías también son posibles en este caso, pudiendo asociarse una probabilidad de cambio de estado. A la hora de simular dicho autómata debe calcularse el siguiente valor de manera aleatoria, siguiendo una distribución de probabilidad (Mikler y col., 2005).

Es común representar a los autómatas mediante grafos de transición de estados. Los grafos permiten representar los estados del autómata y las transiciones posibles (Figura 3-1).

- Los nodos representan los estados;
- Los arcos representan las transiciones entre estados; y
- $C_{ij}$  representan las reglas de transición.

### **Autómatas celulares**

Los autómatas pueden usarse para simular grandes poblaciones de objetos similares si se permite alguna forma de interconexión entre ellos. La forma más común de relacionarlos y organizarlos es a través de redes de autómatas celulares, que consisten en un conjunto de autómatas (o células) que modifican sus estados siguiendo algoritmos sencillos. Los *autómatas celulares* (AC) son redes de autómatas simples conectados localmente (White, del Rey, & Sánchez, 2007). Cada autómata simple produce una salida a partir de varias entradas, modificando en el proceso su estado según una función de transición. Por lo

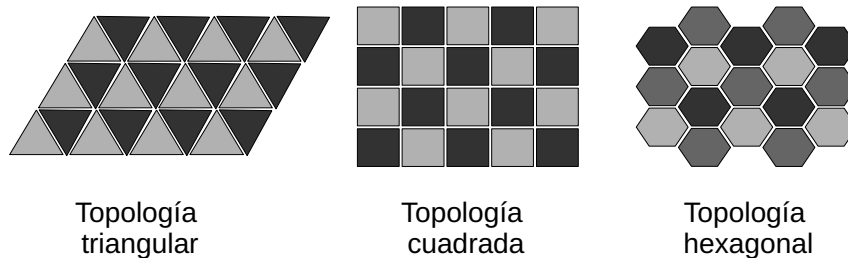


Figura 3-2.: Topologías

general, el estado de una célula en una generación determinada depende única y exclusivamente de los estados de las células vecinas y de su propio estado en la generación anterior. Mediante la interconexión de un gran número de autómatas pueden modelarse diferentes sistemas con un alto grado de complejidad. Constituyen una muy buena alternativa a los sistemas de ecuaciones diferenciales. Una red de autómatas celulares queda definida formalmente por la tupla  $\mathcal{R} = \mathcal{G}(\mathcal{T}, \mathcal{C})$  (Hoya, 2007), donde:

- $\mathcal{T}$  es la topología (forma de interconexión de las células); y
- $\mathcal{C}$  son los grados y tipos de acoplamiento entre células.

El primer paso en la conformación de una red de autómatas celulares es la selección de una geometría específica para la trama o topología (Figura 3-2). La definición de un autómata celular requiere que la trama sea regular.

Una vez determinada la trama, debemos escoger una vecindad en la que las células puedan obrar recíprocamente. La vecindad es descrita al definir el conjunto de células que colindan con una célula determinada  $(i, j)$  como se muestra en la figura 3-3. Para los vecinos más cercanos sobre una trama cuadrada en un espacio bidimensional, la vecindad denominada *vecindad de Von Neumann*, está descrita por:  $N^N(i, j) = \{(k, l) : |k - i| + |l - j| \leq r\}$ ; otras vecindades comúnmente usadas son la de Moore:  $N^M(i, j) = \{(k, l) : |k - i| \leq r, |l - j| \leq r\}$ , en ambos casos con  $r = 1, 2, \dots$  como radio de la vecindad. Por supuesto, puede definirse cualquier vecindad mientras sea uniforme y finita.

El paso siguiente es la definición de las interconexiones entre autómatas. Un autómata podrá tener una conexión con otro si se encuentra dentro de su vecindad. Los valores o pesos de las conexiones entre las células pueden clasificarse según:

- **Isotropía:** las autómatas celulares isótropos son aquellos en los que los pesos de conexión son iguales en todos los sentidos y direcciones. En contraposición, un autómata celular anisótropo tiene al menos un peso diferente en algún sentido o dirección de la trama.
- **Direccionalidad:** se habla de tramas conectadas bidireccionalmente cuando entre cada autómata de la vecindad existen pesos de ida y vuelta. Por el contrario, se

habla de unidireccionalidad cuando hay un peso desde una célula a otra pero no existe el recíproco (Figura 3-2).

Así como un autómata puede representarse mediante un grafo de estados, los autómatas celulares se pueden representar como un grafo en el que cada nodo es un autómata de estados finitos. Este grafo se organiza como una grilla de células en dos dimensiones que evolucionan de acuerdo a una función de actualización que se aplica de manera uniforme a todas las células. De esta forma queda discretizado el espacio y el tiempo. El estado actual de la célula y sus vecinas se toma como argumento de entrada en cada paso de actualización. Los pesos de conexión indican cuánto influyen las vecinas sobre la célula actual (Coburn, Wagner, & Blower, 2009).

Otro aspecto a tener en cuenta son las condiciones de contorno. En la definición formal de un autómata celular comúnmente se requiere que la trama sea infinita en todas las dimensiones. Naturalmente, es imposible simular una trama infinita en una computadora. Por lo tanto, tenemos que definir algunas condiciones de contorno. Se pueden considerar tres de tipos de contornos o límites:

- **Frontera fija:** Todas las células fuera de la trama tienen un valor fijo.
- **Frontera periódica:** se considera la rejilla como si sus extremos se tocaran. Para una trama de una dimensión, esto puede visualizarse en dos dimensiones como una circunferencia. Para una trama de dos dimensiones, podría visualizarse en tres dimensiones como un toroide.
- **Frontera reflectora:** las células fuera de la rejilla reflejan los valores de las que están dentro de la rejilla.

No solamente las condiciones de contorno necesitan ser especificadas, también se deben definir las condiciones iniciales. En la mayoría de los casos las condiciones iniciales determinan en gran medida la evolución. Las condiciones iniciales pueden ser definidas célula a célula en forma determinista o pueden generarse aleatoriamente.

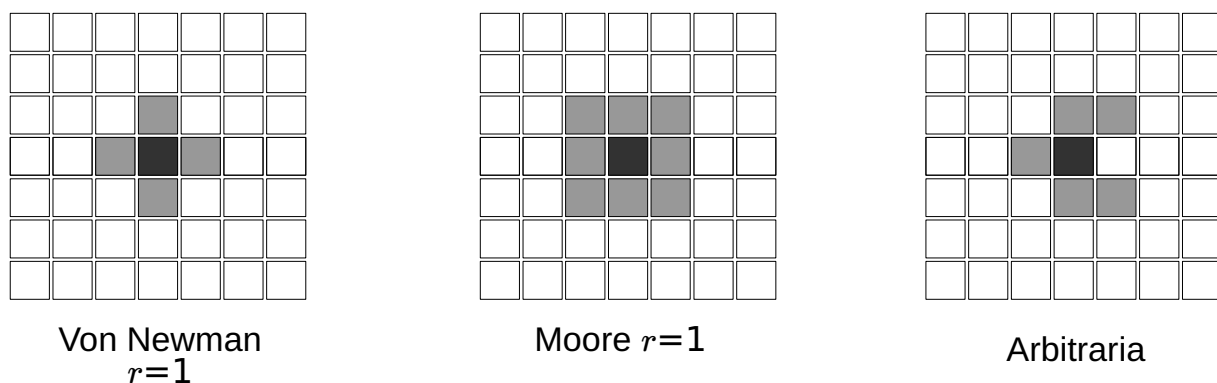


Figura 3-3.: Tipos de vecindades



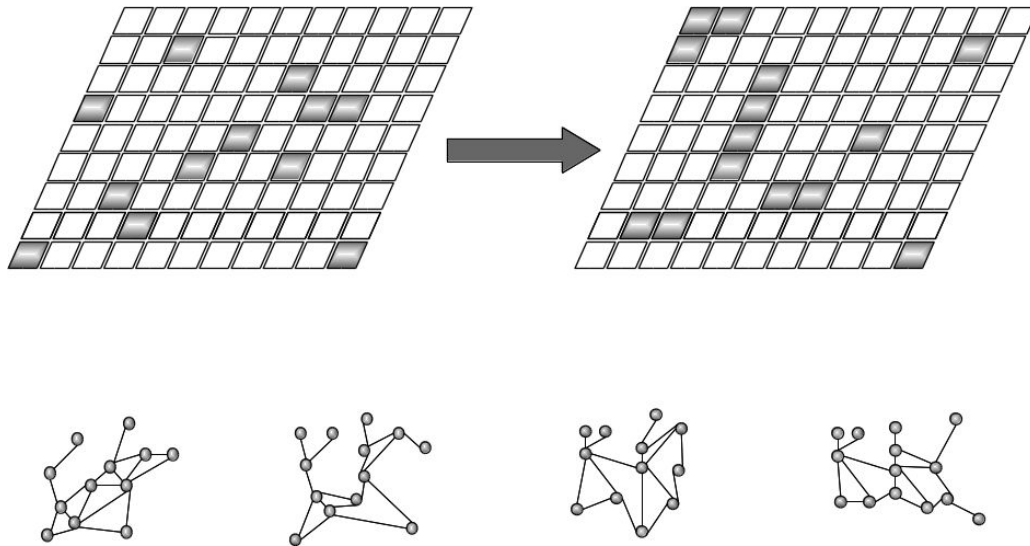
Según la definición de autómata celular, cada célula es un autómata finito y, por lo tanto, el conjunto de estados tiene tamaño finito. Comúnmente el conjunto debe ser bastante pequeño. Hay varias razones para esto: (i) el número de estados posibles, con  $s$  estados por célula y  $n$  células en la vecindad (incluyendo la célula para ser actualizada), es  $s * n$ . Este número rápidamente llega a ser demasiado grande como para explorar todas las posibilidades cuando  $n$  y  $s$  crecen; (ii) cuando hay pocos estados es posible especificar las reglas de transición en forma explícita y almacenarlas en una tabla (el tamaño de la tabla es  $s * n$ ). Una razón para usar un número grande de estados puede ser para aproximar un autómata celular con estados finitos a un sistema continuo.

Otro de los aspectos importantes de un autómata celular son las reglas de transición. Son las que determinan la evolución del sistema. Las reglas de transición dependen de la geometría de la trama, de la vecindad y del conjunto de estados. En general, pueden especificarse de muchas maneras diferentes, la especificación más directa consiste en anotar el resultado de la regla de transición para cada configuración posible de estados en la vecindad. Una clase importante de reglas de transición son las probabilísticas. En este caso la regla de transición no es una función que tiene exactamente un resultado para cada configuración de la vecindad, sino que es una regla que provee uno o más resultados posibles con probabilidades asociadas. Una propiedad importante de los autómatas celulares es que un cambio muy pequeño en la regla de transición puede tener efectos muy grandes sobre la dinámica de la red.

### 3.2.3. Autómatas celulares como redes de contactos

Un autómata celular también puede ser definido como una red compuesta de nodos y enlaces dinámicos entre ellos. Los enlaces pueden ser dirigidos o no dirigidos, estos últimos pueden ser representados como un par de enlaces dirigidos colocados simétricamente. Cada nodo toma un estado de un conjunto finito o infinito de estados ( $\mathcal{X}$ ). Los enlaces describen las relaciones referenciales entre los nodos, que especifica cómo los nodos se afectan entre sí en estado de transición y transformación topológica. Cada enlace también puede tomar uno de los posibles estados en un conjunto de estado de enlace al que se lo llama  $\mathcal{X}'$  (ver figura 3-4). Una configuración de una red de contactos en un tiempo  $t$  determinado es una combinación de estados y topologías de la red, que es formalmente definida como se especifica a continuación:

- $\mathcal{V}_t$ : es un conjunto finito de nodos de la red en el momento  $t$ . Aunque por lo general se supone como invariante en el tiempo en la teoría de los sistemas dinámicos convencionales, este conjunto puede cambiar de forma dinámica en el marco propuesto debido a las adiciones y bajas de nodos.
- $\mathcal{C}_t : \mathcal{V}_t \rightarrow \mathcal{X}$ : es una regla de mapeo que asigna a cada elemento del conjunto de



**Figura 3-4.:** Visualización de una grilla de autómatas como una red de contactos

nodos,  $\mathcal{V}_t$ , otro elemento del conjunto de estados  $\mathcal{X}$ .

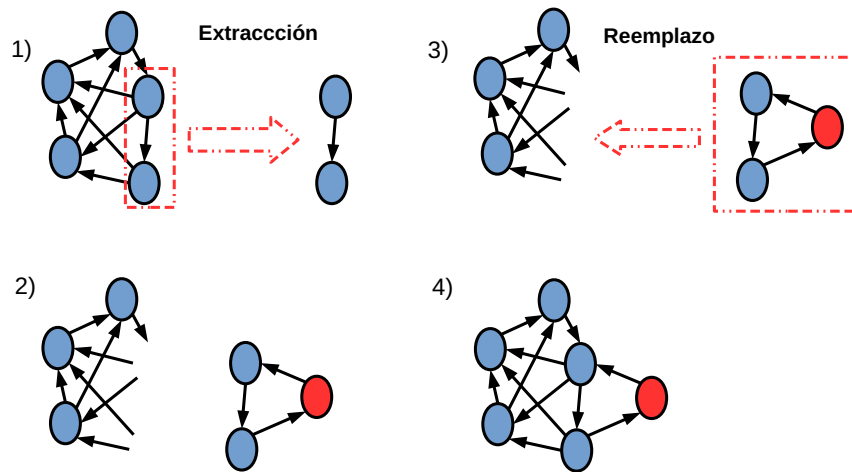
- $\mathcal{L}_t : \mathcal{V}_t \rightarrow \mathcal{V}_t \times \mathcal{X}'^*$ : es una regla de mapeo que asigna a cada elemento del conjunto  $\mathcal{V}_t$  una lista de elementos (enlaces) del conjunto  $\mathcal{X}'$ . Dicho de otra forma, esta regla asigna la topología global de la red en el tiempo  $t$ , que también puede variar con el tiempo.

Los estados y las topologías de una red de contactos se actualizan a través de eventos repetitivos de reescritura de la misma, cada uno de los cuales consiste en tres pasos:

1. La extracción de parte de la red, es decir una sub-red que estará sujeta a cambios.
2. Producción de una nueva sub-red que reemplazará a la sub-red seleccionada anteriormente.
3. Incorporación de la nueva sub-red al resto de la red.

La dinámica temporal de este tipo de modelos puede ser formalmente definida a través del triplete  $\langle E, R, I \rangle$ , donde:

- $E$  es el mecanismo de extracción que determina qué parte de la red se selecciona para la actualización. Se define como una función que toma toda la configuración de la red y devuelve una sub red específica en que se sustituirá.
- $R$  es el mecanismo de sustitución que produce un nuevo segmento de red a partir del segmento seleccionado por  $E$  y también especifica la correspondencia de los nodos entre los viejos y nuevos segmentos. Se define como una función que toma una



**Figura 3-5.:** Proceso de reescritura de la red de contactos

configuración y devuelve una nueva configuración con las correspondiente asignación entre los nodos del segmento de entrada y los nodos del nuevo segmento generado.

- $I$  es una configuración inicial para la red de contactos.

El triplete  $\langle E, R, I \rangle$  introducido antes nos brinda un conjunto de reglas suficientes para definir de forma única un modelo basado en redes de contactos. Toda la imagen de un evento de reescritura se ilustra en la Figura 3-5, que muestra visualmente cómo estos mecanismos trabajan juntos. El proceso de reescritura se puede resumir de la siguiente manera: **(a)** el mecanismo de extracción  $E$ , toma parte de la red de contactos; **(b)** el mecanismo de reemplazo  $R$  produce una nueva sub red como reemplazo de la sub red seleccionada por  $E$  y también especifica la correspondencia de nodos entre la sub red de entrada y la de salida así como las transformaciones de las topologías locales; **(c)** la nueva sub red es embebida con el resto de la estructura de acuerdo a la correspondencia de nodos determinada por  $R$ . Para este ejemplo particular, el nodo gris de mas arriba en la sub red de entrada no tiene correspondencia con ningún nodo en la nueva sub red, por lo tanto, las conexiones correspondientes son eliminadas; **(d)** se obtiene la versión actualizada de la red de contactos. Este proceso, en general, no se puede aplicar de forma sincrónica a todos los nodos o sub redes en una red, porque las modificaciones simultáneas de topologías de redes locales en más de un lugares pueden causar resultados contradictorios que sean incompatibles entre sí. Esta limitación no se aplica, sin embargo, cuando no hay posibilidad de conflictos topológicos, por ejemplo, cuando las reglas de reescritura son libres de contexto, o cuando el autómata se utiliza para simular redes dinámicas convencionales que implican solamente los cambios de estado locales, pero no hay cambios topológicos.

### 3.2.4. Estado del arte

En la literatura actual existen muchas implementaciones de modelos poblacionales en base a modelos celulares y redes de contacto. Las formas de encarar el modelado son diversas, pudiendo agruparse en varias categorías, según modelen en forma continua o discreta el espacio, el tiempo o los individuos. En el caso específico de los autómatas celulares, la evolución temporal se efectúa siempre en pasos discretos, y el espacio está conformado por una grilla fija, donde cada celda se encuentra en un estado de varios posibles dentro de un conjunto finito.

Dado el interés de capturar los efectos de la heterogeneidad poblacional sobre la evolución temporal de la epidemia, el modelado de la población en forma discreta (con un individuo por celda a lo sumo), se evidencia como la forma natural. En este caso, los estados posibles de cada celda se derivan directamente de los compartimentos definidos en el modelo clásico, contemplando además un estado adicional para el caso de las celdas vacías.

Sin embargo, en la literatura abundan los trabajos en los que se considera a cada celda como una distribución homogénea de individuos de cada tipo (Allouche, Kalyuzhny, Moreno-Rueda, Pizarro, & Kadmon, 2012; Murillo, Murillo, & Perelson, 2013; Ovaskainen y col., 2014). En estos casos las celdas representan áreas de igual tamaño que contienen una población específica. Diferentes células tendrán diferentes poblaciones, es decir, diferentes densidades y posiblemente diferentes propiedades de movilidad. Se modela la dispersión de la infección mediante el ajuste de la fuerza de contacto entre cada celda y las incluidas en su vecindad. Las infecciones ocurren por contacto entre individuos de una misma celda o de celdas vecinas. En este caso, el estado de cada autómata puede considerarse como uno solo, y lo que se visualiza es la salida de cada celda, consistente en la cantidad de individuos correspondiente a uno u otro compartimento en cada instante de tiempo. Las ecuaciones diferenciales del modelo clásico están incluidas explícitamente en cada celda, y la evolución de la epidemia en cada una sigue la evolución temporal del modelo clásico, con las modificaciones que puedan surgir del paso de individuos desde o hacia una celda en particular. Sin embargo, tomando a todos los autómatas como un conjunto, la evolución temporal de la epidemia es similar a la del modelo clásico. Muchos modelos usan reglas deterministas para la actualización de las celdas. En general, las actualizaciones de las celdas se realizan en varias fases: fase de movimiento, fase de exposición, infección, recuperación, etc. Si bien existen mínimas diferencias en los resultados, en general es indistinto el orden en el que se implementen las fases. La fase de movimiento es distinta para los casos de población discreta o continua, pero en ambos casos consiste en una simulación de movimiento poblacional mediante el cambio de la salida o del estado de celdas vecinas.

Wang y Chen (2013) proponen un modelo *SEIR* implementado a través de un autómata celular en donde cada célula representa a una población en particular como puede ser un

núcleo urbanístico o rural. La propagación de la epidemia entre nodos es modelada utilizando un enfoque redes de contacto free-scale. Los vecinos de cada célula son aquellos entre los cuales existe una vía de comunicación que permita el flujo de población desde uno hacia el otro. Las suposiciones en las que se basa este modelo son:

- El número total de individuos en el espacio celular y en cada nodo en particular permanece constante a lo largo del tiempo. Es decir que no se tienen en cuenta los procesos migratorios, ni los nacimientos o las muertes.
- La vía de contagio de la enfermedad es en forma directa, por contacto físico entre un individuo infectado y un individuo susceptible (principio de acción de masas).

En este modelo lo que se implementa es un modelo poblacional clásico en cada célula. Presupone una distribución homogénea de la población en cada nodo y heterogénea entre nodos. Es decir, que el enfoque es idéntico al clásico con el agregado de un término de propagación entre celdas producto del marco que otorgan las redes de contactos.

Otro ejemplo con una metodología similar es propuesto por Athithan, Shukla, y Biradar (2014). En este caso se implementa un modelo *SIR* clásico basado en ecuaciones diferenciales ordinarias. El trabajo explora el comportamiento espacial de la epidemia en enfermedades que tienen carácter estacional. Para esto se considera una tasa de infección variable en el tiempo, utilizando una función sinusoidal. No obstante, la tasa es la misma en todas las celdas en un instante dado de tiempo. Para simular el movimiento espacio-temporal asociado a las diferentes olas de una epidemia se implementa un modelo denominado *dependiente de la vecindad* donde el rango de movimiento está limitado por el tamaño de las vecindades utilizadas, de modo que las ecuaciones no evolucionan solo en función del tiempo  $t$  si no también en función del espacio  $r$ . Es, en definitiva, un sistema de ecuaciones diferenciales parciales discretizado en tiempo y espacio. El primer paso en cada iteración corresponde a la determinación de la dinámica local y el estado de cada celda independientemente de los estados de las celdas vecinas. El segundo paso corresponde al movimiento espacial entre vecinos, en el que cada autómatas calcula la suma de migraciones desde y hacia cada celda de su vecindad, definida como de Moore.

Un enfoque similar fue propuesto por Fu (2002), el cual consiste en la implementación de un modelo *SIR*, cuyas celdas tienen los siguientes atributos:

- Capacidad de acarreo.
- Población total.
- Población susceptible.
- Población infectada.
- Población recuperada.

La cantidad de individuos por célula es variable, lo que permite incluir varios individuos por celda, reduciendo el número total de celdas necesarias y por lo tanto el tiempo de cómputo requerido para la simulación. En segundo lugar, provee generalidad al modelo. La capacidad de acarreo es la cantidad máxima de individuos por célula, que define una densidad máxima de población. La grilla es rectangular y el movimiento de individuos sólo se produce entre células adyacentes, el ambiente en sí representa una barrera natural para el movimiento migratorio. Los parámetros que explican el comportamiento de la epidemia se manejan en forma probabilística con excepción del radio efectivo de infección. Este último determina la distancia máxima que el virus puede moverse fuera de un huésped en particular. Por otro lado, el radio de movimiento determina la máxima distancia que un huésped determinado puede moverse y está asociado a una probabilidad de movimiento. Esta probabilidad determina la frecuencia con la que puede ocurrir dicho evento dentro de una célula o entre células vecinas. Se asume una población distribuida en forma homogénea en cada célula. La célula hacia la cual ocurre un proceso migratorio desde la célula actual se elige en forma aleatoria entre las vecinas a la actual. Un parámetro de inmigración permite a este modelo simular un mundo abierto o cerrado dependiendo del valor de éste. Este parámetro define la probabilidad de cada célula de recibir inmigrantes, este evento también se encuentra afectado por la capacidad de carga. El modelo también incorpora una tasa de nacimiento y muerte natural que se refieren a la incorporación de un nuevo individuo a la célula por reproducción natural en el caso del primer parámetro y a la desaparición de uno en el caso del segundo, este último afecta a todos los individuos de la población. Estos eventos pueden ocurrir con una probabilidad asociada a cada uno en cada paso de tiempo. La morbilidad del virus es una medida de cuan rápido el virus mata al huésped, sólo afecta a los que se encuentran en el estado de infectado. En este modelo también se incorporan otros parámetros probabilísticos como la probabilidad de infección por contacto con vectores, la probabilidad de re-infección y la probabilidad de infección espontánea.

Li y Guo (2012) propusieron un modelo basado en autómatas que incluye dos competidores. Un competidor inferior que es inmune a la enfermedad y un competidor superior que es susceptible a la enfermedad. A través del análisis de series temporales y análisis de patrones espaciales, los autores exploran el efecto de la competencia en el comportamiento de una epidemia, viendo si la competencia es favorable para la aplicación de políticas de control. Luego, introducen diferentes estrategias para los competidores para evaluar como evoluciona el sistema. El autómata celular es entonces una grilla en la que cada célula representa un parche diferente que puede estar ocupado por una porción de población de alguna de las dos especies. De este modo, el modelo provee de un marco espacial para la evolución temporal del modelo de parches. Si bien este enfoque brinda la incorporación de la componente espacial de manera explícita, la dinámica de las poblaciones dentro de cada parche sigue evolucionando a través de sistemas deterministas.

Cissé, El Yacoubi, y Tridane (2013) utilizan la lógica de los modelos compartimentales

*SEIR* agregando términos que describen características espaciales. El modelo puede resumirse teniendo en cuenta cuatro características: *i*) el espacio es representado a través de un grilla rectangular en donde cada célula es denotada por un conjunto de coordenadas  $(i, j)$ ; *ii*) el estado de una célula  $(i, j)$  en el instante de tiempo  $t$  se denota como  $s_{i,j}^t$  y está caracterizado por el cuádruple  $(S_{i,j}^t, E_{i,j}^t, I_{i,j}^t, R_{i,j}^t)$  y correspondo a la porción de individuos en cada uno de los cuatro estados posibles determinados por el modelo clásico. *iii*) Se define una vecindad cuadrada  $\mathcal{N}$  y *iv*) La función de transición de estados corresponde a un modelo *SEIR*. Dicho de otra forma, los autores proponen un modelo en el cual cada célula de la grilla representa una población local que evoluciona según la dinámica determinada por el modelo *SEIR*. En el modelo no se contempla el movimiento de los individuos y se dejan de lado otras características espaciales importantes.

Firestone y col. (2012) aplican análisis de redes sociales para describir la propagación de la gripe equina entre caballos infectados durante un brote temprano de la enfermedad, identificando la propagación se produce a través de la red de contactos y una propagación espacial local secundaria. En este trabajo, las redes sociales se construyeron combinando datos del movimiento de los caballos y la matriz de distancia entre los caballos infectados y sanos dentro de los primeros 10 días del brote. Estas redes fueron analizadas para proporcionar una descripción de la epidemia, identificar las premisas que son fundamentales para la propagación de la enfermedad y para estimar la proporción relativa de los grupos infectados a través de movimientos de caballos enfermos y por la extensión espacial local. A continuación, exploraron el efecto de la distancia en la propagación de la enfermedad mediante la estimación del rango de propagación local basado en el nivel de fragmentación de la red y también por estimación directa de la forma de transmisión del núcleo espacial del brote.

Hong y col. (2016) modelan la propagación de una epidemia considerando la movilidad de los nodos en una red dinámica compleja basándose en la teoría de campo medio haciendo uso ecuaciones diferenciales para modelar la dinámica poblacional. Exploró el efecto de la heterogeneidad en el proceso de evolución de la dinámica de propagación, llevando a cabo extensas simulaciones para estudiar los parámetros de influencia clave y la influencia de ellos sobre la dinámica de propagación, que son útiles para la comprensión del mecanismo de propagación de la epidemia y el diseño de estrategias de control eficaces. Los autores proponen un modelo de propagación basado en modelo de movilidad al azar, que puede complementar los puntos insuficientes en el modelo de propagación en una red dinámica compleja. Además, introdujeron la heterogeneidad agregando nodos que resisten la propagación de la epidemia. Pese a todas estas características, algunos aspectos inherentes a la poblaciones no son tenidos en cuenta en este trabajo, por ejemplo de los nodos modelan la dinámica poblacional en términos de ecuaciones diferenciales siguiendo la lógica de los modelos clásicos. De esta manera, el modelo se acerca a la lógica de los modelos metapoblacionales.

Finalmente Song, Jiang, y Song (2014) proponen un modelo de red adaptativa pondera-

da, en el que la topología de la red, los estados de los nodos y los pesos interactúan mutuamente. Por otra parte, también se estudian las dinámicas epidemiológicas en redes de adaptación ponderadas, y dos estrategias adicionales de conexión entre los nodos en base a los pesos son presentadas. Además, dichas estrategias demostraron que el tipo de conexión juega un rol preponderante en la prevención de la propagación de la enfermedad. En este trabajo, los autores se centran en la interacción entre los estados nodos, intensidad de la conexión de red y la topología, y proponen un modelo de red adaptativa ponderado. A continuación, se centran en la dinámica de epidemia en las redes ponderados con el fin de analizar el proceso de difusión

Estos ejemplos muestran una forma de modelado que es útil para visualizar patrones de distribución de la epidemia en las dimensiones espaciales, pero no captura ni explica los patrones y fluctuaciones observados en la dimensión temporal. Además, la incorporación de individuos como una distribución espacial completamente homogénea adolece de la misma simplificación no realista aplicada en los modelos clásicos, que es la que estamos tratando de superar. El modelado de individuos de forma explícita se evidencia como una aproximación más realista. Las diferentes hipótesis con las que se construye el modelo clásico son aplicadas directamente a cada individuo, y las ecuaciones diferenciales del modelo clásico quedan incorporadas en forma implícita a través de reglas. La evolución temporal de la epidemia observada en el modelo clásico surge aquí como propiedad emergente de los comportamientos individuales tomados en conjunto. Esto permite también comprobar hasta qué punto son sostenibles las diversas hipótesis sobre las que descansa el modelo clásico.

### 3.3. Modelo

En la Figura 3-6 se describe mediante un diagrama el modelo propuesto. Teniendo en cuenta el marco propuesto en el capítulo 2 puede apreciarse que las modificaciones introducidas afectan el *comportamiento individual*, que es modelado a través un conjunto de reglas y *FCM*, el *comportamiento social* que es modelada través de una red de contactos implementada como un autómata celular y el *estado epidémico* que es modelado a través de un autómata probabilístico de estados finitos.

Una grilla de  $N \times N$  define el espacio de contactos en donde cada celda representa una porción de dicho espacio que puede estar ocupada por un individuo en alguno de los estados epidémicos posibles o bien vacía. Cada celda dentro de la grilla corresponde a un nodo de la red o un elemento del conjunto  $\mathcal{V}$ . Es posible simular el movimiento de los individuos a través del cambio recíproco de estados entre elementos adyacentes en el espacio de contactos, es decir, una celda en el estado  $E_1$  pasa a  $E_2$  mientras que la celda adyacente pasa del estado  $E_2$  a  $E_1$ . Este movimiento busca emular en forma aproximada el patrón de contactos de los individuos reales y contribuye a la distribución homogénea



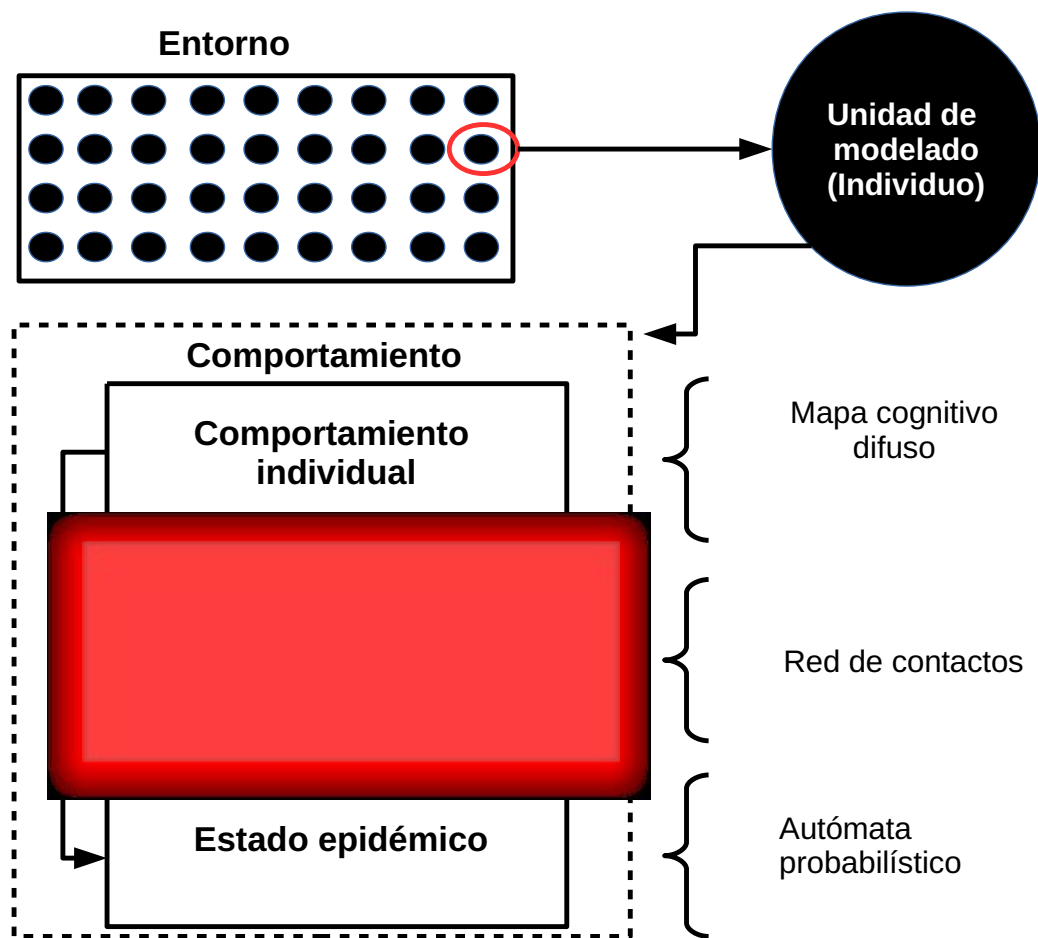


Figura 3-6.: Diagrama del modelo basado en redes de contacto

**Tabla 3-1.:** Matriz de transición

	S	E	I	A	R
S	$1 - \mu$	0	0	0	0
E	0	$1 - (\varepsilon + \mu)$	0	0	0
I	0	0	$1 - (\gamma + \mu)$	0	$\gamma$
A	0	0	0	$1 - (\gamma + \mu)$	$\gamma$
R	0	0	0	0	$1 - \mu$
D	$\mu$	$\mu$	$\mu$	$\mu$	$\mu$

de contactos. El movimiento de individuos en el el espacio de los contactos repercute de manera directa en la estructura de la red en cada instante de tiempo  $t$ . La conexión de los nodos en la red de contactos es unidireccional.

El estado epidémico es modelado como una autómatas estocástico de estados finitos. La función de transición de estados se define como  $\mathcal{R} : \mathcal{X} \rightarrow \mathcal{X}$ . Las probabilidades de transición derivan de los parámetros del modelo clásico (Chowell, Ammon, Hengartner, & Hyman, 2006). Estos parámetros son deterministas en dicho modelo, pero como resultado de una agregación de probabilidades individuales bajo la hipótesis de una población suficientemente grande. Por lo tanto, es la transición probabilística la que sustenta al modelo clásico y no al revés. Si  $\mathcal{R}$  se aplica e forma directa a cada nodo del conjunto  $\mathcal{V}$ , se evidencia una variabilidad decreciente al aumentar la población, hasta coincidir con la evolución determinista del modelo clásico. La distribución inicial de los individuos en la red puede ser al azar, cumpliendo con la hipótesis de distribución homogénea para tamaños grandes de población. El modelo también permite explorar diferentes distribuciones espaciales para los individuos así como el número de individuos. El contacto potencialmente infeccioso se efectúa entre los individuos infecciosos (sintomáticos y asintomáticos) y los susceptibles que se encuentren dentro de la vecindad definida como zona de influencia de los individuos infecciosos. La infección es también probabilística y esta relacionada con el parámetro correspondiente del modelo clásico. Por simplicidad, la grilla usada para la construcción del autómatas celular es de tipo rectangular, y la vecindad es de tipo Moore, con un tamaño  $\nu$  que puede variar, sin embargo es posible utilizar otros tipos de grillas y vecindades. Esto cambia grado de de conectividad entre nodos en la red de contactos. La condición de frontera es fija, con un contorno compuesto por celdas vacías no interactuantes.

El estado epidémico queda entonces definido como una máquina de Moore estocástica mediante el séxtupla  $\mathcal{A} = (\mathcal{X}, \mathcal{U}, \mathcal{E}, \mathcal{R}, \Lambda, \mathcal{P}_0)$ , donde  $\mathcal{X} = \{S, E, I, A, R, D\}$ , es el conjunto finito de estados, comprende seis estados:  $S$  (susceptible),  $E$  (expuesto),  $I$  (infeccioso sintomático),  $A$  (infeccioso asintomático),  $R$  (recuperado) y  $D$  (muerto o vacío).  $\mathcal{U}$  es el conjunto de entrada (números reales). Un individuo únicamente recibe una entrada no

nula cuando se encuentra en estado  $S$ , emitida por otro en estado  $I$  o  $A$ , cuando el receptor se encuentra dentro de la vecindad del emisor.  $\mathcal{R}$  es la función de transición de estados, aplicada sobre el estado activo en la iteración  $k$ , decide probabilísticamente el estado activo en la iteración  $k + 1$ . La función se aplica en dos pasos, uno correspondiente al cambio de estados por infección y recuperación y otro correspondiente al movimiento. Para decidir los cambios de estado se definen dos matrices de probabilidad, una para la transición sobre entrada vacía  $\varepsilon$  (Tabla 3-1, donde  $\mu$  es la probabilidad de muerte natural,  $\gamma$  es la probabilidad de recuperación) y otra para la transición por contacto con infecciosos (Tabla 3-2) en la cual teniendo en cuenta el tamaño de vecindad  $\nu$  y al valor de entrada como  $\lambda \in \mathcal{U}$ . Teniendo en cuenta el tamaño de vecindad  $\nu$  y al valor de entrada como  $\lambda$ .  $\mathcal{E}$  es el conjunto de salida, la salida es no nula cuando el estado del nodo es  $I$  o  $A$ .

**Tabla 3-2.:** Matriz de transición por contacto

	S	E
S	$1 - \lambda/\nu$	$\lambda/\nu$
E	0	1

La función de salida  $\Lambda$ , definida como:  $\Lambda : \mathcal{X} \rightarrow \mathbb{R}$  emite la tasa de infección del autómata, si éste se encuentra en estado  $I$  ( $\beta$ ) o  $A$  ( $q\beta$ , siendo  $q$  la probabilidad de infección por contacto con un individuo asintomático) dicha tasa es no nula. Este valor está relacionado con el  $\beta'$  del modelo clásico, pero no es exactamente el mismo, es la probabilidad de transmisión por contacto. El número potencial de contactos infecciosos  $c$  se relaciona con el tamaño de la vecindad lo que determina la conectividad con otros nodos en la red, pero no es exactamente el número de vecinos. Para poder realizar la equivalencia entre ambas metodologías, el número de celdas adyacentes o nodos vecinos en el modelo de autómatas se multiplica por la densidad de celdas ocupadas en la grilla (que puede variar según el tamaño de grilla usado y la cantidad de individuos a simular). La combinación de la probabilidad  $\beta'$  con el número potencial de contactos  $c$  es lo que dará como resultado la tasa de infección de cada individuo infeccioso. Como el valor que se obtuvo a partir del ajuste de parámetros, el valor de  $\beta$  se obtiene a partir de la elección del tamaño de vecindad y del parámetro  $\beta'$  usado en el modelo clásico, utilizando la relación  $\beta = \beta'c$ . En la figura 3-7 se muestra el grafo de transición de estados del autómata que modela el estado epidémico.

El vector de estado inicial  $\mathcal{P}_0$  está compuesto por las probabilidades de cada estado de ser el inicial en un autómata dado. Definiendo como  $G_t$  a la cantidad total de celdas en la grilla y como  $S_i, E_i, I_i, A_i, R_i, D_i$  al número inicial de individuos en cada estado en la grilla (siendo su suma igual a  $G_t$  y no a la población total, ya que  $D_i$  incluye las celdas vacías), el vector queda definido como se muestra en la Tabla (3-3). En nuestro caso sólo se

**Tabla 3-3.:** Vector de estado inicial

$S_i/G_t$	$E_i/G_t$	$I_i/G_t$	$A_i/G_t$
-----------	-----------	-----------	-----------

consideran probabilidades para los estados  $S$ ,  $E$ ,  $I$  y  $D$ , derivadas de los valores iniciales de individuos en el modelo clásico, y las celdas que quedan vacías en consecuencia. Todos los otros estados no cuentan con individuos iniciales.

La función  $\mathcal{L}_t : \mathcal{V}_t \rightarrow \mathcal{V}_t \times \mathcal{X}'^*$  establece el conjunto de reglas que asignan a cada elemento del conjunto  $\mathcal{V}_t$ , es decir, cada nodo en el espacio de contactos, una lista de elementos (enlaces) de un conjunto  $\mathcal{X}'$ . Es decir, esta regla asigna la topología  $\mathcal{T}$  a la red en el tiempo  $t$ . Para una grilla cuadrada y una vecindad de tipo Moore, el numero de enlaces que se le asignan a un elemento determinado en la red esta dado por  $\mathcal{X}'t = (2r + 1)^2 - 1$ , siendo  $r$  el radio de la vecindad utilizado. Esto quiere decir que el radio utilizado determina el grado de cada nodo en la red de contactos. De esta forma si se quieren modelar contactos locales, el radio no solo determina el grado del nodo, sino que también establece que el nodo se conecta con todos los nodos dentro de ese radio. Si se quieren modelar contactos homogéneos en la red, como en los modelos poblacionales, el radio limita el número de contactos de un nodo pero no su espacio dentro de la red de contactos.

A su vez se establece que la conexión  $\mathcal{R}$  es unidireccional y no nula desde los nodos en estado  $I$  o  $A$  hacia los nodos en estado  $S$  que estén dentro del radio de influencia  $r$ . La conexión es isótropa e igual en cualquier punto de la vecindad, y aporta una entrada a cada nodo en estado  $S$  que consiste en el valor de  $\lambda$  que tenga el nodo en estado  $I$  o  $A$ , utilizado para efectuar la transición probabilística hacia el estado  $E$ . Los nodos en estado  $S$  que estén incluidas en varias vecindades en un instante de tiempo  $t$  tendrán tantas oportunidades para cambiar de estado como la cantidad de vecindades en las que estén incluidas. La red de contactos queda definida entonces por  $\mathcal{R} = \mathcal{G}(\mathcal{T}, \mathcal{C})$ .

El comportamiento individual de los autómatas se modela a través de un **mapa cognitivo difuso** como se introduce en el capítulo 2. El valor del nodo  $C_{10}$  de este mapa limita el numero de contactos dentro del espacio de contactos, lo cual afecta la tasa de contactos efectivos. Los valores de los pesos  $w_{ij}$  de la matriz  $W$  son los que se detallan en el Capítulo 2. Los valores de los conceptos de entrada  $C_1$  (radio de individuos infectados que tienen contacto con un individuo determinado) y  $C_2$  (radio de individuos recuperados que tienen contacto con un individuo determinado) son estimados en cada iteración como densidad de infectados sobre tamaño de vecindad ( $\nu$ ) y densidad de recuperados sobre tamaño de vecindad respectivamente. El concepto  $C_5$  (conocimiento de la situación epidemiológica local) se fija en 0,5 de acuerdo a las fases epidémicas establecidas por la OMS que distingue 6 fases. Estas fases van de 1 a 6 siendo  $C_5 = 0,5$  el valor correspondiente a una fase 4 correspondiente a la etapa de brotes epidemiológicos locales en

comunidades pequeñas. El resto de los conceptos son inicializados de forma aleatoria en el rango  $[0, 1]$  usando distribución uniforme de probabilidad. El valor de  $C_{10}$  cuantifica la percepción de la situación epidemiológica global por parte del individuo. En el trabajo de Lee y col. (2012), los autores utilizan una función  $\phi$  para calcular la probabilidad de infección según el valor de  $C_{10}$ . En el modelo implementado en este capítulo si el valor de  $C_{10}$  se encuentra dentro del intervalo  $[0,6, 0,9]$  entonces el individuo deja de moverse en el espacio de los contactos. Esta estrategia tiene sentido cuando el radio de la vecindad no solo limita el grado del nodo, sino también el área de influencia en la red. Cuando se simulan contactos homogéneos en la red, esta estrategia no tiene sentido. En este caso, el valor de  $C_{10}$  limita directamente el grado del nodo de la forma  $\lceil (1 - C_{10}) \nu \rceil$ .

---

**Algoritmo 2** Dinámica del autómata celular
 

---

$G = \mathcal{M}_{M \times N}$ {Grilla cuadrada de  $N \times N$  en donde evolucionan los autómatas}  
 $Nt = \text{IndividuosTotales}$   
 $N = \text{DiasTotales}$   
 $G = I$ {Configuración inicial de la red}  
**para**  $k = 1 \rightarrow N$  **hacer**  
     Extracción  
     Reemplazo  
**fin para**

---

El algoritmo 2 describe en términos generales el proceso epidemiológico. Para cada iteración  $k$  la dinámica temporal del modelo puede ser formalmente definida a través de un triplete  $\langle \mathcal{E}, \mathcal{R}, \mathcal{I} \rangle$ . Donde  $\mathcal{E}$  es el mecanismo de extracción que la sección de la red que se selecciona para ser actualizada. En nuestro caso, la sección de la red seleccionada corresponde a los nodos en estado  $I$  o  $A$  y los nodos dentro de la vecindad.  $\mathcal{R}$  toma la configuración del segmento seleccionado y devuelve una nueva configuración topológica.  $\mathcal{I}$  es la configuración inicial que toma la red de contactos.

En el Apéndice A se describe cada uno de los pasos del algoritmo 2 en forma mas detallada. En la siguiente sección se procede a la validación del modelo presentado.

En la figura 3-8 se puede tener una visión mas detallada sobre la dinámica del modelo, en la parte superior puede verse la grilla del autómata celular en diferentes instantes de tiempo  $t$ , en la parte del medio se puede ver de que manera va evolucionado la red de contactos booleana correspondiente a una sección determinada de la grilla, finalmente en la parte inferior de la figura puede observarse un modelo de crecimiento de la red de contactos, el mismo brinda una idea de como se expande la epidemia a partir de un nodo infectado a lo largo del tiempo  $t$ . La figura 3-8 ilustra la dinámica cuando el radio limita el grado del nodo y el área de influencia en el espacio de los contactos. Pro simplicidad se adopta un radio  $\nu = 1$  par este ejemplo, aunque no es el radio utilizado luego. Por otro lado puede verse de que forma el movimiento aleatorio de los autómatas en la grilla afecta

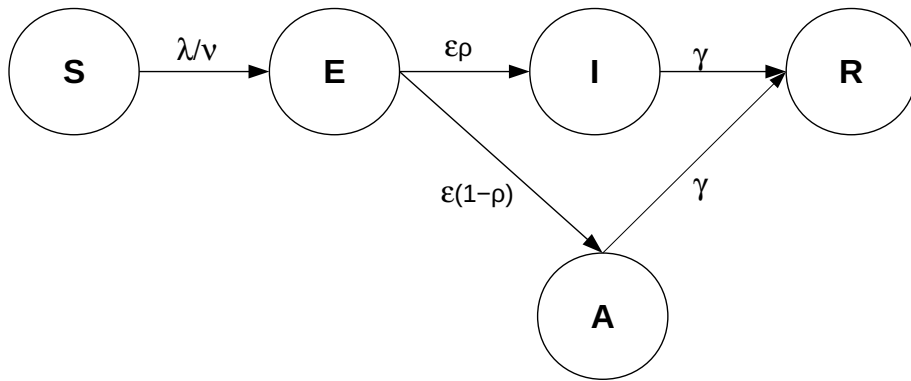


Figura 3-7.: Grafo de transición de estados del autómata que modela el estado epidémico.

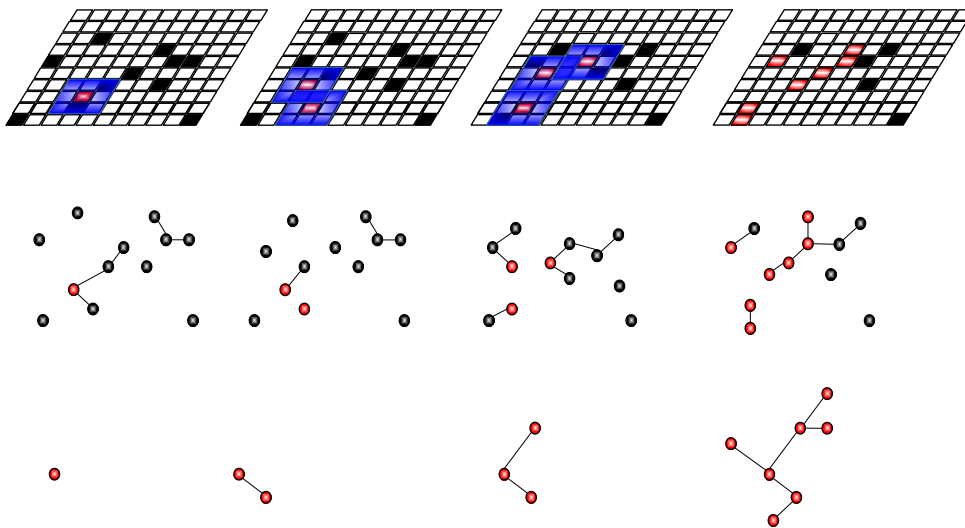


Figura 3-8.: Diferentes representaciones de la red de contactos como (a) grilla, (b) red booleana y (c) red de propagación

la topología de la red ya que debe tenerse en cuenta que la grilla determina el espacio de los contactos y no necesariamente guarda relación directa con el espacio físico.

### 3.4. Estimación de parámetros

Los parámetros del modelo fueron estimados siguiendo un procedimiento de dos etapas: (i) utilizando un método global de optimización estocástico, mas precisamente recocido simulado (Deb, Pratap, Agarwal, & Meyarivan, 2002) y luego (ii) algoritmos de optimización locales basados en gradiente (Byrd, Gilbert, & Nocedal, 2000). Este procedimiento nos permite explorar todo el espacio de parámetros buscando buenos candidatos (búsqueda global), que se utilizan para encontrar los mejores parámetros para el modelo a través de una búsqueda local. Los métodos de optimización estocástica proporcionan buenos puntos de partida para los métodos de optimización de basados en gradiente. La función objetivo utilizada fue el *error cuadrático normalizado* (NMSE):

$$NMSE = \sum_{k=1}^m \frac{\|m_I(k) - d_I(k)\|_2^2}{\|m_I(k)\|_2^2}, \quad (3-1)$$

donde  $m_I(k)$  es la dinámica temporal obtenida por el modelo (infectados) y  $d_I(k)$  son los infectados reportados en reales. En la tabla 3-4 se resumen los parámetros del modelo clásico donde  $\beta$  es la tasa de transmisión,  $\rho$  es la fracción de individuos en estado latente que pasan a estado infeccioso,  $\gamma$  es la tasa de recuperación,  $\alpha$  es la tasa de reportados (se asume que no todos los infectados son detectados por el sistema de salud),  $q$  es la tasa de infección de individuos asintomáticos,  $N_e$  es la población inicial de individuos expuestos y  $N_i$  es la población inicial de individuos infecciosos.

Debido a la gran dificultad que existe para disponer de series temporales que describan epidemias de gripe, se optó por utilizar los datos de la gripe española de 1918 registrada en la ciudad suiza de Ginebra que ha sido ampliamente difundida (Chowell y col., 2006) A la hora de intentar reproducir el modelo clásico (Chowell, Miller, & Viboud, 2008) con los parámetros minimizados, se presentan varios obstáculos y alternativas a emplear. En primer lugar, se presenta el problema de la escala a utilizar. Es evidente que a mayor tamaño de grilla y población empleada, más cerca estaremos de la hipótesis de población grande del modelo clásico (Chowell, Nishiura, & Bettencourt, 2007) (así como mayor será el costo computacional de la simulación). Como la extensión de la validez de esta hipótesis es justamente algo que se está poniendo a prueba al aplicar el modelo de autómatas, no es necesario ni deseable usar tamaños de grilla demasiado grandes a la hora de analizar el comportamiento temporal de la epidemia, pero sí lo es a los efectos de validar el modelo de autómatas usando como parámetro el modelo clásico.

Otra elección necesaria y que muestra efectos importantes en el comportamiento global es la del tamaño y forma de la vecindad. El uso de formas alternativas no muestra ningún cambio significativo, por lo que se elige la de Moore por razones de simplicidad de

programación y de cómputo. El tamaño de la vecindad, en cambio, determina el grado de influencia global de la heterogeneidad. Cuanto mayor sea el tamaño de vecindad usado, más se acercan los resultados a la hipótesis de homogeneidad espacial del modelo clásico. Esto se debe a que con un tamaño muy grande de vecindad los infecciosos pueden influir aun en zonas de baja densidad de infecciosos, suavizando el efecto de heterogeneidad. Si usásemos un tamaño de vecindad que abarque toda la grilla, no importaría la distribución espacial de los individuos; el resultado sería el equivalente a una distribución perfectamente homogénea porque cada infeccioso influiría en toda la grilla, sin importar su ubicación. El tamaño de vecindad también representa un límite para el valor de  $\beta$  usado. Como el valor de  $\beta$  ajustado en el modelo clásico es mayor a uno pero en cada celda la probabilidad de contagio no puede ser mayor a uno, la tasa de contagio dependerá de la cantidad de celdas con las que pueda interactuar el individuo infeccioso en cada unidad de tiempo. Por ejemplo, con una vecindad de Moore de radio  $r = 2$ , un individuo infeccioso no podrá contagiar a más de ocho individuos susceptibles en cada iteración. Para asegurar un límite superior apropiado a la hora de ajustar el modelo de autómatas, se eligió un radio de Moore de  $r = 3$ , que permite un máximo de 48 contagios (contactos) por unidad de tiempo.

**Tabla 3-4.:** Parámetros minimizados del modelo clásico

$\beta$	$\rho$	$\gamma$	$\alpha$	$q$	$N_e$	$N_i$
8,3	0,087	0,246	0,465	0	207	136

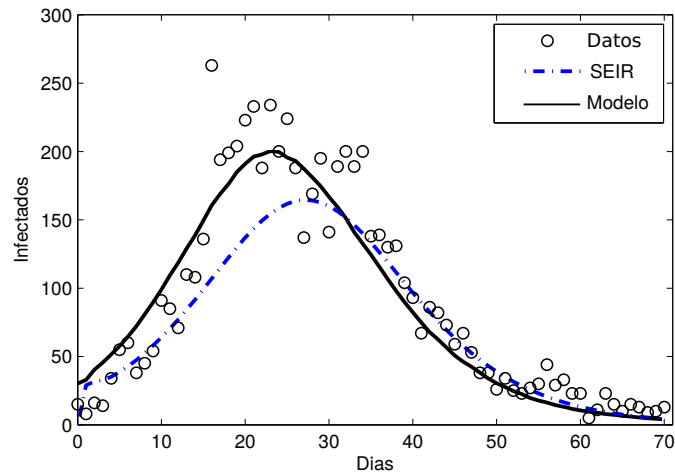
Con todas estas consideraciones, se efectuó primeramente una simulación que cumpla con las hipótesis del modelo clásico en el mayor grado posible. Para ello, se usó un tamaño de población total igual al de la población de Ginebra en 1918, uniformemente distribuida en la grilla, acercándose a la hipótesis de población grande. El tamaño de vecindad se agrandó al máximo, de modo que cada infeccioso interactúe en forma homogénea con todos, limitando solo el grado del nodo a 48, correspondiente a  $\nu = 3$  (Figura 3-9).

Vemos cómo el modelo de autómatas coincide muy bien con el modelo clásico, verificando su aplicabilidad para poblaciones grandes y con distribución espacial homogénea.

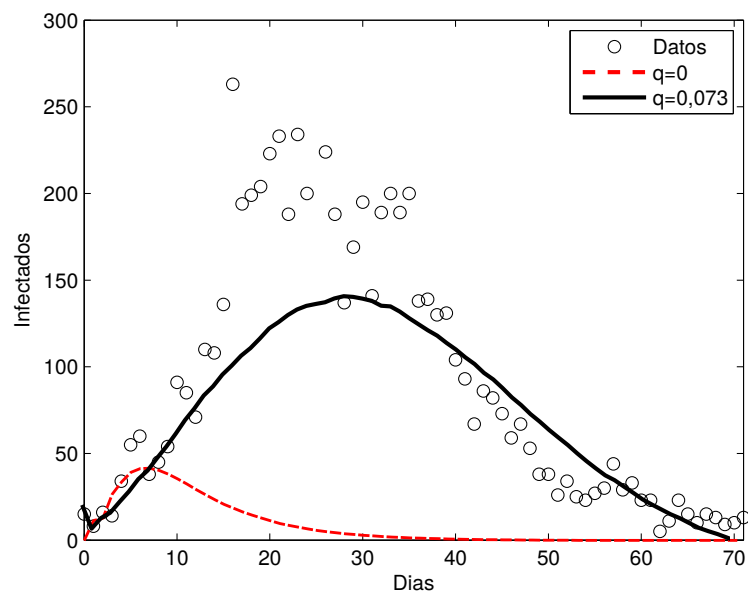
Para acercarse a condiciones más realistas, se realizaron luego simulaciones con el tamaño de población real y vecindades pequeñas. La distribución inicial de individuos se mantuvo uniforme. Se evidencia entonces un primer problema. Usando un radio de vecindad  $r = 3$  se obtienen evoluciones de reportados como la de la Figura 3-10, que no se relacionan con los datos reales.

Usando tamaños de vecindad mayores no se obtienen mejores resultados a menos que sean muy grandes, del orden del tamaño de la grilla (que se encontró que coincide muy bien con el modelo clásico, pero no está acorde a la realidad), y el aplicar tamaños mayores de población tampoco logra una modificación del comportamiento. El problema es que el ajuste de parámetros en el modelo clásico funcionará para cualquier conjunto de





**Figura 3-9.:** Evolución de la epidemia para una población homogénea, promedio de 500 realizaciones



**Figura 3-10.:** Evolución de la epidemia para vecindad de Moore, promedio de 500 realizaciones.

valores que minimice el error, sin importar si éstos tienen significado físico o son factibles de ocurrir y, por supuesto, no dicen nada de su comportamiento bajo otras condiciones, en este caso, una población discreta, donde la homogeneidad total no es posible. En el juego de parámetros usado en este caso, el parámetro  $q$  que afecta a la tasa de infección al pasar al estado de asintomáticos es igual a cero, lo que implica que estos individuos no infectan en absoluto. Sumado a la baja proporción de expuestos que pasan a ser infecciosos sintomáticos (menor a la décima parte), el efecto en el modelo de autómatas es el de un rápido descenso de la tasa efectiva de infección, porque cada individuo infeccioso queda, al cabo de un corto tiempo, rodeado de una gran proporción de individuos asintomáticos, que no pueden contagiar a nadie y a los cuales no puede contagiar. Puede intentarse resolver el problema ajustando la tasa de infección, pero no se obtienen mejoras porque el efecto de reducción de infectividad persiste al mantenerse  $q$  en el valor ajustado originalmente. En cambio, con un ajuste de  $q$  y  $\alpha$ , el comportamiento vuelve a ser el esperado. En la Figura 3-10 puede verse el resultado promedio de 500 realizaciones, utilizando  $q = 0,073$  y  $\alpha = 0,25$ . El ajuste de  $\alpha$  fue necesario para lograr una mejor correspondencia con los datos. De esta manera se logra una mayor homogeneidad usando un valor de  $q$  ubicado dentro del rango esperable. No obstante, queda evidenciada la discrepancia entre el modelo clásico y el de autómatas al relajar la hipótesis de homogeneidad. Un juego de parámetros que se ajusta muy bien a los datos en el modelo clásico se desvía totalmente cuando se considera una situación más realista. El ajuste del modelo clásico a los datos puede entonces resultar en un juego de parámetros que no se acerque al real aunque parezca ser el más adecuado.

### 3.5. Validación del modelo

La validación del modelo es uno de los pasos más importantes en el proceso de construcción del mismo. Se pueden utilizar dos metodologías para llevar adelante este proceso. Por un lado, los métodos gráficos ilustran una amplia gama de aspectos complejos de la relación entre el modelo y los datos observados. Por otro lado, los métodos numéricos ilustran de forma precisa aspectos particulares de la relación entre el modelo y los datos observados. Los métodos numéricos para la validación de un modelo revelan un número único o resultado que representa el grado de confiabilidad del mismo.

#### 3.5.1. Validación visual

Probar las hipótesis subyacentes de un modelo ayuda a asegurar la validez de las mismas. Debido a que la validez del modelo final está ligada de forma directa a la validez de las suposiciones subyacentes, se hace evidente la necesidad de que todos y cada uno de los supuestos sean evaluados.

Con el fin de comprobar la capacidad de nuestro modelo de ajustarse a dinámicas epidemiológicas reales, se muestran los resultados del ajuste con y sin comportamiento social. La Figura 3-11 muestra el ajuste del modelo sin el comportamiento social contra el ajuste del modelo *SEIR* clásico. En la misma figura también puede observarse el ajuste del modelo cuando se contempla comportamiento social de los individuos. Ambos modelos logran capturar la dinámica real muy bien como puede apreciarse en la figura, las aproximaciones su muy similares a los datos reales.

Los gráficos residuales se pueden utilizar para evaluar la calidad de un modelo. Los diferentes tipos de gráficos de residuos se pueden utilizar para comprobar la validez de las hipótesis del modelo y proporcionar información sobre cómo mejorarlo. En la Figura 3-12a se pueden observar los residuos correspondientes al modelo *SEIR*, el modelo basado en redes de contacto sin comportamiento social y el modelo basado en redes de contacto con comportamiento social. En la Figura 3-12b se observan las distribuciones de los errores para cada modelo respectivamente. Puede apreciarse que el error cometido por el modelo, cuando se lo compara con los datos reales es significativamente despreciables.

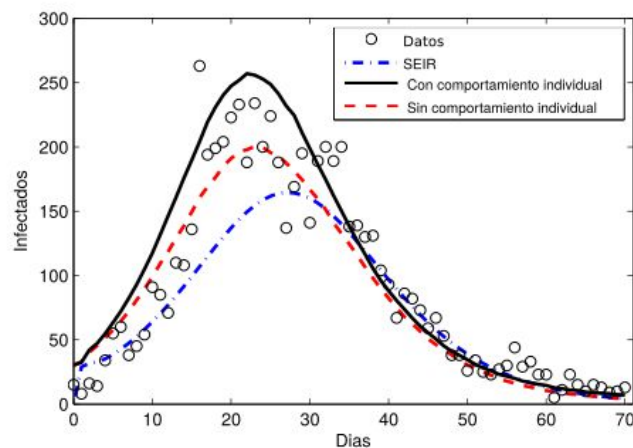
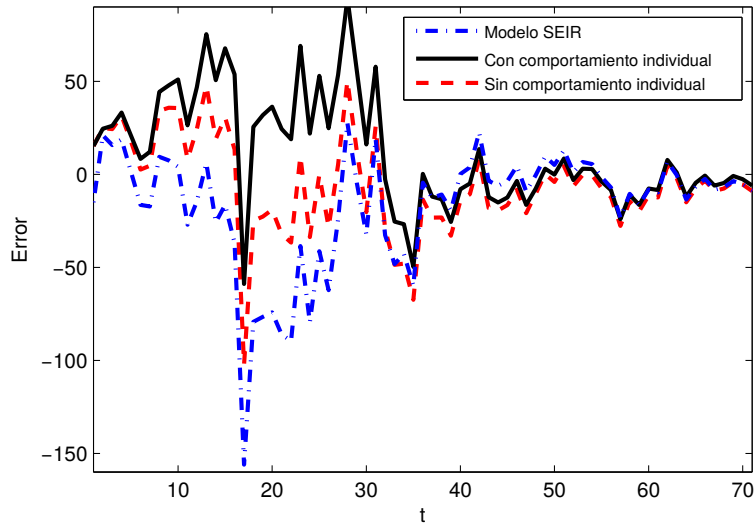


Figura 3-11.: Respuestas promedio del modelo con y sin conducta social

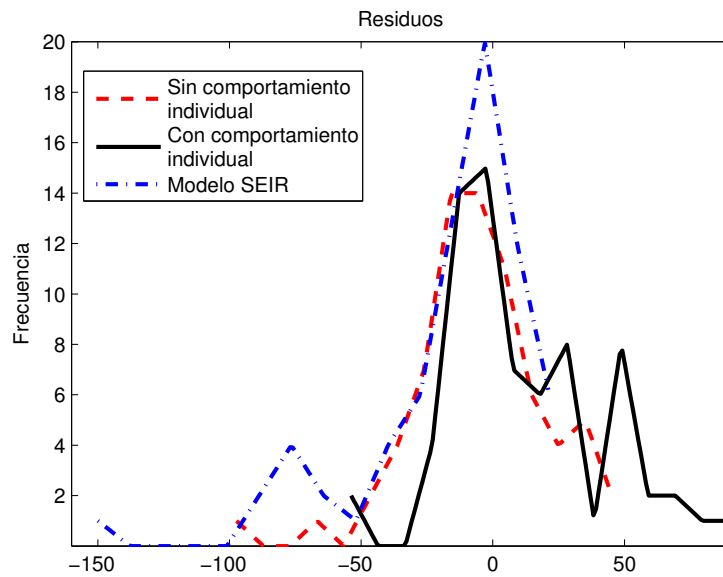
Tabla 3-5.: Estadísticos del error

Modelo	$\bar{X}$	$M$	$std$	$S$	Curtosis
SEIR	-19,57	-9,15	54,75	0	1,78
Con FCM	18	-5,08	45,87	0	1,75
Sin FCM	-5,8	-3,13	45,5	0	1,76

En la Tabla 3-5 se cuantifican las principales medidas estadísticas que caracterizan a la distribución del error para cada uno de los modelos, donde  $\bar{X}$  es la media,  $M$  es la moda,  $std$  es el desvíos estándar y  $S$  es el sesgo de la distribución. El sesgo  $S$  se calcula a



(a)



(b)

**Figura 3-12.:** (a) Gráfico de residuos (b) Distribución del error

través del segundo *coeficiente de Pearson* de la forma  $S = 3(\bar{X} - Md)/std$ , donde  $Md$  es la mediana. El coeficiente de sesgo da una idea de la simetría de la distribución, si  $S > 0$  la simetría de la distribución es positiva, si  $S < 0$  la simetría es negativa y si  $S = 0$  la distribución es simétrica. Esto quiere decir que mientras mas cerca de 0 esté el valor de  $S$ , mejor es la aproximación del modelo. En ambos casos, tanto el modelo con comportamiento social (con FCM) como el sin comportamiento social (sin FCM) el valor del sesgo es muy cercana a 0, con una leve asimetría negativa para el segundo caso. En tanto a la media y la mediana, puede observarse que  $\bar{X}$  y  $M$  son considerablemente menores para los dos modelos basados en redes de contactos.

### 3.5.2. Validación numérica

El modelo se validó numéricamente mediante el criterio de información de Akaike (*AIC*), que proporciona una medida de la calidad del modelo teniendo en cuenta la precisión y la complejidad del mismo. Este criterio es ampliamente utilizado para medir la calidad de los modelos (Symonds & Moussalli, 2011). Los modelos que tienen un valor de coeficiente *AIC* dentro del rango 1 – 2 soportan de forma consistente la variación estructural en los datos. Los modelos que tienen su valor dentro del rango 3 – 7 soportan de manera considerable la variación estructural en los datos. Finalmente, aquellos modelos que tienen sus coeficiente *AIC* > 10 no explican algunas variaciones estructurales importantes en los datos. El índice se calcula de acuerdo a la siguiente ecuación:

$$AIC = \log(\det(\frac{1}{m} \sum_1^m \varepsilon(t, \Theta)(\varepsilon(t, \Theta))^T)) + \frac{2n}{m}, \quad (3-2)$$

donde:  $\Theta$  es el conjunto de  $n$  parámetros estimados,  $m$  es el número de muestras y  $\varepsilon(t, \Theta)$  es la medida de error.

La razón por la que se optó por éste método y no por un método bayesiano, como el Criterio de Información bayesiano (*BIC*), subyace en el hecho de que el índice obtenido mediante *AIC* es una estimación de una constante más la distancia relativa entre la verdadera función de verosimilitud desconocida de los datos y la función de probabilidad ajustada del modelo, de modo que un *AIC* inferior representa un modelo mas cercano a la realidad. El índice *BIC*, por otro lado, es una estimación la probabilidad a posteriori de un modelo considerado verdadero, bajo una cierta configuración bayesiana, de modo que un *BIC* más bajo significa que el modelo se considera es mas probable de ser verdadero. El criterio utilizado resulta en un coeficiente *AIC* = 6,3 para el modelo que no tiene en cuenta el comportamiento social de los individuos y un *AIC* = 6,1 cuando el modelo contempla el comportamiento social de los individuos. Para el modelo clásico, usando el mismo criterio el *AIC* = 7,5. En los dos casos propuestos con el modelo basado en redes de contacto el índice está dentro del rango satisfactorio. El modelo se comporta de manera robusta ante la variación estructural de los datos.

El criterio de información de Akaike es equivalente a una validación cruzada *leave-one-out* en modelos de datos longitudinales, por ejemplo, selección de ancho de banda (Fang, 2011). Esta técnica es utilizada para evaluar los resultados de análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y los de prueba. Consiste en calcular la media aritmética obtenida de las medidas de evaluación sobre diferentes particiones. Se utiliza en entornos donde el objetivo principal es la predicción y se quiere estimar el grado de precisión del modelo implementado.

Teniendo esto en cuenta hemos evaluado la significación estadística de estos resultados mediante el cálculo de la probabilidad de que el error en la aproximación para los modelos implementados en este capítulo es mejor que el error del modelo clásico. Para realizar esta prueba asumimos la independencia estadística de los errores de ajuste para diferentes data sets y nos aproximamos a los errores de *distribución binomial* por medio de una distribución gaussiana. Esto es posible porque tenemos un número suficientemente alto de registros para cada experimento (1000 experimentos por data set).

Se dividió el conjunto de datos en tres subconjuntos. En primer lugar se entrenaron los modelos basados en redes de contactos y el modelo clásico con el mismo data set pero con  $n = 10$  datos menos que fueron seleccionados aleatoriamente siguiendo una distribución de probabilidad uniforme. Con el data set resultante se entrenó cada modelo con el propósito de obtener 7 generaciones de parámetros. Luego, se realizaron 1000 corridas por cada generación de parámetros obtenida en el paso anterior con el propósito de obtener una buena aproximación de la respuesta promedio del modelo basado en redes de contacto. Finalmente el error promedio de cada se calcula teniendo en cuenta la respuesta promedio del modelo utilizando los datos que fueron extraídos del data set original. Este proceso se llevó luego a cabo para  $n = 20$  y  $n = 30$ . De esta forma para cada conjunto de parámetros de ajuste con diferentes huecos en el data set tenemos que la hipótesis propuesta es que  $P(Error_{modelo} < Error_{SEIR}) > p$ . En el apéndice C se muestra la probabilidad de que el modelo basado en redes de contacto tenga una mejor performance que el modelo clásico.

En las Tablas (3-6) y (3-7) se resumen los resultados del test de insignificancia estadística para los diferentes conjuntos de validación para el modelo sin y con comportamiento social respectivamente. Puede verse que los modelos implementados son mejores que el modelo clásico si se observa el error de aproximación. El intervalo de confianza del error, en general, esta por encima del 90%. Puede observarse una leve disminución de la probabilidad  $P$  en para el caso de cuando se extraen 20 datos. Este comportamiento no es el esperado y requiere de un análisis más detallado que se realizará en el futuro.

**Tabla 3-6.:** Significancia de los errores para el modelo sin comportamiento social

10 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Redes de contacto SEIR	8,09	0,919	0,0044	4,91	99,96
	10,04	0,896	0,005		
20 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Redes de contacto SEIR	5,01	0,95	0,0035	2,02	94,15
	5,75	0,943	0,0038		
30 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Redes de contacto SEIR	4,9	0,951	0,0035	2,67	97,23
	5,88	0,941	0,0038		

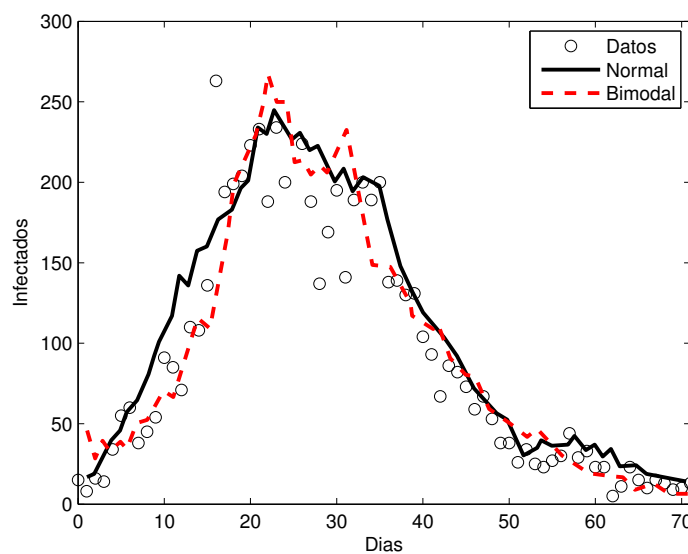
**Tabla 3-7.:** Significancia de los errores para el modelo con comportamiento social

10 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Redes de contacto SEIR	8,05	0,92	0,0044	5,01	99,9
	10,04	0,896	0,005		
20 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Redes de contacto SEIR	5,18	0,948	0,0036	1,52	88,4
	5,75	0,943	0,0038		
30 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Redes de contacto SEIR	5,11	0,949	0,0036	2,08	94,15
	5,88	0,941	0,0038		

### 3.6. Otros Resultados

A continuación, presentamos los resultados de diferentes realizaciones realizadas con el modelo descrito en esta sección con diferentes configuraciones y estudiando diferentes hipótesis (López, Burguener, & Giovanini, 2014).

#### Heterogeneidad y superspreading



**Figura 3-13.:** Evolución de la epidemia para una distribución normal y bimodal de la tasa de infección. Resultados de una realización por caso.

El uso de una tasa constante de infección no tiene en cuenta el rol de la variación individual en el proceso de infección. Esta variación puede ser de gran influencia cuando la distribución espacial tampoco es uniforme y cuando hay una presencia importante de *superspreaders*. Se define como *superspreaders* a aquellos individuos dentro de la población que son capaces de infectar a un número mayor de individuos susceptibles que el promedio, en otras palabras, si  $R_0$  es el número reproductivo de la epidemia, un *superspreader* posee un número reproductivo mayor que el promedio (Galvani & May, 2005a; Lloyd-Smith y col., 2005). De esta forma podemos definir el número reproductivo individual,  $\nu$ , como el número de casos secundarios esperados por un infectado en particular. Este parámetro puede ser descrito por medio de funciones de densidad de probabilidad continuas con una o varias medias según sea el caso de la función utilizada. Con un grado de infectividad distribuido continuamente, se logra agregar una mayor heterogeneidad al modelo. Estas variaciones en el número reproductivo pueden deberse a factores tales como:



- Individuos infectados y no diagnosticados.
- Altas tasas de contacto entre individuos.
- Alta carga viral en algunos individuos infectados.

En nuestro caso, la variación de  $R_0$  se logra únicamente a través de la variación de la tasa de infección. En un caso se utilizó una distribución de probabilidad normal para generar una grilla de individuos infectados con valores variables, dicha distribución de probabilidad posee media  $\beta = 8,3$  y varianza  $\sigma = 0,2$ . En el otro se usó una distribución bimodal para generar una grilla de individuos infectados, dicha distribución posee dos modas:  $\beta_I = 8,3$  y  $\beta_S = 8,7$  correspondientes a la probabilidad de infección por contacto con individuos infectados y superspreaders respectivamente.

**Tabla 3-8.:** Estadísticos del error

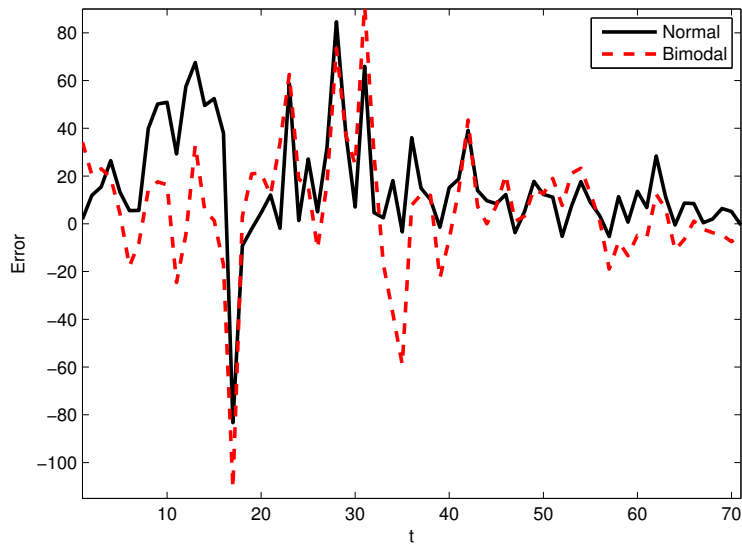
Modelo	$\bar{X}$	$M$	$std$	$S$	Curtosis
Normal	6,63	-77,69	50,05	0	1,76
Bimodal	-9,47	-103,92	60,33	0	1,75

En la Tabla (3-8) se describen los principales estadísticos de la distribución del error para ambos casos. Puede apreciarse que sin bien ambas distribuciones brindan un sesgo muy cercano al ideal, ideal en el caso de la distribución normal, la distribución bimodal da valores considerablemente mas bajos para la media y la mediana del error.

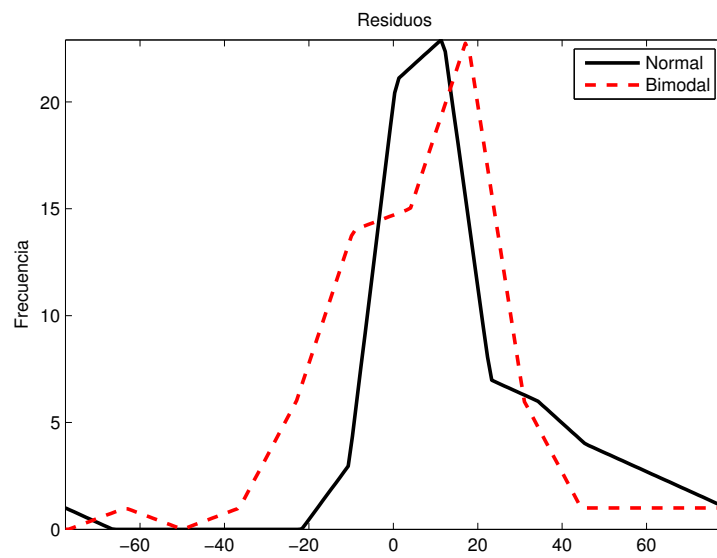
Si bien no se considera expresamente, están presentes también los individuos asintomáticos, que presentan una infectividad muy baja pero influyen en la dinámica de la epidemia ya que representan un porcentaje importante de la población. En este caso la tasa de infectividad de los individuos asintomáticos se mantiene en un valor fijo, pero podría también considerarse como una tercera población heterogénea, modelando el conjunto como una distribución bimodal, con modas  $\beta_I$  y  $\beta_S$ , correspondientes a los individuos infectivos y superspreaders, respectivamente. Tanto para la distribución normal como para la bimodal se ajustaron los parámetros de promedio (o modas) y varianza. Las siguientes figuras muestran la evolución en el tiempo de los individuos infectados por el patógeno. El primer caso corresponde al de la distribución normal y el segundo caso corresponde a la distribución bimodal. Puede observarse cómo al agregar mayor heterogeneidad al modelo los resultados obtenidos se ajustan más al comportamiento de los datos reales.

El primer caso, correspondiente a la distribución de probabilidad normal con media  $\beta = 8,3$ , se ejemplifica en la Figura 3-13. En la misma figura se puede ver el segundo caso, correspondiente a la distribución de probabilidad bimodal.

En este contexto, los patrones que aparecen como pequeñas olas epidemiológicas dentro de un periodo temporal corto, se interpretan como rebrotes debidos al paso de una cierta cantidad de superspreaders al estado infeccioso. En la Figura 3-14a pueden observarse



(a)



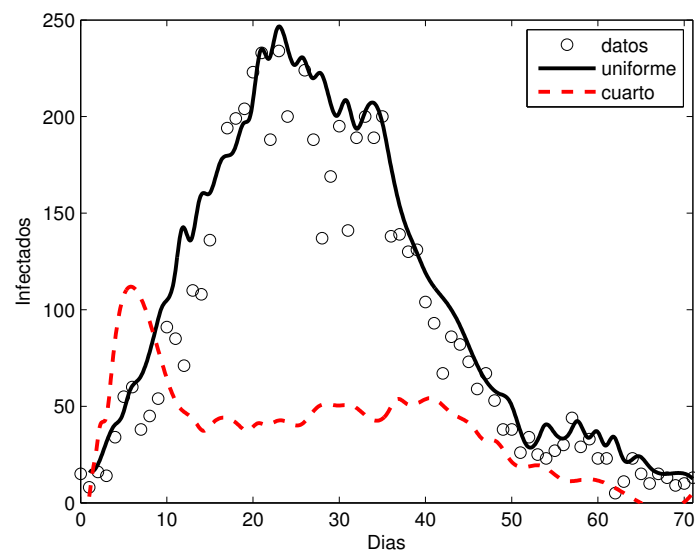
(b)

**Figura 3-14.:** (a)Residuos para el modelo con distribución normal y bimodal. (b) Distribución del error para el modelo normal y bimodal

los residuos para el modelo con distribución normal y los residuos para la distribución bimodal, mientras que en la Figura 3-14b se puede ver la distribución del error para cada caso. El análisis numérico de estos casos arrojó un índice  $AIC = 6,3$  cuando se modela la infectividad a través de una distribución normal y un índice  $AIC = 6$  cuando se modela la infectividad como una distribución bimodal que contempla la presencia de *superspreaders*

### **Heterogeneidad espacial**

Una suposición que se hace en los modelos epidemiológicos clásicos es que la distribución espacial de la población es homogénea; diferentes escenarios son planteados en donde se evalúa el impacto de la dinámica temporal para diferentes distribuciones iniciales de los individuos infecciosos. Se supone una distribución inicial homogénea de individuos susceptibles. En el caso de que una epidemia comience en una o en unas pocas regiones aisladas de la grilla, la homogeneidad poblacional no es algo que se logra de manera instantánea cuando se asume movimientos cortos por parte de los actores. Variando la distribución espacial inicial de los individuos en la grilla se obtienen dinámicas temporales y espaciales diferentes. Si consideramos a los individuos infecciosos distribuidos



**Figura 3-15.:** Dinámicas temporales para diferentes distribuciones espaciales iniciales de individuos infectados asumiendo movimiento aleatorio.

inicialmente en un cuarto de la grilla, la evolución de la población de infectados muestra picos de valor inferior que el observado con una distribución homogénea, del mismo modo, puede observarse como la epidemia se prolonga en el tiempo debido al confinamiento espacial de los individuos infectados, lo que lleva a que se incremente temporalmente la frecuencia de encuentros entre individuos sanos y enfermos. Es importante observar sin embargo, que a medida que los individuos infecciosos se mueven en la grilla, se producen

contactos efectivos y la epidemia se prolonga en el tiempo. En la Figura 3-15 pueden observarse las dinámicas temporales para diferentes distribuciones iniciales de infectados. Para estos experimentos se asumió una tasa de contagio  $\beta = 8,3$  constante.

### 3.7. Conclusiones

Los modelos basados en autómatas son especialmente útiles cuando se intenta modelar un sistema como una colección masiva de objetos que en su totalidad afectan el comportamiento de todo el sistema. La implementación de diferentes distribuciones de probabilidad para la infectividad de los individuos resulta en un aumento en la heterogeneidad poblacional y en una dinámica mas cercana a la real. Cuando se limita el área de influencia de los individuos infectados el movimiento aleatorio contribuye a la distribución homogénea de la población en el espacio de los contactos, lo cual incrementa la probabilidad de contacto entre los individuos susceptibles e infecciosos y por lo tanto la aumenta la probabilidad de contagio, otro factor determinante en este aspecto es el radio de la vecindad. El tamaño de la vecindad contribuye a determinar el nivel de heterogeneidad global, si la vecindad resulta muy pequeña la epidemia merma muy rápido debido a la falta de contacto entre individuos infecciosos y susceptibles, por otro lado, si el tamaño de la vecindad es demasiado grande aumenta el nivel de homogeneidad resultando en un comportamiento similar al de un modelo compartimental clásico. La introducción de superspreaders resulta en un aumento de la estocasticidad en el sistema. Esto se debe a que el número de individuos infecciosos en el tiempo  $t$  no siguen un patrón de crecimiento y disminución normal produciendo un pequeños rebrotes a lo largo de la ola. La incorporación del comportamiento individual a través del FCM brinda un mejor ajuste del modelos como pudo observarse en los resultados expuestos anteriormente. El modelo con comportamiento individual captura mejor la dispersión de los datos en el pico de la epidemia como puede apreciarse en las curvas obtenidas para la dinámica temporal. Este comportamiento puede atribuirse a que a medida que el número de infectados aumenta al principio de la ola, los individuos tienden a limitar su movimiento en el espacio de los contactos y por ello puede verse que hacia el final de la ola, la dinámica que describe el modelos con FCM tiende a sub-estimar el número de infectados. Finalmente puede decirse que la importancia de este tipo de modelado, como ya se mencionó antes, radica en el hecho de que permite ver a un sistema complejo como una colección masiva de objetos simples. A su vez, esos objetos (individuos) pueden ser modelados como la agregación de tres bloques diferentes que pueden modelarse de diferentes formas. La heterogeneidad es una propiedad de todas las poblaciones y es de vital importancia entender su rol en fenómenos poblacionales como son las epidemias, ya que permite un mejor análisis del mismo.



## 4. Modelos espacialmente explícitos

### 4.1. Introducción

Los modelos epidemiológicos tradicionales representan la dinámica de las enfermedades infecciosas a través de un enfoque no espacial y basado en la población, donde el tiempo siempre ha sido representado de forma explícita. Estos modelos no abordan explícitamente los factores causales para el desarrollo de epidemias. En su lugar, se estiman directamente el tamaño de la población afectada. En el proceso de modelado, una población se divide en susceptibles, infectados, y recuperados. El desarrollo de una epidemia se representa como el cambio en el tamaño del segmento de la población infectada debido a variaciones en los tamaños de los otros segmentos. Estos cambios se supone que ser continua y la velocidad de cambio en un momento dado se expresa por las ecuaciones diferenciales parciales (Anderson, May, & Anderson, 1992). Este enfoque y sus derivados toman a la población como la base de modelado y proporcionan una base para la epidemiología moderna y una herramienta para evaluar la dinámica de las epidemias y la salud de la población. Los modelos basados en ecuaciones poseen una estructura simple y un número reducido de parámetros que permiten un fácil proceso de modelado. Mediante el ajuste de los parámetros, los modelos se pueden aproximar razonablemente los datos de salud observados. Durante más de medio siglo, estos modelos tradicionales han permanecido como soporte para la epidemiología. En la última década, sin embargo, estos modelos no espaciales han atraído cada vez más críticas. A pesar de las ventajas de estos para predecir la dinámica de salud a nivel de población, no producen resultados realistas y útiles para problemas complejos (Holmes, 1997; Cordero y col., 2004). Los modelos basados en autómatas celulares surgen para subsanar muchas de las falencias que poseen los modelos clásicos. Un autómata celular es un sistema dinámico en el cual tanto el tiempo como el espacio se modelan de manera discreta. Los mismos están conformados por un conjunto de unidades básicas denominadas células que son idénticas unas de otras. Estos modelos intentan aumentar la heterogeneidad de la población con el fin de producir resultados más realistas que los modelos epidemiológicos tradicionales. Una de las principales limitaciones de estos modelos es que las unidades permaneces inmóviles en la grilla no contemplando de esta forma el movimiento natural de los individuos de una población.

El modelado basado en agentes es una metodología que intenta superar las limitaciones de los autómatas celulares. Utiliza como base del modelado a entidades virtuales denomi-

nadas agentes. Tradicionalmente, los agentes se han empleado en la resolución de problemas en forma distribuida, donde cada agente tiene la capacidad para resolver una parte del problema pero sólo se puede llegar a la resolución global trabajando en forma conjunta. Más recientemente se ha comenzado a utilizar a los agentes para realizar simulaciones, permitiendo modelar individuos con interacciones y comportamientos complejos dentro de un entorno dinámico. En el presente capítulo se desarrollará una revisión teórica de los principales conceptos referidos a agentes y sistemas multi-agente con el objetivo de incorporar estos conocimientos y poder implementar dicha técnica de modelado en una sección posterior. El presente capítulo se introduce una nueva modificación al marco teórico sugerido en el capítulo 2. Se utiliza un enfoque espacialmente explícito para modelar el comportamiento social de los individuos. En primer lugar se realiza una revisión sobre los conceptos más importantes que hacen al modelado basado en agentes. Luego se introducen los aspectos más importantes sobre el comportamiento social de los individuos. Finalmente se describe el modelo y se lo ajusta a un caso real para evaluar su desempeño. Luego se flexibilizan las hipótesis de ajuste y en la sección conclusiones se exploran las posibilidades de modelar fenómenos más complejos utilizando este enfoque.

## 4.2. Marco teórico

### 4.2.1. Agentes

Diversas definiciones de agente pueden encontrarse en la literatura, pese a que no existe una definición consensuada, mantienen aspectos en común. Se puede definir como agentes a sistemas informáticos que actúan en forma autónoma en un ambiente (Eyob, Tshilidzi, & Conrad, 2005), entendiéndose por autonomía la capacidad poseer recursos y habilidades propios que le permitan desenvolverse sin intervención externa. En forma más acotada, se consideran agentes inteligentes a aquellos capaces de actuar de forma autónoma y flexible en un ambiente, entendiéndose por flexibles a agentes que son reactivos (mantienen una interacción constante con su ambiente y responden a los cambios que ocurren en éste), pro-activos (no siguen sólo reglas de estímulo-respuesta, sino que generan y tratan de alcanzar metas, toman la iniciativa) y sociales (se comunican e interactúan con otros agentes y en algunos casos cooperan para lograr metas que no pueden alcanzarse individualmente) (Epstein, 2006).

Para poder actuar sobre el ambiente, los agentes deben poder percibir de alguna forma su estado y poseer una función que relacione la percepción con la acción. Para ello poseen lo que se conoce como *sistema de percepción* que es la capacidad de un agente de observar su entorno se pone de manifiesto en la función *see*. La relación entre percepción y acción se plasma en la función *action*, que representa el proceso de toma de decisiones del agente (Epstein, 2006). La salida de la función *see* mapea estados del ambiente en *preceptos*:

$$see : E \rightarrow Per \quad (4-1)$$

La función *action* mapea secuencias de *preceptos* en acciones:

$$action : perceptos^* \rightarrow Acciones \quad (4-2)$$

Otro concepto importante que es necesario definir a la hora de hablar sobre agentes es el de *Agentes con estado*, estos son agentes que modifican su comportamiento según el estado en que se encuentren, en forma similar a los autómatas. Tienen una estructura de datos interna usada para almacenar información sobre el ambiente y su historia (Epstein, 2006). Se define a  $I$  como el conjunto de todos los estados internos del agente. La función *see* no se modifica, pero la función *action* ahora se define como un mapeo de estados internos en acciones.

$$action : \mathcal{X} \rightarrow Ac \quad (4-3)$$

La función *next* mapea un estado interno y una percepción en un estado interno:

$$next : \mathcal{X} \times Per \rightarrow I \quad (4-4)$$

El agente comienza en un estado  $\mathcal{X}_0 \in \mathcal{X}$ , observa el estado de su ambiente y genera un precepto,  $see(e)$ . El estado interno es actualizado usando la función  $next(i_o, see(e))$  y la acción seleccionada es  $action(next(\mathcal{X}_0, see(e)))$ . El ciclo se repite.

### 4.2.2. Ambientes

Todo agente se desenvuelve en un ambiente, que puede o no ser explícitamente creado para él. El rango de los ambientes de trabajo es muy grande. Es posible clasificar estos ambientes a través de diferentes características que determinan el tipo de interacción con que tendrán con los agentes (Epstein, 2006).

- **Accesible/No accesible:** un ambiente accesible es aquel en el cual el agente puede obtener información completa y actualizada sobre el estado del ambiente. La mayoría de los ambientes moderadamente complejos (incluyendo el mundo físico e Internet) son inaccesibles. Mientras más accesible sea el ambiente es más sencillo construir agentes que funcionen sobre él.
- **Determinista/Aleatorio:** un ambiente determinista es aquel en el cual el siguiente estado está totalmente determinado por el estado actual y la acción a ejecutar por el agente. No hay incertidumbre. El mundo físico es no determinista. Representa un gran reto para el diseño de agentes.



- **Episódico/No episódico:** considerando como episodio a un evento de percepción-acción, un ambiente episódico implica que los episodios siguientes no dependen de las acciones que ocurrieron en episodios previos (como ocurre en las cadenas de Markov). Este caso es más simple que el no episódico, porque el agente puede decidir qué acción tomar basándose sólo en el episodio actual, sin necesidad de reaccionar en base a las interacciones entre el episodio actual y los futuros.
- **Discreto/Continuo:** un ambiente es discreto si tiene un número fijo y finito de acciones y percepciones. De lo contrario, es continuo.
- **Estático/Dinámico:** el ambiente es dinámico si puede modificarse durante el tiempo en que el agente decide. De lo contrario es estático.
- **Competitivo/Cooperativo:** el ambiente es competitivo cuando el aumento de eficiencia de un agente reduce la eficiencia de los demás. Si es cooperativo, en cambio, el aumento de eficiencia de un agente aumenta la eficiencia del resto de los agentes.

El ambiente puede estar en cualquiera de los estados del conjunto finito  $E$  (*estados del ambiente*), compuesto por estados instantáneos y discretos  $e, e', \dots$ , y los agentes tienen un repertorio de acciones posibles  $Ac = \alpha, \alpha' \dots$  que transforman el estado del ambiente. Una corrida  $r$  de un agente en un ambiente es una secuencia de estados del ambiente alternados con acciones del agente.

Definiendo  $R$  como el conjunto de todas las posibles secuencias sobre  $E$  y  $Ac$ ,  $R^{Ac}$  es el subconjunto de  $R$  que termina con una acción y  $R^E$  el subconjunto de  $R$  que termina en un estado del ambiente.

Una función de transformación de estados representa el comportamiento del ambiente:

$$\tau = R^{Ac} \rightarrow p(E) \quad (4-5)$$

Si  $\tau(r) = \emptyset$ , no hay posibles estados sucesores de  $r$ . Se dice que la corrida ha terminado. Un ambiente ( $Env$ ) es entonces un triplete  $Env = \langle E, e_0, \tau \rangle$  (donde  $e_0$  es el estado inicial) y un agente puede considerarse una función que mapea corridas en acciones:

$$Ag = R^E \rightarrow Ac \quad (4-6)$$

Es decir que el agente decide la acción a tomar basándose en la historia del sistema.

### 4.2.3. Sistemas multi-agentes

Un sistema multi-agente puede definirse como una red de agentes que trabajan en conjunto para resolver un problema que escapa a las posibilidades de su capacidad individual (Yergens, Hiner, Denzinger, & Noseworthy, 2006; Armano & Martin, 2006). En general, un sistema multi-agente consta de los siguientes elementos:

- Un ambiente de trabajo;
- Un conjunto de objetos ubicados en el ambiente;
- Un conjunto de agentes (los objetos activos del sistema);
- Un conjunto de relaciones que comunica a los agentes;
- Un conjunto de operadores que hacen posible que los agentes; puedan percibir, producir, transformar o modificar objetos.

En el contexto de las simulaciones mediante agentes, un sistema multi-agente es un conjunto de agentes que se desenvuelven e interactúan en un ambiente determinado (Iantovics, 2010). Las simulaciones con sistemas multi-agente permiten simular en forma directa a los individuos, sus comportamientos y sus interacciones, así como otros elementos propios del sistema (Epstein, 2006). Están basadas en la idea de que es posible representar en forma computarizada el comportamiento de entidades que se encuentran activas en un determinado ambiente y que es posible obtener consecuencias válidas de las interacciones.

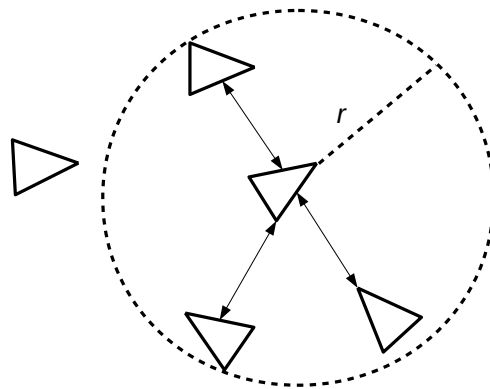
En un sistema multi-agente existen propiedades a nivel individual y a nivel sistema. Las propiedades a nivel individual corresponden a los comportamientos y características de cada agente. Las propiedades a nivel sistema son las propiedades globales del ambiente en el que viven los agentes, y emergen como resultado de la acción e interacción de los agentes entre sí y con el ambiente (Epstein, 2006).

Desde el punto de vista dinámico, los sistemas multi-agentes pueden analizarse como un conjunto de  $N$  agentes móviles con una dinámica simple dada por:

$$\begin{aligned} \dot{r}_i &= v_i, & r_i(0) &= r_i^0 \\ \dot{v}_i &= u_i, \end{aligned} \tag{4-7}$$

donde  $r_i = (x_i, y_i)^T$  es la posición,  $v_i = (\dot{x}_i, \dot{y}_i)^T$  es la velocidad y  $u_i = (u_i^x, u_i^y)^T$  es la entrada que determina el comportamiento del agente  $i$ . El ambiente puede modelarse como agentes cuyo comportamiento está predefinido, por ejemplo agentes inmóviles en el caso de que se tratase de obstáculos. Además, esta forma de modelar el comportamiento nos permite incorporar los efectos de la percepción y la comunicación entre agentes. El comportamiento individual de cada agente en relación a los otros (agentes y entorno) y a sus objetivos esta dado por  $w_i$ , suponiendo que cada agente interactúa sólo con aquellos agentes que se encuentran dentro de un radio de interacción  $r$  (Figura 4-1).

Entonces, el comportamiento del sistemas urge de la agregación de los comportamientos individuales de los agentes que lo componen (auto-organización). La idea de auto-organización descansa en la premisa de *racionalidad individual* (Lee, 2011; Cristiani, Priuli, & Tosin, 2015), doctrina en la que cada individuo persigue el mejor resultado posible para si mismo. De este modo, cada individuo trata de hacer lo mejor para si mismo sin tener



**Figura 4-1.:** Vecindad de un agente

en cuenta los efectos de sus acciones en el resto del grupo. En este sentido, el comportamiento del sistema resulta de la superposición de los comportamientos individuales, comportamiento emergente del sistema, el cual resulta más complejo que la suma de los comportamientos individuales (Helbing, Yu, & Rauhut, 2011).

El comportamiento emergente diferencia entre los niveles micro y macro en un proceso auto-organizado. Se considera que de las interacciones locales entre los componentes de una red (nivel micro) (Figura 4-1) emerge una estructura o patrón global (nivel macro). Por ejemplo, un cardumen de peces puede considerarse un proceso emergente. El nivel micro está constituido por los peces en movimiento y el nivel macro por el patrón espacial que observamos.

En relación al reduccionismo, cabe destacar que gran parte de la literatura sobre la emergencia está relacionada con las propiedades no lineales de los sistemas auto-organizados. En concreto, algunos tipos de redes, cuyos componentes interactúan de forma no lineal, resultan analíticamente intratables. Las ecuaciones diferenciales que rigen su comportamiento no se pueden resolver analíticamente y calcular el cambio de una variable en el sistema exige calcular el cambio simultáneo en las demás variables. Por tanto, cualquier transición de micro-estados requiere hacer referencia a todo el sistema, dando lugar a un holismo irreductible. Por este motivo, los sistemas complejos no están sujetos a una reducción localizacionista (Helbing y col., 2011) y sus propiedades suelen considerarse emergentes. Las simulaciones por ordenador (que hacen uso del cálculo numérico para realizar un estudio cualitativo del comportamiento) se utilizan para naturalizar estos sistemas y estudiarlos científicamente sin reducirlos a agregados de sus componentes.

### 4.3. Estado del arte

Epstein, Institution, on Social, y Dynamics (2002) simula una epidemia de tipo *SEIR* suponiendo una tasa de contagio constante en dos estados infecciosos posibles (uno de baja

y otro de alta tasa de contagio), se modelan dos ciudades de 400 habitantes cada una. Cada ciudad contiene 100 casas con dos adultos trabajadores y dos niños en cada una. En cada ciudad hay una escuela y un lugar de trabajo. Todos los niños acuden a la escuela en su ciudad. La mayoría de los adultos acuden al trabajo en su ciudad, mientras que una pequeña fracción trabaja en la ciudad vecina. Existe un solo hospital usado por ambas ciudades, con 5 trabajadores de cada ciudad. Cada día se divide en dos fases iguales. En la primera fase, correspondiente al día, los adultos y niños acuden al trabajo y la escuela, respectivamente. En la segunda fase, correspondiente a la noche, los miembros de cada familia interactúan en sus casas exclusivamente. Cada fase consta de muchas iteraciones. En cada iteración se efectúan una interacción al azar entre cada individuo y un vecino (colega de trabajo, compañero de escuela o miembro de la familia). La probabilidad de infección depende no sólo del estadio del agente infeccioso sino de la densidad de población a su alrededor. Luego de 20 horas en estado de infección altamente contagiosa, el individuo es hospitalizado. Los individuos recuperados retoman su rutina habitual, pero en estado inmunizado. Es importante destacar que, en este ejemplo, el espacio se modela de forma discreta y los agentes se ubican siempre en posiciones fijas en una grilla de forma similar a los autómatas celulares (por ejemplo, en las casas hay 4 posiciones posibles). Otro aspecto notorio en este ejemplo es la falta de transición entre el día y la noche. Durante el día están todos distribuidos entre el hospital, lugar de trabajo y escuela; a la noche vuelven a sus hogares y ciudades instantáneamente, en la primera iteración.

Por otro lado Teweldemedhin, Marwala, y Mueller (2004) implementa un modelo en el cual todos los elementos del sistema, incluido el ambiente son implementados como agentes. Un agente controlador inicia la simulación y supervisa la creación o borrado de los otros agentes, así como su evolución. Un agente estadístico recolecta toda la información obtenida de la simulación y efectúa cálculos con ella. Los agentes que simulan a los individuos se mueven por un espacio continuo e interactúan con otros eligiendo uno o más compañeros al azar entre sus vecinos. Se realiza una petición de interacción, que puede ser aceptada o rechazada, lo cual tiene sentido en este caso, al modelarse contactos sexuales o de otro tipo que puedan ocasionar una infección de HIV. Luego el agente puede decidir si infectarse o no, en base a la probabilidad de infección usada (con la posibilidad de ser un contacto con o sin protección). Los individuos pueden modificar su comportamiento luego de una interacción dada (por ejemplo, comenzar a tener contactos con protección solamente). En este caso, la tasa de infección también es constante, pero contemplando la variación para el caso de contacto con protección. El ambiente no tiene regiones especiales y los individuos se mueven en él en forma aleatoria. El tiempo es modelado en forma discreta.

Yergens y col. (2006), también implementa todos los elementos como agentes, pero en este caso se modelan individuos, ciudades y carreteras que las unen, en forma flexible y detallada. Se ejemplifica para el caso de gripe aviaria, pero el sistema puede adaptarse a otros casos. Los individuos tienen gran cantidad de atributos ajustables, como edad, se-

xo, estado de infección por la enfermedad bajo estudio, estado de HIV (que influye en el estado del individuo ante la enfermedad bajo estudio) o interacción con otros individuos. Las ciudades tienen entre sus atributos la población o el estado de conexión con otras ciudades (que puede variar ante una estrategia de intervención por aislamiento). El sistema utiliza una base de datos para cargar los atributos de las ciudades y su conectividad en base a datos reales. Se obtiene una representación en base a grafos, con los nodos representando a las ciudades y los arcos con pesos asociados representando a las carreteras. La distribución de edades y sexos al crear los agentes también se efectúa a partir de la base de datos. Los individuos interactúan entre sí en forma distinta antes y después de infectarse. Esto parte de la hipótesis de un cambio de hábitos del individuo al saber que está enfermo. Los individuos pueden pasar de una ciudad a otra y dispersar la enfermedad, también en forma ajustable.

En los tres trabajos descritos anteriormente se utiliza una tasa de infección constante, con ciertas variaciones ocasionales, pero que surgen de afectar en mayor o menor grado a esa tasa constante. El uso de una tasa de infección constante es una simplificación. En el modelado clásico con ecuaciones diferenciales, el uso de tasas variables puede dar como resultado ecuaciones muy complicadas para analizar, por lo que es natural el uso de esta simplificación. Cuando el modelado es en base a individuos, el variar la tasa en cada iteración y en forma independiente en cada individuo con algún criterio, es una tarea sencilla y que no supone un gran consumo de recursos computacionales.

Comúnmente, las condiciones necesarias para que las infecciones perduren en las poblaciones huéspedes han sido definidas a través de umbrales. En particular, el tamaño de la población o umbrales de densidad suelen utilizarse a menudo para especificar la abundancia huéspedes suficiente para mantener la transmisión ininterrumpida del agente infeccioso sin importaciones subsiguientes. Sin embargo, el mantenimiento de agentes infecciosos depende tanto de las tasas demográficas como del tamaño de la población, especialmente cuando la intensidad fluctúa lejos del equilibrio epidémico. Esto plantea retos particulares en los sistemas en donde la densidad poblacional o las migraciones estacionales causan la abundancia de huéspedes y las tasas demográficas relevantes fluctúan (Lloyd-Smith y col., 2005). La confianza en los umbrales de abundancia poblacional se dificulta aún más cuando el riesgo de enfermedad está influenciado por factores como la edad o el sexo, la territorialidad, o comportamiento de grupo, por lo que la definición y cuantificación de la abundancia de huéspedes se dificulta. En consecuencia, el análisis de la prevalencia en una sola población debe ser refinado para reconocer los mecanismos más allá de los umbrales poblacionales simples. Los efectos espaciales como la difusión de individuos (Davis, Trapman, Leirs, Begon, & Heesterbeek, 2008) o la estructura metapoblacional (Lambert y col., 2013) pueden ser importantes para la persistencia del agente infeccioso. Además, se necesitan estudios de modelos empíricos para entender cómo el mantenimiento está influenciado por diversas interacciones huésped / agente, incluyendo fenómenos tales como infecciones crónicas, excreción intermitente, o disminución de

la inmunidad.

De los modelos actuales, los métodos basados en número reproductivo han sido utilizados para explorar el mantenimiento en múltiples huéspedes en interfaces humano-animal (Radivojac y col., 2013) e identificar las contribuciones específicas de cada especie a la transmisión (Nishiura, Castillo-Chavez, Safan, & Chowell, 2009). Estos enfoques típicamente asumen un sistema en equilibrio endémico, aunque los métodos más recientes han relajado esta suposición (Streicker, Fenton, & Pedersen, 2013). Estas dinámicas que no se encuentran en equilibrio pueden ser cruciales para la persistencia, especialmente en sistemas con brotes variantes temporalmente. Este problema es particularmente pronunciado cuando la infección sólo puede determinarse después de la muerte, poniendo de relieve la importancia de evaluar el estado de la enfermedad a través de un muestreo no destructivo.

## 4.4. Modelo

Los sistemas complejos tienen muchos componentes que interactúan de una manera no lineal, estos sistemas no pueden ser representados y analizados con los métodos tradicionales de modelado. Una forma de poder explicar y conocer sus características es a través del modelado espacialmente explícito basado en individuos. Es un tipo de modelado descentralizado que utiliza agentes, en el que se pueden dividir a estos agentes por clases y si bien todos los agentes de una misma clase tienen las mismas reglas, se comportan de una manera independiente ya que también dependen de su entorno, por lo que al agregar más agentes se eleva el nivel de complejidad de la simulación. Estas simulaciones se basan en especificar el comportamiento de las estructuras más pequeñas pero muestran el comportamiento global más complejo. El modelo basado en agentes presenta varias ventajas con respecto a otras técnicas de modelado. El comportamiento de estos está determinado por un conjunto de reglas, si bien esas reglas son las mismas para cada agente del sistema, las condiciones locales hacen que surjan diferentes patrones de comportamiento a nivel global, al igual que en los sistemas reales. Por otro lado, al ser un sistema espacialmente explícito, esto permite definir claramente los límites del ambiente en el que se desenvuelven los agentes y concentrarse en las interacciones entre los mismos. El enfoque posee cuatro características que lo convierten en una alternativa atractiva a la hora de modelar una epidemia: (i) Los agentes son autónomos, es decir que no existen mecanismos externos que controlen el comportamiento de los mismos, el mismo es regido por las reglas internas. (ii) Tienen capacidad social, es decir que los agentes son capaces de interactuar unos con otros. (iii) Son reactivos, es decir que reaccionan ante estímulos externos provenientes del entorno o de otros agentes y actúan de acuerdo a dichos estímulos.

### **Definición del sistema**

Con el fin de incorporar la dinámica propuesta, se puede definir un agente con los si-

güentes atributos:

- **Estado:** susceptible, expuesto, infeccioso o recuperado. Sin efecto para el comportamiento.
- **Susceptibilidad:** nivel de predisposición del individuo a infectarse por otros.
- **Infectividad:** predisposición a infectar.
- **Posición:** coordenadas del agente en el ambiente donde se desenvuelve.
- **Radio de vecindad:** radio de interacción con otros agentes.
- **Velocidad y dirección:** definen el movimiento del agente. Pueden cambiar en cada iteración según el comportamiento elegido y el encuentro con otros agentes.
- **Comportamiento:** puede definirse un movimiento al azar o un movimiento hacia objetos pasivos que se encuentran en el ambiente (casas, lugares de trabajo, hospitales, escuelas).

Todos los valores iniciales son ajustables y pueden modificarse durante la simulación. Los agentes pasivos que constituyen el ambiente pueden colocarse en diferentes posiciones y cantidades y los agentes pueden dirigirse a uno u otro en diferentes momentos y según el comportamiento asignado a cada agente.

El *ambiente* es un espacio con la única función de servir de contenedor para ubicar los objetos y los agentes durante la simulación. La percepción de cada agente se efectúa con los agentes que se encuentren en cierto radio predefinido de distancia respecto de él, y con el objeto hacia el que deba dirigirse. Con el fin de facilitar la interacción, cada agente publica su posición. La posición de los objetos pasivos también se encuentra disponible para todos los agentes.

El comportamiento social de los agentes es modelado a través de auto-organización, es decir, definiendo el comportamiento de cada agente, luego el comportamiento del sistema emerge de la agregación de los comportamientos de todos los agentes. Una posible forma de definir el comportamiento a partir de las reglas de introducidas por Reynolds (Reynolds, 1987).

- Mantener a los agentes vecinos lo más próximo que se pueda;
- Evitar las colisiones de los agentes vecinos;

Estas reglas son conocidas como las reglas de cohesión, separación y alineamiento. El principal problema con la implementación de estas reglas es que pueden interpretarse de

una manera muy laxa. La interpretación de estas reglas fue descrita en trabajos posteriores de Reynolds (Reynolds, 1999, 2000). Considerando un grupo de  $N$  agentes cuyo comportamiento dinámico viene dado por

$$\begin{aligned} \dot{p}_i &= q_i, & p_i(0) &= p^0, i \in \mathcal{N} \\ \dot{q}_i &= (1 - C_{10})u_i, & q_i(0) &= q^0, \end{aligned} \quad (4-8)$$

donde  $u_i$  es la entrada que determina el comportamiento de cada agente. Puede observarse que dicha entrada es afectada de manera directa por la salida del FCM por medio del término  $(1 - C_{10})$ , limitando de esta forma el movimiento de los agentes. Los trabajos más recientes se han enfocado en conceptos Lagrangianos (Mogilner, Edelstein-Keshet, Bent, & Spiros, 2003; Cucker & Smale, 2005; Olfati-Saber, 2006), que se enfoca en la relación entre individuo. La entrada  $u_i$  se obtiene a partir del gradiente de una función de potencial que se genera automáticamente a partir de la superposición de los potenciales de agentes individuales. Por lo tanto, las acciones de cada agente no solo está determinada por la información local, sino que es también producto de la combinación de la información disponible en todo el sistema. De este modo el movimiento de un agente  $i$  es descrito por las ecuaciones (4-8). Los agentes móviles tienen la tendencia a conservar una distancia  $d > 0$  de sus vecinos dentro de un radio de interacción previamente definido conformando las estructuras definidas como lattices o enrejados. En un sistema sin obstáculos, cada agente recibe una entrada que puede ser resumida de la siguiente forma:

$$u_i = f_i^g + f_i^\gamma, \quad (4-9)$$

donde  $f_i^g = -\nabla_{q_i} V(q)$  es el término que determina la cohesión del grupo y  $f_i^\gamma$  es una función de feedback que determina el objetivo del grupo. El comportamiento del agente en un ambiente sin obstáculos puede definirse como:

$$u_i = \sum_{j \in N_i} \phi_\alpha(\|q_j - q_i\|) \vec{n}_{ij} + u_i^\gamma(q_i, p_i), \quad (4-10)$$

donde  $p_i$  es la componente de velocidad del agente y  $\vec{n}_{ij}$  es un vector a lo largo de la línea que une a  $q_i$  y  $q_j$  y esta dado por la siguiente expresión:

$$\vec{n}_{ij} = \frac{q_j - q_i}{\sqrt{1 + \epsilon \|q_j - q_i\|^2}}, \quad 0 < \epsilon < 1, \quad (4-11)$$

lo cual define el protocolo de comportamiento entre agentes. El término  $u_i^\gamma$  es la función de navegación definida como:

$$u_i^\gamma := f_i^\gamma(q_i, p_i, q_r, p_r) = -c_1(q_i - q_r) - c_2(p_i - p_r), \quad c_1, c_2 > 0, \quad (4-12)$$

el par  $(q_r, p_r) \in \mathbb{R}^2$  es el estado de un tipo de agente estático o dinámico al que se lo denomina agente  $\gamma$  y representa un objetivo del grupo. Si el agente  $\gamma$  es un agente dinámico,



su movimiento viene definido por el sistema de ecuaciones similar a (4-8)

$$\begin{aligned}\dot{q}_r &= p_r, \\ \dot{p}_r &= f_r(q_r, p_r),\end{aligned}\tag{4-13}$$

cuando el agente  $\gamma$  es estático, entonces tiene un estado estático ( $\dot{q}_r = 0$ ) que se conserva para todo instante de tiempo. La dinámica colectiva de un grupo de agentes cuyo comportamiento está dado por las reglas establecidas en (4-10) tiene la siguiente forma:

$$\begin{aligned}\dot{q} &= p, \\ \dot{p} &= -\nabla V(q) - \hat{L}(q)p + f_\gamma(q, p, q_r, p_r),\end{aligned}\tag{4-14}$$

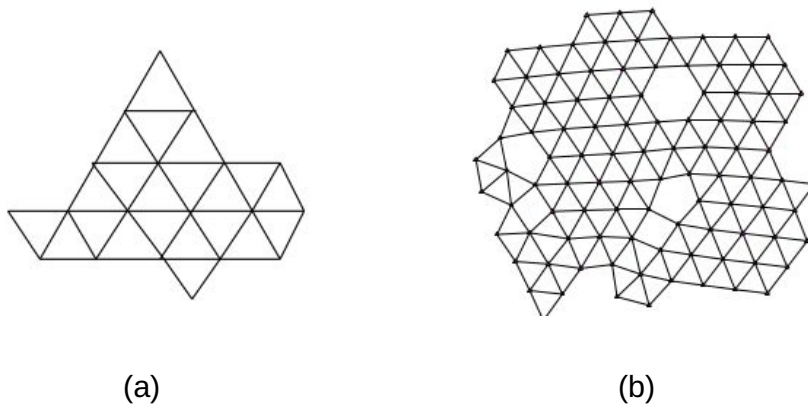
donde  $V(q)$  es el potencial colectivo definido en (4-20) y  $\hat{L}(q)$  es el laplaciano del grafo  $G(q)$  con matriz de adyacencias dependiente de la posición  $A(q) = [a_{ij}(q)]$

### Formación de grupos

Una de las características de este tipo de sistemas es la tendencia a formar grupos por parte de los individuos que lo componen como consecuencia de la dinámica que define el comportamiento de cada uno. Sea  $r > 0$  el rango de interacción entre dos agentes que determina el conjunto de vecinos espaciales de un agente  $i$  y se define como:

$$N_i = \{j \in \mathcal{V} : \|q_j - q_i\| < r\},\tag{4-15}$$

donde  $\| \cdot \|$  es la norma euclídea en  $\mathbb{R}^2$ . Dado un rango de interacción  $r > 0$ , un grafo inducido espacialmente  $G(q) = (\mathcal{V}, \mathcal{E}(q))$  puede ser descripto por  $\mathcal{V}$  y el conjunto de conexiones  $\mathcal{E}(q) = \{(i, j) \in \mathcal{V} \times \mathcal{V} : \|q_j - q_i\| < r, i \neq j\}$ , que claramente depende de  $q$ .



**Figura 4-2.:** (a)  $\alpha$ -lattice (b) cuasi  $\alpha$ -lattice

El grafo  $G(q)$  se define como red la estructura  $(G(q), q)$  se define como marco de la red. Si el rango de interacción para todos los agentes de la red es el mismo, la red  $G(q)$  se convierte en un grafo no dirigido Los enrejados (lattices) (ver Figura 4-2), se definen con

el objetivo de diseñar protocolos que permitan dinámicas de grupo. Esta idea motiva las siguientes restricciones algebraicas entre los agentes:

$$\|q_j - q_i\| = d, \forall j \in N_i(q), \quad (4-16)$$

Un  $\alpha$ -lattice es una configuración  $q$  que satisface las restricciones en (4-16), siendo  $d$  la escala y  $k = r/d$  y el radio de la  $\alpha$ -lattice. Debe tenerse en cuenta que no se requiere que la red inducida por una  $\alpha$ -lattice esté conectada. Una cuasi  $\alpha$ -lattice es una configuración  $q$  que satisface las restricciones:

$$-d \leq \|q_j - q_i\| - d \leq \delta, \forall j \in \mathcal{E}(q), \quad (4-17)$$

donde  $\delta \ll d$  es la medida de incertidumbre de conexión.

Para medir en que grado en que una configuración  $q$  difiere de una  $\alpha$ -lattice se usa la siguiente función de energía:

$$E(q) = \frac{1}{(|\mathcal{E}(q)| + 1)} \sum_{i=1}^n \sum_{j \in N_i} z^2 (\|q_j - q_i\| - d), \quad (4-18)$$

en donde  $z^2$  es el potencial de pares. Se pueden usar diferentes funciones como potenciales de acuerdo al comportamiento que se desea representar. Esta medida de energía puede ser visto como un potencial para un sistema de  $n$  partículas. Para las cuasi- $\alpha$ -lattices la energía de una configuración  $q$  puede ser calculada por medio de:

$$E(q) \leq \frac{|\mathcal{E}(q)|}{|\mathcal{E}(q) + 1|} \delta^2 \leq \varepsilon^2 d^2, \varepsilon \ll 1 \quad (4-19)$$

La función de potencial  $V(q)$  para un grupo de agentes es una función no negativa  $V : \mathbb{R}^{2n} \rightarrow \mathbb{R}_{\geq 0}$  con la propiedad de que cualquier solución del conjunto de restricciones algebraicas en (4-16) está estrechamente relacionada con un mínimo local de la función  $V(q)$  y vice versa. La forma en que estas dos cantidades están relacionadas se mide usando la energía de desviación que actúa como una función de distancia. Un potencial colectivo puede ser visto como una función de energía de desviación con un potencial escalar parrejo que tiene un corte finito. La función de potencial para una configuración  $q$  puede ser definida por:

$$V(q) = \frac{1}{2} \sum_i \sum_{j \neq i} \psi_\alpha(\|q_j - q_i\|), \quad (4-20)$$

donde  $\psi_\alpha$  determina el potencial de atracción-repulsión de la  $\alpha$ -lattice y se define como una regla de mapeo  $\psi(z) : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  siendo  $z$  la dimensión.

### **Movimiento en el entorno**

La idea principal subyacente en esta forma de flocking, que es la principal ventaja del enfoque es la de usar una aproximación basada en agentes para representar obstáculos.

Estos agentes son llamados agentes  $\beta$ , los mismos se definen como agentes cinemáticos inducidos por agentes  $\alpha$  cuando un agente  $\alpha$  se encuentra en las proximidades de un obstáculo. El método que utilizan los agentes para evadir obstáculos puede ser resumido en tres pasos:

1. Determinar los índices  $N_i^\beta$  del conjunto de obstáculos  $O_k$  que se encuentran en las proximidades de un agente  $\alpha$   $i$ ,
2. Crear un agente cinemático  $\beta$  en  $\hat{q}_{ik}$  en la vecindad de un obstáculo  $O_k$  donde  $\hat{q}_{ik}$  satisface:  $\hat{q}_{ik} = \text{armin}_{x \in O_k} \|x - q_i\|$ .
3. Agregar un término  $\psi_\beta(\|\hat{q}_{ik} - q_i\|)$  a la función de potencial para un grupo de agentes  $\alpha$  por la influencia de los agentes  $\beta$

Para lograr flocking en presencia de obstáculos se puede usar una función de potencial colectivo de la siguiente forma:

$$V(q) = c_1^\alpha V_\alpha(q) + c_2^\beta V_\beta(q) + c_3^\gamma V_\gamma(q), \quad (4-21)$$

donde  $c_1^\alpha$ ,  $c_2^\beta$ ,  $c_3^\gamma$ , son constantes positivas y los potenciales de interacción se definen como:

$$\begin{aligned} V_\alpha &= \sum_{i \in \mathcal{V}_\alpha} \sum_{j \in \mathcal{V}_\alpha \setminus \{i\}} \psi_\alpha(\|q_j - q_i\|), \\ V_\beta &= \sum_{i \in \mathcal{V}_\alpha} \sum_{k \in N_i^\beta} \psi_\beta(\|\hat{q}_{i,k} - q_i\|), \\ V_\gamma &= \sum_{i \in \mathcal{V}_\alpha} (\sqrt{1 + \|q_i - q_r\|^2} - 1), \end{aligned} \quad (4-22)$$

donde  $N_i^\beta$  es el conjunto de los agentes  $\beta$  vecinos,  $N_i^\beta = \{k \in \mathcal{V}_\beta : \|\hat{q}_{i,k} - q_i\| < r'\}$ , donde  $r$  es el rango de interacción,  $\mathcal{V}_\beta$  denota el conjunto de agentes  $\beta$  dentro del radio  $r'$  y  $\mathcal{V}_\alpha$  denota el conjunto de agentes  $\alpha$  dentro del radio  $r'$

De esta forma el algoritmo de flocking para un ambiente con obstáculos queda definido de la siguiente forma:

$$u_i = u_i^\alpha + u_i^\beta + u_i^\gamma, \quad (4-23)$$

donde  $u_i^\alpha$ , denota las interacciones entre agentes del grupo  $(\alpha, \alpha)$ ,  $u_i^\beta$  denota interacciones entre agentes y obstáculos  $(\alpha, \beta)$  y  $u_i^\gamma$  es el término de feedback de navegación. Los términos de la ecuación (4-23) pueden calcularse de la siguiente forma:

$$\begin{aligned} u_i^\alpha &= c_1^\alpha \sum_{j \in N_i^\alpha} \phi_\alpha(\|q_j - q_i\|) \vec{n}_{ij} + c_2^\alpha \sum_{j \in N_i^\alpha} a_{ij}(q)(p_j - p_i), \\ u_i^\beta &= c_1^\beta \sum_{k \in N_i^\beta} \phi_\beta(\|\hat{q}_{i,k} - q_i\|) \hat{n}_{ik} + c_2^\beta \sum_{j \in N_i^\beta} b_{ik}(q)(\hat{p}_{ik} - p_i), \\ u_i^\gamma &= -c_1^\gamma \sigma_1(q_i - q_r) - c_2^\gamma (p_i - p_r), \end{aligned} \quad (4-24)$$

donde  $\sigma_1 > 0$ .

Este enfoque representa al modelo como una colección de agentes de diferentes clases y la dinámica del sistema global puede verse como el producto de la interacción entre ellos. Otro factor fundamental que afecta al comportamiento de los individuos que componen el modelo es la percepción que ellos tienen de su entorno y de la situación epidemiológica global y local. Esto se modela por medio del *mapa cognitivo difuso*, el detalle de implementación del mapa puede verse en el capítulo 2. Dependiendo del estado del mapa, el agente evalúa si es conveniente o no moverse hacia los atractores o si es mejor alejarse de los mismos y tratar de evitar el agrupamiento con otros agentes. Este comportamiento puede simularse simplemente ponderando de manera diferente los parámetros del modelo que afectan la repulsión y atracción en del agente en el instante  $t$  por medio del sistema (4-20). El algoritmo (3) muestra la secuencia de pasos en para cada instante de  $t$ . En primer lugar, se setea los parámetros iniciales de la simulación (poblaciones iniciales, días de de simulación, entre otros.). Luego se produce el ciclo de infección a continuación se simula el movimiento de los agentes por medio de los métodos antes expuestos, finalmente se censa la población. En este diagrama  $k$  es el instante de tiempo actual y  $T$  es la duración en días de la simulación. Las rutinas implementadas están disponibles en el apéndice B. Recapitulando esta sección, en la Figura 4-3 se describe mediante un diagrama el modelo

---

#### Algoritmo 3 Dinámica del autómata celular

---

```

Inicializar{Se inicializan los parámetros de la simulación}
para  $k = 1 \rightarrow T$  hacer
    {Infección}
    {Actualización de estado infeccioso}
    {Movimiento de individuos}
    {Actualización de dinámicas temporales}
fin para

```

---

propuesto. Como puede observarse, nuevamente al igual que en el capítulo anterior, se modifica solo un bloque del esquema que corresponde al comportamiento social. El comportamiento individual y el estado epidémico son los mismos introducidos en capítulos previos. Teniendo en cuenta el marco propuesto en el capítulo 2 puede apreciarse que las modificaciones introducidas no afectan el *comportamiento individual*, que es modelado a través de *FCM*, el *comportamiento social* que es modelado como interacciones entre los individuos en el espacio explícito por medio de algoritmos de flocking y el *estado epidémico* que es modelado nuevamente mediante un autómata estocástico de estados finitos como en el capítulo anterior.

En las siguientes secciones se procede a mostrar los resultados del ajuste del modelo, la validación y algunos resultados cuando se relajan las hipótesis de los modelos clásicos.

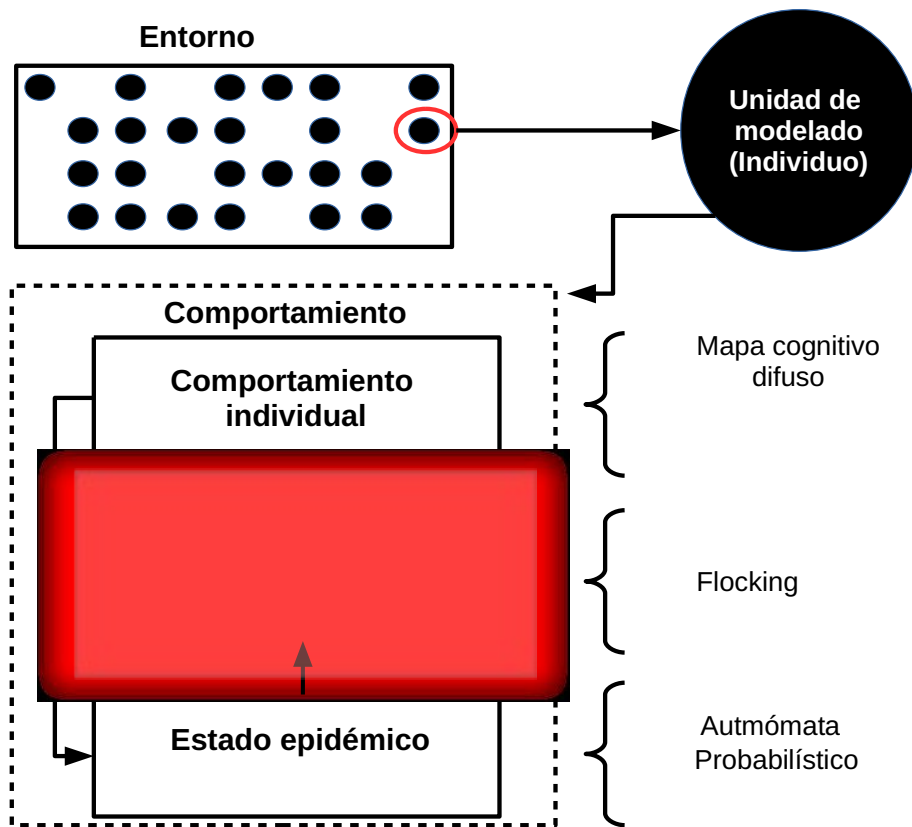


Figura 4-3.: Diagrama del modelo basado en individuos con espacio explícito

## 4.5. Estimación de parámetros

Para modelar la transición entre los diferentes estados epidemiológicos (estado epidémico), se utilizó un autómata de estados finitos con el propósito de describir el estado epidémico de cada agente cuyas probabilidades de transición fueron obtenidas a partir de los parámetros del modelo clásico (Tabla 4-1) Al igual que el modelo basado en redes de contacto, los parámetros del modelo fueron estimados siguiendo un procedimiento de dos etapas: (i) utilizando un método global de optimización estocástico, mas precisamente recocido simulado (Deb y col., 2002; Houck, Joines, & Kay, 1995) y luego (ii) algoritmos de optimización locales basados en gradiente (Powell, 1978; Byrd y col., 2000). Este procedimiento nos permite explorar todo el espacio de parámetros buscando buenos candidatos (búsqueda global), que se utilizan para encontrar los mejores parámetros para el modelo a través de una búsqueda local. Los métodos de optimización estocástica proporcionan buenos puntos de partida para los métodos de optimización de basados en gradiente. La función objetivo utilizada fue el *error cuadrático normalizado*(NMSE):

$$NMSE = \sum_{k=1}^m \frac{\| m_i(k) - d_i(k) \|_2^2}{\| m_i(k) \|_2^2}, \quad (4-25)$$

donde  $m_i(k)$  es la dinámica temporal obtenida por el modelo y  $d_i(k)$  es la dinámica temporal de los datos reales. En la Figura 4-4 se pueden observar los resultados del ajuste.

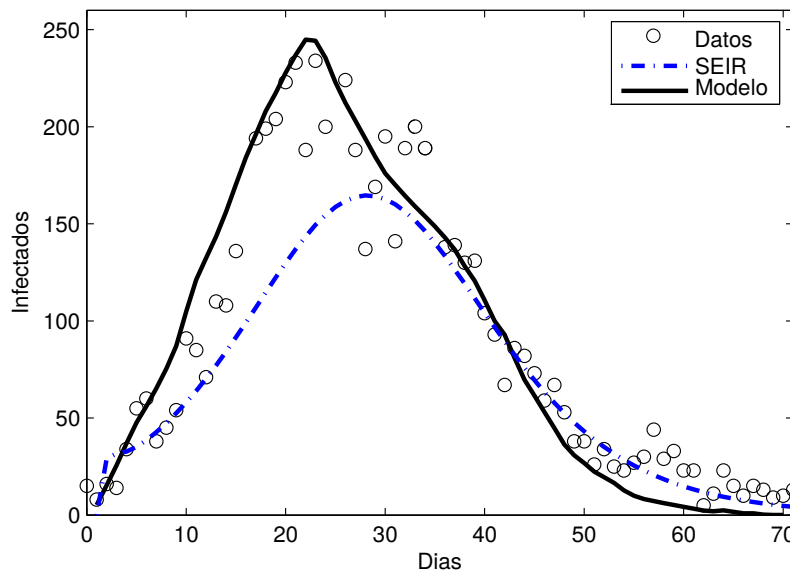
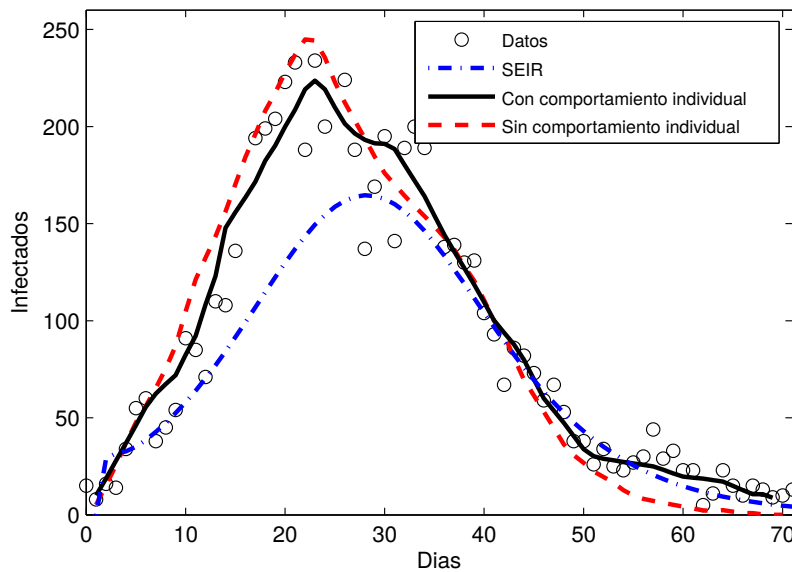


Figura 4-4.: Ajuste del modelo, respuesta promedio de 500 corridas

Nuevamente, al igual que en el capítulo anterior, se modela el comportamiento individual de los agentes a través de un FCM usando el mismo mapa. En este caso no se utiliza el

**Tabla 4-1.:** Parámetros minimizados del modelo clásico

$\beta'$	$\rho$	$\gamma$	$\alpha$	$q$	$N_e$	$N_i$
8,1	0,09	0,26	0,506	0,075	207	136

**Figura 4-5.:** Ajuste del modelo. Respuesta promedio sobre 500 corridas.

parámetro  $\omega$ . Si  $0,6 \geq C_{10} > 0,9$  los agentes se mueven a menor velocidad tienden a evitar el contacto con otros agentes.

## 4.6. Validación

La validación del modelo es un paso importantes dentro del proceso de construcción del mismo. Se pueden utilizar varias metodologías para llevar adelante este proceso. Por un lado, los métodos gráficos ilustran una amplia gama de aspectos complejos de la relación entre el modelo y los datos observados. Por otro lado, los métodos numéricos ilustran de forma precisa aspectos particulares de la relación entre el modelo y los datos observados. Los métodos numéricos para la validación de un modelo revelan un número único o resultado que representa el grado de confiabilidad del mismo.

### 4.6.1. Validación visual

Al igual que en capítulo anterior por medio de la validación visual del modelo se plantea las hipótesis subyacentes del mismo para asegurar la validez de las mismas.

Con el fin de comprobar la capacidad de nuestro modelo de ajustarse a dinámicas epidemiológicas reales, se muestran los resultados del ajuste en la Figura 4-5, allí puede observarse el ajuste del modelo espacialmente explícito basado en individuos móviles con y sin conducta social contra el ajuste del modelo *SEIR* clásico. Para ambos casos se observa que hay algunos puntos correspondientes a los datos reales que no se ajustan de forma precisa y puede observarse que tanto el modelo con conducta social como el modelo sin conducta social capturan el pico de la ola epidémica de forma precisa mientras que el modelo clásico no. Como modelo representa un enfoque estocástico realizamos 500 experimentos con el fin de capturar el comportamiento medio.

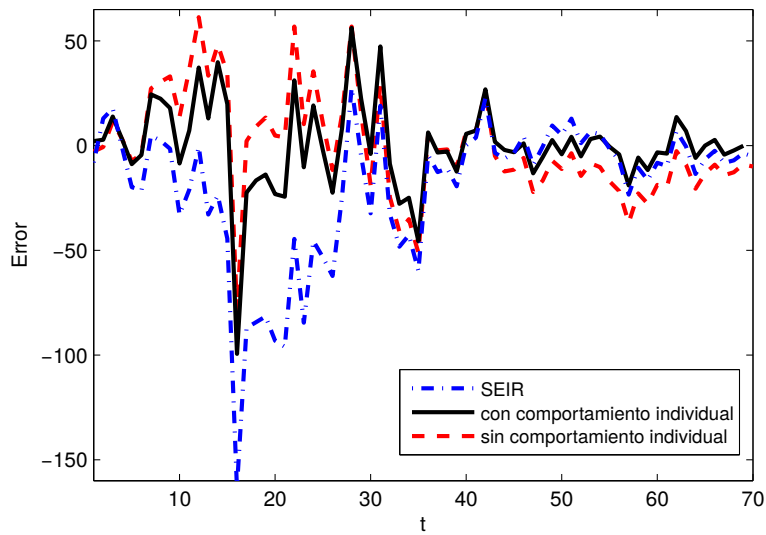
La Figura 4-6a muestra los residuos de los modelos y permite evaluar la calidad de los mismos para aproximar la dinámica real. Por otro lado la Figura 4-6b muestra la distribución de los errores cometidas por cada modelo incluido el modelo clásico.

**Tabla 4-2.:** Estadísticos del error

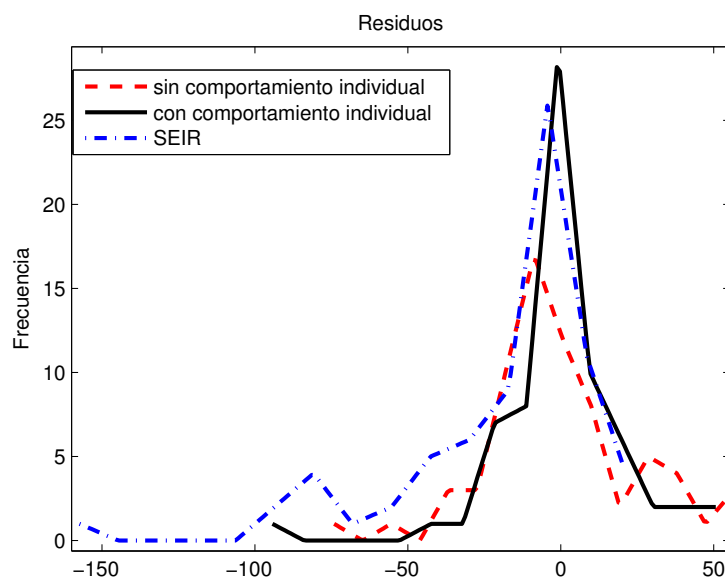
Modelo	$\bar{X}$	$M$	$std$	$S$	Curtosis
SEIR	-19,57	-9,15	54,75	0	1,78
Con FCM	-21,63	1,79	46,41	0	1,75
Sin FCM	-8,73	-8,12	41,77	0	1,77

En la Tabla (4-2) se cuantifican las principales medidas estadísticas que caracterizan a la distribución del error para cada uno de los modelos, donde  $\bar{X}$  es la media,  $M$  es la moda,  $std$  es el desvíos estándar y  $S$  es el sesgo de la distribución. Al igual que en el capítulo anterior, el sesgo  $S$  se calcula a través del segundo *coeficiente de Pearson* de la forma  $S = 3(\bar{X} - Md)/std$ , donde  $Md$  es la mediana. El coeficiente de sesgo da una idea de la simetría de la distribución, si  $S > 0$  la simetría de la distribución es positiva, si  $S < 0$  la simetría es negativa y si  $S = 0$  la distribución es simétrica. Esto quiere decir que mientras mas cerca de 0 esté el valor de  $S$ , mejor es la aproximación del modelo. En ambos casos, tanto el modelo con comportamiento social (con FCM) como el sin comportamiento social (sin FCM) el valor del sesgo es muy cercana a 0, con una leve asimetría negativa para el segundo caso. En tanto a la media y la mediana, puede observarse que  $\bar{X}$  y  $M$  son considerablemente menores para los dos modelos basados en individuos.





(a)



(b)

Figura 4-6.: (a) Residuos del modelo con espacio explícito. (b) Residuos del modelo clásico.

### 4.6.2. Validación numérica

El modelo fue numéricamente validado utilizando nuevamente el criterio de información de Akaike ( $AIC$ ) al igual que en el capítulo anterior. Es importante recordar que dicho criterio proporciona una medida de la calidad del modelo teniendo en cuenta la precisión y la complejidad del mismo. Como ya fue mencionado antes, los modelos que tienen un coeficiente  $AIC$  dentro del rango  $[1 - 2]$  soportan de forma consistente la variación estructural en los datos, si el valor está dentro del rango  $[3 - 7]$  soportan de manera considerable la variación estructural en los datos y si el coeficiente  $AIC > 10$  no explican algunas variaciones estructurales relevantes en los datos. El índice se calcula al igual que fue descrito en la sección anterior de acuerdo a la siguiente ecuación:

$$AIC = \log(\det(\frac{1}{m} \sum_1^m \varepsilon(t, \Theta)(\varepsilon(t, \Theta))^T)) + \frac{2n}{m}, \quad (4-26)$$

donde:  $\Theta$  es el conjunto de  $n$  parámetros de incerteza,  $m$  es el número de muestras y  $\varepsilon(t, \Theta)$  es la medida de error.

El criterio utilizado resulta en un coeficiente  $AIC = 6,4$  para el modelo basado en individuos con espacio explícito sin comportamiento social, mientras que para el modelo que contempla comportamiento social el  $AIC = 6,1$ . Para el modelo clásico, usando el mismo criterio el  $AIC = 7,5$ . En los casos propuestos con el modelo el índice está dentro del rango satisfactorio. El mismo se comporta de manera robusta ante la variación estructural de los datos.

Como ya se mencionó antes, el criterio de información de Akaike es equivalente a una validación cruzada *leave-one-out* en modelos de datos longitudinales, por ejemplo, selección de ancho de banda (Fang, 2011). Esta técnica es utilizada para evaluar los resultados de análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y los de prueba. Consiste en repetir y calcular la media aritmética obtenida de las medidas de evaluación sobre diferentes particiones. Se utiliza en entornos donde el objetivo principal es la predicción y se quiere estimar cuan preciso es un modelo implementado.

Teniendo esto en cuenta hemos evaluado la significación estadística de estos resultados mediante el cálculo de la probabilidad de que el error en la aproximación para los modelos implementados en este capítulo es mejor que el error del modelo clásico. Para realizar esta prueba asumimos la independencia estadística de los errores de ajuste para diferentes conjuntos de datos y nos aproximamos a los errores de *distribución binomial* por medio de una distribución gaussiana. Esto es posible porque tenemos un número suficientemente alto de registros para cada experimento (500 experimentos por data set).

Nuevamente, se dividió el conjunto de datos en tres subconjuntos. En primer lugar se entrenaron los modelos basados en individuos espacialmente explícitos y el modelo clásico con el mismo data set pero con  $n = 10$  datos menos que fueron seleccionados aleatoriamente siguiendo una distribución de probabilidad uniforme. Con el conjunto de datos

resultante se entrenó cada modelo con el propósito de obtener 7 generaciones de parámetros. Luego, se realizaron 500 corridas por cada generación de parámetros obtenida en el paso anterior con el propósito de obtener una buena aproximación de la respuesta promedio del modelo basado en individuos espacialmente explícito. Finalmente el error promedio de cada se calcula teniendo en cuenta la respuesta promedio del modelo utilizando los datos que fueron extraídos del data set original. Este proceso se llevó luego a cabo para  $n = 20$  y  $n = 30$ . De esta forma para cada conjunto de parámetros de ajuste con diferentes huecos en el data set tenemos que la hipótesis propuesta es que  $P(Error_{modelo} < Error_{SEIR}) > p$ . En el apéndice C se muestra la probabilidad de que el modelo basado en individuos descripto en este capítulo tenga un mejor desempeño que el modelo clásico.

En las Tablas (4-3) y (4-4) se resumen los resultados del test de significancia estadística para los diferentes conjuntos de validación para el modelo sin y con comportamiento social respectivamente. Puede verse que los modelos implementados son mejores que el modelo clásico si se observa el error de aproximación. El intervalo de confianza del error, en general, esta por encima del 90 %. Puede observarse una leve disminución de la probabilidad  $P$  en para el caso de cuando se extraen 20 datos. Este comportamiento no es el esperado y requiere de un análisis más detallado que se realizará en el futuro.

**Tabla 4-3.:** Significancia de los errores para el modelo sin comportamiento social

10 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Modelo basado en individuos SEIR	8,5	0,915	0,0045	4	99,55
	10,04	0,896	0,005		
20 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Modelo basado en individuos SEIR	5,27	0,947	0,0036	1,29	80,78
	5,75	0,943	0,0038		
30 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Modelo basado en individuos SEIR	5,18	0,948	0,0036	1,88	88,84
	5,88	0,941	0,0038		

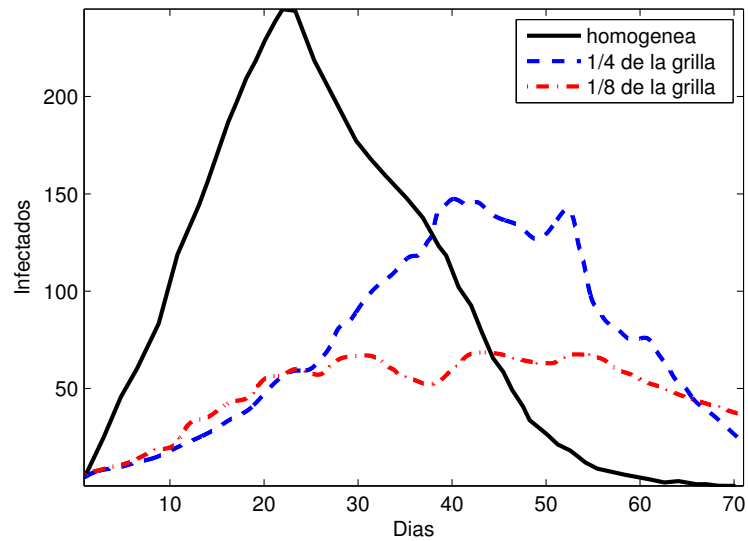
**Tabla 4-4.:** Significancia de los errores para el modelo con comportamiento social

10 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Modelo basado en individuos SEIR	8,69	0,913	0,0046	3,58	99,55
	10,04	0,896	0,005		
20 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Modelo basado en individuos SEIR	5,07	0,949	0,0036	1,85	88,84
	5,75	0,943	0,0038		
30 datos					
Modelo	Error	$\mu$	$\sigma$	$\frac{\mu_1 - \mu_2}{\sigma}$	$P(p_1 < p_2)$
Modelo basado en individuos SEIR	4,86	0,951	0,0035	2,81	97,23
	5,88	0,941	0,0038		

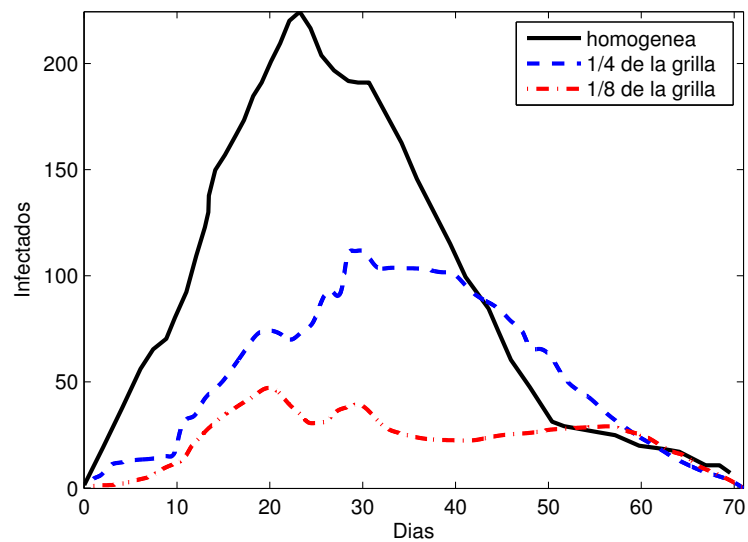
## 4.7. Otros Resultados

Como en el caso de autómatas, al relajar la condición de homogeneidad perfecta, se observan comportamientos alejados de lo ideal. Comenzando con los individuos expuestos sólo en una cuarta de la región de interacción, se observa una caída en la incidencia de la epidemia debido al retardo que se reduce en la homogeneización en la distribución de los individuos. Lo mismo ocurre si los expuestos ocupan inicialmente un octavo de la región. En la Figura 4-7a puede verse un promedio de 100 realizaciones para los dos casos antes mencionados y para distribución homogénea de los individuos infecciosos.

Los modelos basados en individuos introducidos en este capítulo permiten describir sistemas complejos como una colección de objetos simples, lo que lleva a comportamientos emergentes de gran complejidad. Los FCM, son modelos con retroalimentación utilizados para representar causalidad, combinan herramientas teóricas de los mapas cognitivos, la lógica difusa, las redes neuronales, las redes semánticas, los sistemas expertos y los sistemas dinámicos no lineales. En los experimentos anteriores no se tuvo en cuenta el comportamiento social de manera explícita, es decir a través del movimiento de los agentes. A través del algoritmo de flocking propuesto los agentes tienden a formar grupos mientras que el FCM influye en sus percepción del ambiente y media su relación con el resto de los agentes induciendo comportamientos mas complejos según la situación epidemiológica local. Al igual que en el caso anterior, la población de individuos infectados se distribuye de forma homogénea, en un cuarto o en un octavo de la grilla. En la Figura 4-7b se puede observar los resultados de las dinámicas resultantes con diferentes distribuciones iniciales para los individuos infecciosos. Nuevamente se observa el efecto de recuperación con retraso debido al comportamiento de los individuos que tienden a alejarse de los grupos poniéndose en un estado de aislamiento voluntario. En las figuras 4-8 y 4-9 pueden observarse las dinámicas espaciales para una distribución localizada y una distribución inicial uniforme de los individuos infecciosos respectivamente.



(a)



(b)

**Figura 4-7.:** Evolución de la epidemia para diferentes distribuciones espaciales de los individuos infecciosos (a) sin conducta social, (b) con conducta social.

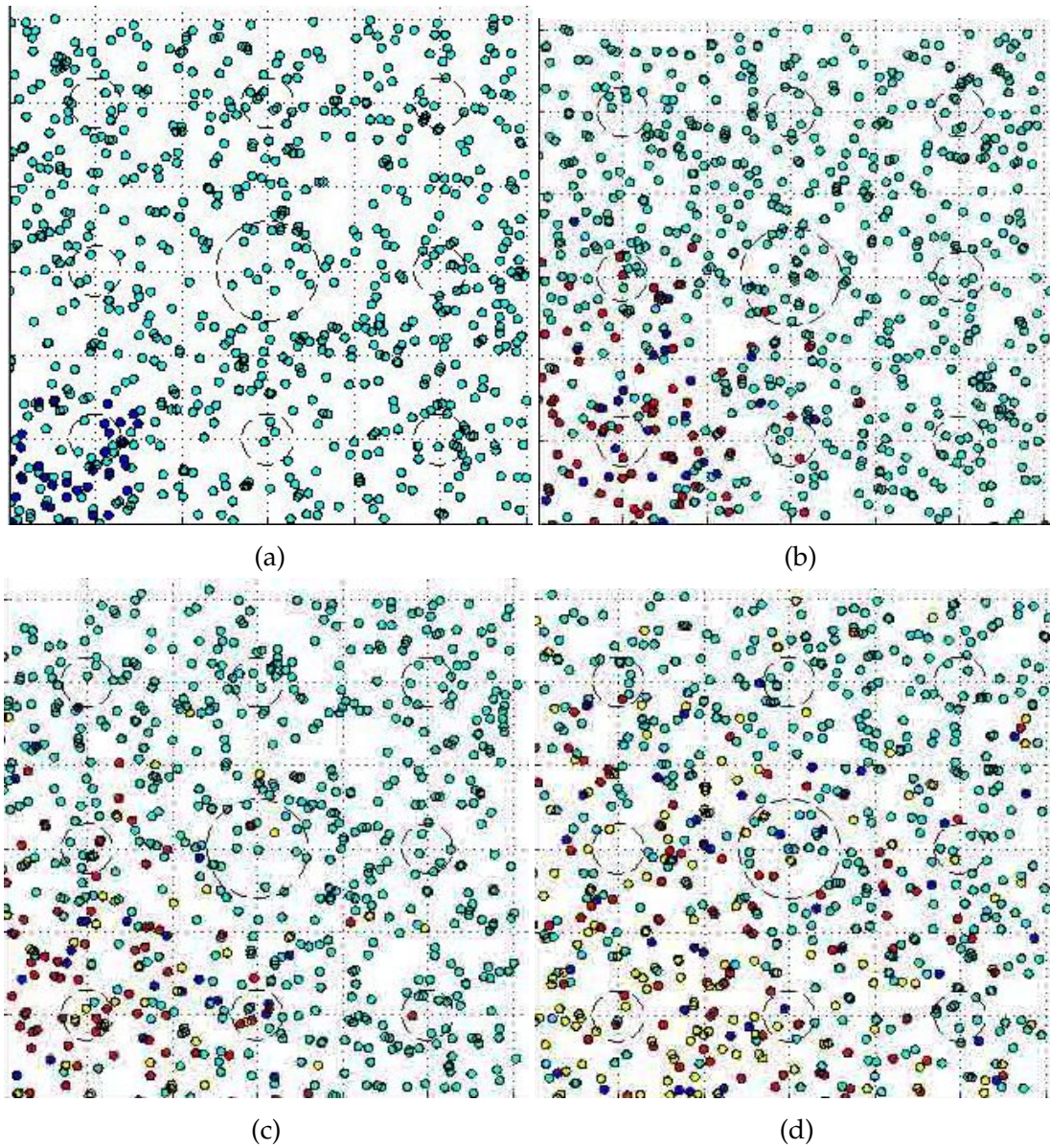
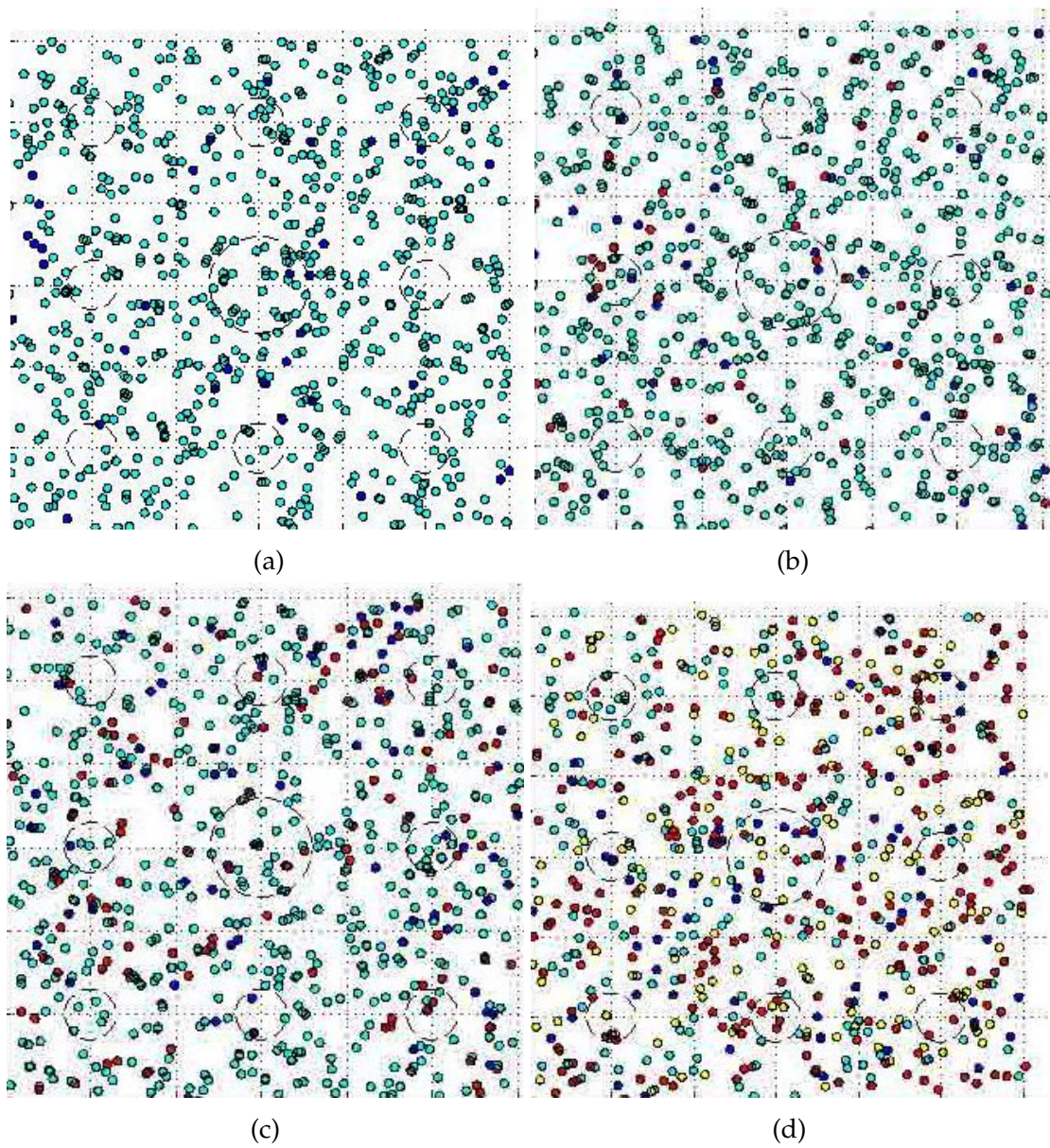


Figura 4-8.: Dinámica temporal para distribución espacial localizada para (a)  $t = 0$  (b)  $t = 10$  (c)  $t = 30$  (d)  $t = 50$



**Figura 4-9.:** Dinámica temporal para distribución espacial uniforme para (a)  $t = 0$  (b)  $t = 10$  (c)  $t = 30$  (d)  $t = 50$



## 4.8. Conclusiones

Los agentes son entidades que actúan en forma autónoma en un ambiente determinado, para poder actuar los agentes deben poder percibir dicho ambiente y todo lo que hay en el, por lo que son capaces de interactuar entre si. Para ello los agentes poseen un conjunto de funciones que le permiten desenvolverse como tal. Al igual que los autómatas los agentes modifican su comportamiento de acuerdo al estado en el que se encuentren gracias a una estructura de datos interna que le permite almacenar cierto tipo de información y actuar en base a las percepciones actuales y a la historia del sistema.

Los sistemas multi-agente permiten utilizar múltiples agentes unidos en una red para la resolución de un determinado problema. En el contexto de las simulaciones puede verse a un sistema multi-agente como un conjunto de agentes que interactúan en un ambiente determinado, permitiendo simular en forma directa a los individuos, sus comportamientos e interacciones.

Todos estos factores hacen a este tipo de modelado una buena alternativa a la hora de pensar en una forma de simular la pandemia de gripe española, por otro lado, dicha técnica ya ha sido utilizada con anterioridad para modelar otro tipo de enfermedades en poblaciones humanas, lo cual nos brinda un background lo suficientemente amplio a la hora de implementar un modelo propio.

Al modelar una epidemia en base a agentes, la población a estudiar está compuesta por individuos que interactúan entre sí, implementando explícitamente un agente por individuo. Los ambientes en los que se desenvuelven los agentes son descritos de manera explícita en este caso y si bien en el presente capítulo se utiliza una grilla vacía, es posible modelar diferentes regiones, ciudades o zonas más pequeñas, dependiendo del objetivo del modelado, y cada ambiente puede tener diferentes características (Bauer, Beauchemin, & Perelson, 2009; Auchincloss & Roux, 2008).

El ambiente ya no es una grilla discreta como en el caso de los autómatas celulares, sino que se trata de un espacio continuo, y los agentes se mueven libremente en el mismo. Los individuos interactúan al encontrarse dentro de una vecindad, ya sea en cuanto a la forma de moverse como a la posibilidad de ser infectados. Al igual que en el autómata, las vecindades pueden ser de tamaño variable. El movimiento de los individuos puede modelarse en forma puramente estocástica o seguir algún lineamiento acorde con lo observado para el individuo que se está modelando (ida o vuelta del trabajo, migración de aves, etc.) cambiando los potenciales de acción utilizados.

Al incorporar el comportamiento individual en el modelo a través del FCM puede observarse una mejora notable en la aproximación del modelo a los datos reales. Puede observarse que cuando se incorporan estos aspectos conductuales a los individuos, el modelo logra capturar la dispersión de los datos durante toda la ola. Incluso al final de la ola epidémica puede observarse que el modelo logra capturar los rebotes epidémicos. Este comportamiento puede ser adjudicado a las alteraciones que surgen en el movimiento

de los individuos a medida que la situación epidemiológica es mas grave. Hacia el final de la ola, cuando la situación mejora, los individuos tienden a recuperar sus patrones de movimiento naturales y es por ello que surgen nuevos brotes.

Los resultados estadísticos muestran una mejores resultados para el modelo con comportamiento individual si se los compara contra el modelo clásico o el modelo sin comportamiento individual. Esto puede verse en la media y la mediana del error cometido por el modelo.

La homogeneidad de contactos no pude ser asumida en este enfoque. Al realizar el modelado del espacio de manera explícita esta hipótesis no resulta natural. El uso de flocking para modelar el comportamiento social de los individuos contribuye a que sucedan contactos efectivos entre susceptibles e infecciosos como resultados de la formación de grupos cuando se asume que los individuos se distribuyen de manera homogénea en el ambiente.

Teniendo en cuenta que los contactos no son homogéneos, se hace mas evidente el rol de los individuos asintomáticos en el desarrollo de la epidemia. El valor del parámetro  $q$  refleja este aspecto.



# 5. Múltiples cepas y estrategias de control

## 5.1. Introducción

La gripe es una infección vírica aguda que se transmite fácilmente de una persona a otra por contacto directo o por micro gotas en el aire. Los virus de la gripe circulan por todo el mundo y pueden afectar a cualquier persona de cualquier edad, aunque hay grupos etarios que se pueden ver mas afectados que otros con respecto a las consecuencias que dicha enfermedad puede causar en los mismos. El virus de la gripe es el causante epidemias anuales que en las regiones templadas alcanzan su auge durante el invierno. La influenza es principalmente una infección comunitaria que se contagia en los hogares y en los ámbitos comunitarios. Las infecciones por influenza asociadas a los cuidados de salud pueden suceder en cualquier ámbito de cuidados de la salud y son más comunes cuando la influenza circula en la comunidad. Por lo tanto, las medidas de prevención de la influenza deben implementarse en todos los ámbitos de cuidados de la salud. Este virus es un problema grave de salud pública que puede ser causa de muerte en poblaciones de alto riesgo. Las epidemias pueden ejercer gran presión sobre los servicios de salud y tener importantes repercusiones económicas debido a la reducción de la productividad laboral. Existen varias estrategias de control a ser implementadas por parte de los estados y los organismos encargados de linear políticas de salud pública para mitigar los efectos de esta enfermedad en las poblaciones. La vacunación es la forma más eficaz de prevenir la infección, también existen drogas antivirales para tratar la enfermedad.

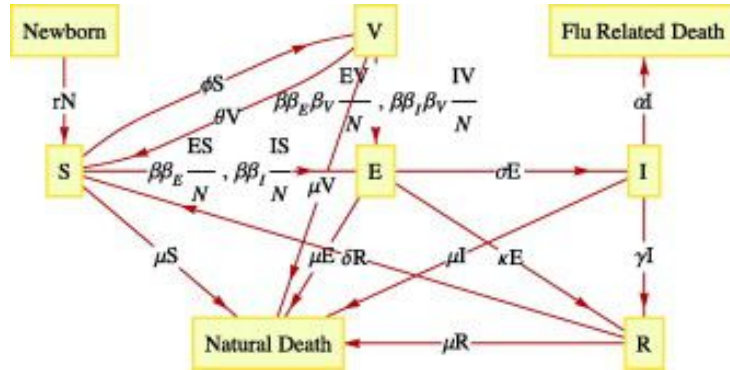
En el presente capítulo se organiza de la siguiente manera. Primero se realiza una breve revisión del estado del arte, repasando los diferentes enfoques empleados por otros autores para modelar estrategias de control epidemiológico y la coexistencia de múltiples cepas de gripe y su impacto en la dinámica de la enfermedad. Luego se procederá a explicar de qué manera pueden ser implementadas estas hipótesis en los modelos desarrollados en este trabajo, viendo sus principales ventajas por sobre otros enfoques. Finalmente se muestran los resultados de diversos experimentos y se desarrollan as conclusiones.

## 5.2. Estado del arte

En la bibliografía existen numerosos modelos que contemplan diferentes estrategias de control. Samsuzzoha, Singh, y Lucy (2012) proponen un modelo poblacional difusivo con vacunación basado en ecuaciones diferenciales ordinarias. Este enfoque incorpora un compartimento explícito que modela a la población vacunada. De esta forma el flujo poblacional del modelo *SEIR* clásico se ve modificada por la incorporación de un nuevo compartimento denominado *V* que modela a la población vacunada. Este modelo contempla que no toda la población susceptible es vacunada, solo una porción de la misma es vacunada y de esa porción un porcentaje determinado por una tasa de reincidencia puede volver al compartimento de susceptible y contagiarse el patógeno. El modelo propuesto en este trabajo modela la dinámica epidemiológica de la siguiente manera: La población susceptible se incrementa por los nacimientos y la incorporación de población que presenta pérdida de la inmunidad. La población susceptible se reduce mediante la acción de vacunación (porción de la población que se mueve a la clase *V*), infección (que se mueve a la clase *E*) y la muerte natural. La población vacunada se incrementa con la vacunación de personas susceptibles. La población vacunada tiene susceptibilidad a la infección reducida en un factor  $(1 - \beta_V, \text{ con } \beta_V \in [0, 1])$ . La clase vacunada se reduce por la infección (que se mueve a la clase *E*), muerte natural y disminución de inmunidad por vacunación. La población expuesta se incrementa debido a la infección de que permanecen susceptibles incluso después de haber sido vacunados. La población expuesta también se incrementa por la infección de población que no han sido vacunada todavía. La población expuesta se reduce por la recuperación (población que se mueve a la clase *R*), la muerte natural y el pasaje al compartimento de infectados (que se mueve a la clase *I*). La población, durante la etapa de expuesta, tiene un bajo nivel de infectividad. Técnicamente, este compartimento representa una etapa ligeramente infecciosa. La población infecciosa se incrementa por la fracción de individuos expuestos que se convierten a clase antes mencionada. La población de individuos infecciosos se reduce por la muerte natural, muerte relacionada con la enfermedad y la recuperación de la enfermedad. La población recuperada se incrementa cuando la población infectada avanza al compartimento de recuperado desde los compartimentos expuesto e infeccioso. La población recuperado se reduce por muerte natural y pérdida de inmunidad. Este trabajo evalúa el impacto que la vacunación de población susceptible tiene en el número reproductivo básico de la epidemia.

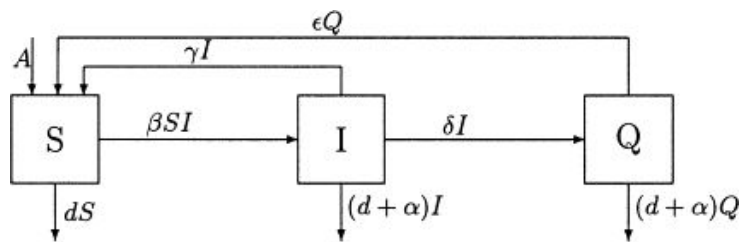
Otro tipo de estrategia que suele utilizarse como método de control de epidemias es la cuarentena. Formalmente cuarentena es la acción de aislar o apartar a personas o individuos durante un período, para evitar el riesgo de que extiendan una determinada enfermedad ya que están cerca de personas enfermas, los periodos de cuarentena pueden variar y no necesariamente llegan a durar cuarenta días.

El modelos *SIQS* fue introducido por Herbert en 2002 (Hethcote, Zhien, & Shengbing,



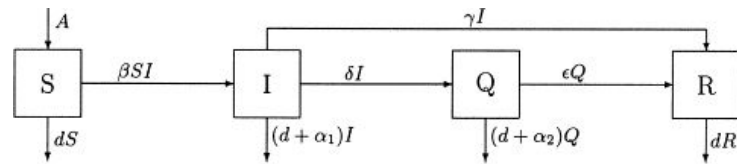
**Figura 5-1.:** Diagrama de flujo del modelo de Samsuzzoha, Singh, y Lucy (2012).

2002). En este modelo  $S$  es el número de individuos en la clase susceptible,  $I$  es el número de individuos en la población infectada y que no han sido aislados,  $Q$  es la población de individuos infectados que han sido aislados del resto de la población (población en cuarentena). En este modelo, la población en el compartimento  $Q$  crece a una tasa constante  $\delta$  que mide el paso desde el compartimento de población infectada. En este modelo las poblaciones de los compartimentos  $Q$  e  $I$  disminuyen a una tasa  $\alpha$  que corresponde a las muertes por enfermedad. En el trabajo de (X.-B. Zhang, Huo, Xiang, & Meng, 2014) los autores introducen mejoras al modelo de Herbert a través de la incorporación de términos no lineales para medir la tasa de infección y cuarentena de la población. En dicho trabajo los autores proponen dos enfoques para el modelo. Un modelo determinista en donde modifican la tasa de incidencia de la enfermedad por medio de un factor no lineal para darle más realismo al modelo y un modelo estocástico que denominan modelos  $SIQS$  estocástico. En las figuras



**Figura 5-2.:** Diagrama de flujo del modelo con estado de cuarentena  $SIQS$

El virus de la influenza es causado por un virus que puede ser clasificado en tres tipos diferentes ( $A$ ,  $B$  y  $C$ ). Entre estos tipos, el virus  $A$  es el más importante en términos epidemiológicos para los seres humanos, ya que puede recombinar sus genes con los de las cepas que circulan en las poblaciones de animales (aves, cerdos y caballos). Estas recombinaciones relativamente raras dan lugar cada pocos decenios a nuevos subtipos virales a través del denominado mecanismo de cambio antigénico. Los nuevos subtipos, que se clasifican de acuerdo a la naturaleza antigénica y genética de sus glicoproteínas de super-



**Figura 5-3.:** Diagrama de flujo del modelo con estado de cuarentena  $SIQR$  que contempla población recuperada

ficie, por lo general son antigénicamente muy diferentes de sus antepasados para escapar completamente de las defensas del sistema inmunitario de los anfitriones infectados previamente. En consecuencia, cada cambio antigénico está asociado con pandemias graves: la gripe española en 1918 o la gripe de Hong Kong en 1968, causada por la génesis de los subtipos  $H1N1$  y  $H3N2$ , respectivamente, son sólo dos ejemplos importantes. Al igual que la influenza  $B$ , el virus de influenza  $A$  también está sujeto a mutaciones puntuales en los genes  $HA$  y  $NA$  que inducen una deriva antigénica de las cepas víricas en cada subtipo. Estas mutaciones menores, que se producen en una escala de tiempo más rápido con respecto a los cambios, son las causas subyacentes de las epidemias de gripe que experimentamos todos los años. Hay mucha evidencia que muestra que la distancia antigénica entre dos cepas diferentes influye en el grado de inmunidad parcial, a menudo llamada inmunidad cruzada, conferida a los huéspedes que ya fueron infectados por una de las cepas inmunidad para la otra.

Muchos modelos matemáticos se han propuesto en la literatura para describir la ecología de la gripe en los seres humanos. Un enfoque típico es el de modelado de las interacciones entre los individuos que son (o han sido) infectados por diferentes cepas virales. Para este fin, varios modelos de  $SIR$  se pueden conectar a través de parámetros de inmunidad cruzada. En la Figura 5-4 muestra el modelo introducido por (Nuno, Feng, Martcheva, & Castillo-Chavez, 2005), en el mismo los autores modelan el tránsito de dos cepas del virus en una población. El análisis de estos modelos ha demostrado que varias cepas de influenza pueden persistir en las poblaciones humanas y que su prevalencia puede presentar oscilaciones auto-sostenidas en el tiempo. El principal inconveniente de estos modelos es que su análisis puede llegar a ser difícil analíticamente o numéricamente engorroso si el número de cepas es lo suficientemente grande para imitar la variedad real en la naturaleza.

Una ruta alternativa a la visualización de la aparición de nuevas cepas virales es el de introducir en el modelo una pérdida de la inmunidad por parte del anfitrión, como se sugiere originalmente por Pease para el caso de la gripe  $A$ . Otros autores introducen un nuevo compartimento al que llaman  $C$ , que puede ser llamado inmunidad cruzada, para diseñar un estado intermedio entre el estado completamente susceptible ( $S$ ) y el completamente recuperado ( $R$ ). En el trabajo presentado por (Casagrandi y col., 2006) puede observarse un ejemplo de este enfoque. Sólo mediante la ampliación del modelo  $SIR$  con

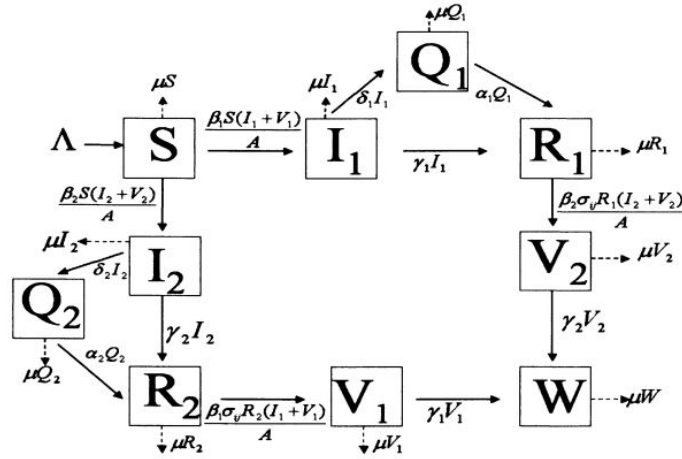


Figura 5-4.: Diagrama de flujo del modelo propuesto en (Nuno, Feng, Martcheva, & Castillo-Chavez, 2005),

una nueva clase de hecho uno puede tener en cuenta el hecho de que, cuando los anfitriones inmuno-cruzados están expuestos a cepas antigénicamente similares, pero diferentes, sus respuestas inmunes preexistentes son impulsadas. En la Figura 5-5 se muestra el diagrama de flujos para este enfoque compartimental.

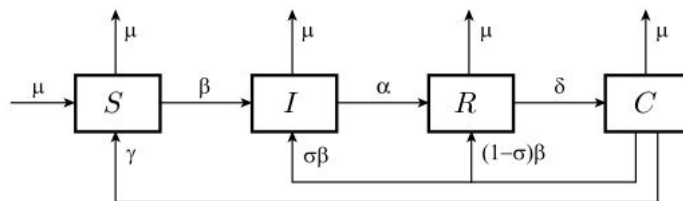


Figura 5-5.: Diagrama de flujo del modelo *SIR* que contempla múltiples cepas e inmunidad cruzada (Casagrandi, Bolzoni, Levin, & Andreasen, 2006)

### 5.3. Implementación

En la presente sección se procede a plantear diferentes escenarios hipotéticos sobre los cuales los modelos propuestos son evaluados. Las hipótesis bajo las cuales se corren los modelos son tres. En primer lugar se plantean un escenario de vacunación en donde una porción de la población es inmunizada previamente a la ola de influenza. En el segundo escenario se propone un escenario de cuarentena en donde se limita el movimiento de ciertos individuos en la población. Finalmente, el último escenario evaluado es la circulación de múltiples cepas del virus.



El aspecto fundamental a tener en cuenta de esta tesis es que lo que se presenta no son dos técnicas de modelado diferentes para un mismo problema. Lo que en realidad se está planteando es un marco de trabajo que puede ser aplicado a diferentes situaciones. En la Figura 5-6 se resume el enfoque propuesto.

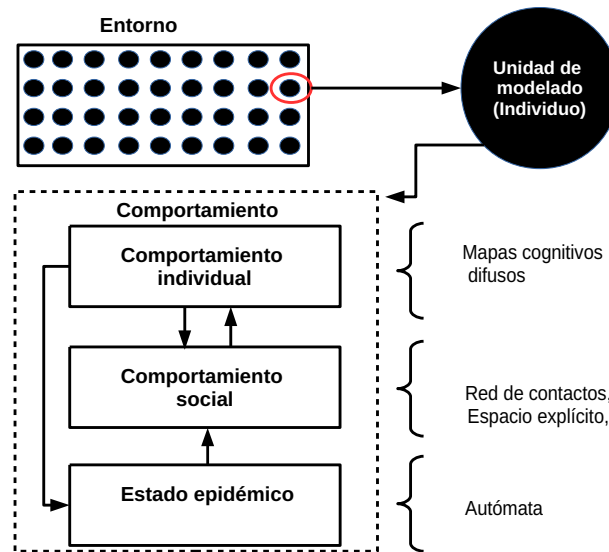


Figura 5-6.: Diagrama esquemático del marco de trabajo propuesto

El marco de trabajo presentado propone ver los modelos como construcciones modulares. Puede verse al sistema como una colección de objetos más simples que determinan la dinámica global del sistema a través de la integración de la dinámica individual. Dicha dinámica depende del comportamiento del individuo y ese comportamiento puede ser visto como la unión de tres bloques esenciales: (i) El comportamiento individual, el cual puede ser definido mediante un conjunto de reglas, a través de la lógica difusa o por medio de un enfoque mixto. (ii) El comportamiento social, que está determinado por la forma en la que se modelan las interacciones entre las unidades de modelado. (iii) El estado epidémico está determinado por la forma en la que se modela el mecanismo de contagio.

### 5.3.1. Múltiples cepas

Aunque los patógenos suelen considerarse entidades uniformes, cada vez hay más datos de que las infecciones suelen estar provocadas por más de una especie o de una cepa viral. Sin embargo, hay pocas publicaciones referidas a las infecciones por múltiples cepas de virus, posiblemente por las dificultades técnicas para distinguir las cepas individuales. Las cepas son importantes porque pueden diferir notablemente en diversos rasgos, como sus tasas de crecimiento, virulencia, poder patógeno y la capacidad antigénica. Si bien se han propuesto gran cantidad de modelos matemáticos que siguen el enfoque de

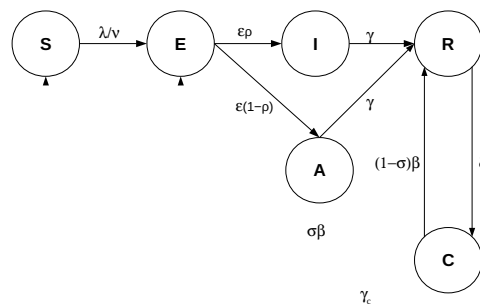
los modelos *SIR* con el fin de modelar este fenómeno, una gran variedad de estos modelos, basados en el acople de varios modelos *SIR* clásicos por medio de parámetros de inmunidad cruzada, resultan analíticamente complicados a medida que se añaden cepas. El problema de modelar múltiples cepas se puede pensar de forma alternativa como la incorporación al modelo de un parámetro que represente la pérdida de inmunidad por parte del huésped. Diferentes estudios experimentales han demostrado que la probabilidad de recuperación de un individuo infectado con una nueva cepa circulante se incrementa linealmente con el tiempo transcurrido desde la infección. En el caso del virus de la influenza A, tanto la evolución genética como la antigénica ocurren de manera lineal a lo largo del tiempo. Los resultados experimentales han demostrado que los modelos de simple cepa *SIRS* clásicos no pueden ser utilizados para el estudio de la influenza cuando en el ambiente está circulando más de una a la vez. Otros autores han propuesto considerar un estado de "inmunidad parcial temporal" para los individuos recuperados para resolver esto. Otra forma de modelar el fenómeno es agregando un nuevo estado *C*, al cual algunos autores definen como inmunidad cruzada, de esta forma se crea un estado intermedio entre un individuo completamente susceptible *S* y un individuo completamente recuperado. De esta forma, extendiendo el enfoque clásico añadiendo este nuevo estado uno puede tener en cuenta que cuando un individuo con inmunidad cruzada es expuesto a una cepa antigénicamente similar pero diferente, su respuesta inmune preexistente actuará sobre la cepa similar.

Desde el punto de vista epidemiológico se puede decir que para una cepa dominante de influenza cada individuo de la población puede encontrarse en cualquiera de los siguientes 5 estados:

1. Susceptible: cuando es completamente susceptible a la cepa dominante y las cepas secundarias.
2. Expuesto: posterior al contacto con un individuo infeccioso, se encuentra incubando la enfermedad.
3. Infeccioso: se encuentra infectado por la cepa principal.
4. Recuperado: cuando se encuentra recuperado totalmente.
5. Cross-inmune: cuando se encuentra recuperado parcialmente pero es susceptible a cepas secundarias.

Teniendo en cuenta esto, el grafo de estados del modelo se puede observar en la Figura 5-7. La novedad de este enfoque es que luego de recuperarse de la infección, los individuos pasan a un nuevo estado, el estado *C*. Los individuos que se encuentran en estado *R* son aquellos que se han recuperado de la cepa principal y tienen inmunidad total contra esa cepa, pero transcurrido un lapso de tiempo, los individuos pasan a un estado *C*, ya que solo tienen inmunidad parcial ante el surgimiento de nueva cepa dominante desde

el instante que ocurrió la primera infección. Si bien los individuos que se encuentran en la clase  $C$  se ven expuestos a la cepa dominante a la misma intensidad que aquellos individuos que se encuentran en el estado  $I$ , solo una fracción de ellos se infecta en un instante de tiempo determinado. Con este enfoque, los individuos que se encuentran en el estado  $R$  se asumen completamente inmunes.



**Figura 5-7.:** Grafo de estados del modelo de infección con múltiples cepas

En la Figura 5-8 se puede ver el ajuste del modelo basado en redes de contactos con el estado epidémico modificado como se describió anteriormente. Al autómata descrito en el capítulo 3 se le agrega un nuevo estado que es el estado de inmunidad cruzada  $C$ . Un parámetro  $\gamma$  con distribución uniforme en el intervalo  $[0,1, 0,2]$  determina la probabilidad de transición de  $C$  a  $S$ , un parámetro  $\delta$  distribuido de manera uniforme en el intervalo  $[0,2, 0,4]$  que determina la probabilidad de paso de  $R$  a  $C$  y un parámetro que  $\sigma$  distribuido uniformemente en el intervalo  $[0,05, 0,09]$  que determina la probabilidad de transición de  $C$  a  $E$  y de  $C$  a  $R$ , siendo dichas probabilidades  $\sigma\beta$  y  $(1-\sigma)\beta$  respectivamente con siendo  $\beta$  la probabilidad de transmisión derivada del modelo clásico.

En la Figura 5-9a se observan los resultados de los residuos del modelo con comportamiento social, sin comportamiento social y el modelo SEIR para múltiples cepas y también puede observarse puede verse la distribución de los errores.

**Tabla 5-1.:** Estadísticos del error

Modelo	$\bar{X}$	$M$	$std$	$S$	Curtosis
SEIRC	-33,45	-129,51	61,36	0	1,78
Con FCM	-7,89	-95,58	56,02	0	1,75
Sin FCM	1,97	-98,79	64,38	0	1,77

En la Tabla (5-1) se cuantifican las principales medidas estadísticas que caracterizan a la distribución del error para cada uno de los modelos, donde  $\bar{X}$  es la media,  $M$  es la

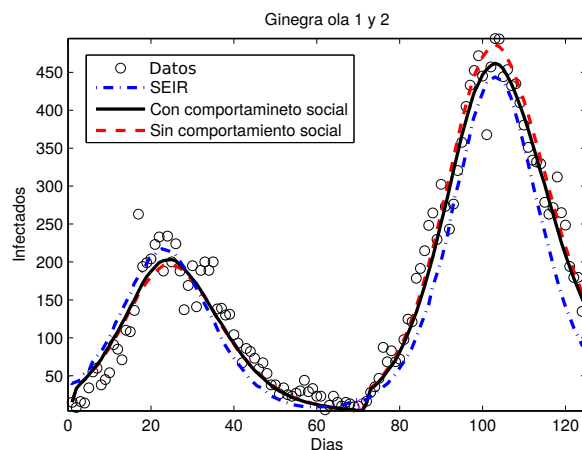


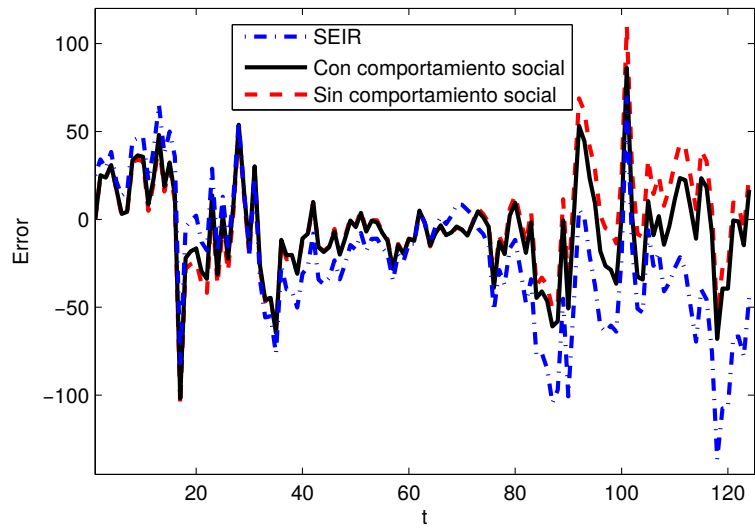
Figura 5-8.: Ajuste del modelo para múltiples cepas

moda,  $std$  es el desvíos estándar y  $S$  es el sesgo de la distribución. Al igual que en el capítulo anterior, el sesgo  $S$  se calcula a través del segundo *coeficiente de Pearson* de la forma  $S = 3(\bar{X} - Md)/std$ , donde  $Md$  es la mediana. En ambos casos, tanto el modelo con comportamiento social (con FCM) como el sin comportamiento social (sin FCM) el valor del sesgo es muy cercana a 0, con una leve asimetría negativa para el segundo caso. En tanto a la media y la mediana, puede observarse que  $\bar{X}$  y  $M$  son considerablemente menores para los dos modelos basados en redes de contactos.

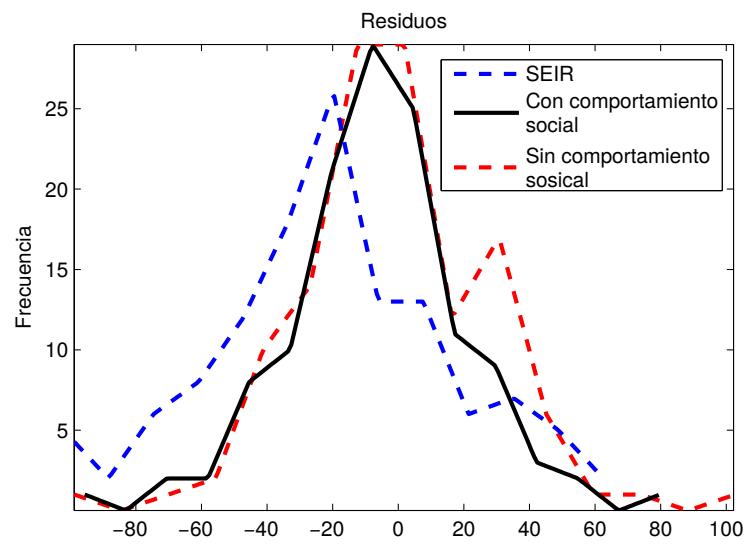
El modelo se validó numéricamente mediante el criterio de información de Akaike, que proporciona una medida de la calidad del modelo teniendo en cuenta la precisión y la complejidad del mismo. El criterio utilizado resulta en un coeficiente  $AIC = 6,5$  para el modelo sin comportamiento social, mientras que para el modelo que contempla comportamiento social el  $AIC = 6,4$ . Para el modelo clásico, usando el mismo criterio el  $AIC = 7,7$ . En los casos propuestos con el modelo basado en redes de contacto el índice está dentro del rango satisfactorio. El modelo se comporta de manera robusta ante la variación estructural de los datos. La matriz  $W$  utilizada para modelar el comportamiento individual es la descrita en el capítulo 2.

### 5.3.2. Vacunación

La vacunación es considerada como la política de intervención de mayor éxito, así como una estrategia rentable para reducir tanto la morbilidad y la mortalidad de los individuos. Se ha utilizado para combatir enfermedades como el sarampión, las paperas, la rubéola, el tétanos, la difteria, la hepatitis B y la influenza. Sin embargo, la vacunación no implica necesariamente inmunidad de por vida para una persona vacunada. En algunos casos, la vacunación es ineficaz y no proporciona inmunidad en absoluto (el llamado "fallo primario"), a veces la inmunidad adquirida mediante la vacunación disminuye y no



(a)

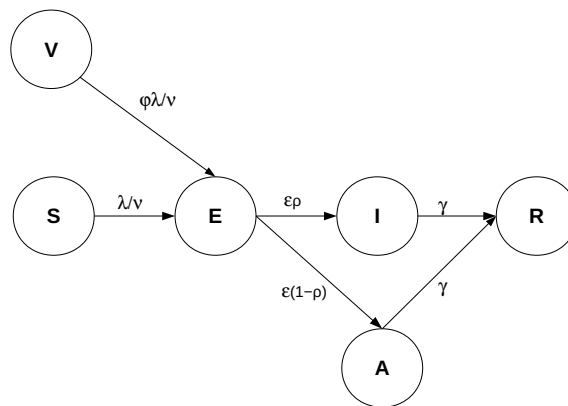


(b)

**Figura 5-9.:** (a) Gráfico de residuos (b) Distribución del error

proporciona la protección esperada ("fallo secundario"). Por otra parte, es prácticamente imposible vacunar a todos los individuos susceptibles en una comunidad dada, sobre todo en países donde estas vacunas no son de fácil disponibilidad o accesibilidad.

En el caso del modelo basado en redes de contacto presentado en el capítulo 3 de esta tesis, puede simularse esta estrategia de control epidémico realizando una mínima modificación del mismo. Si se tiene en cuenta que la vacuna contra la influenza disminuye en el un 50 % o 60 % la probabilidad de padecer la infección durante la época de mayor circulación de la cepa principal (McLean y col., 2014; Ohmit y col., 2014). Esto se puede capturar con el enfoque propuesto agregando un nuevo estado al autómata. Llamemos a ese estado  $V$ , o vacunados. Los individuos que se encuentren en dicho estado cuentan con una probabilidad de infección  $\varphi$  menor que aquellos que no han sido vacunados.

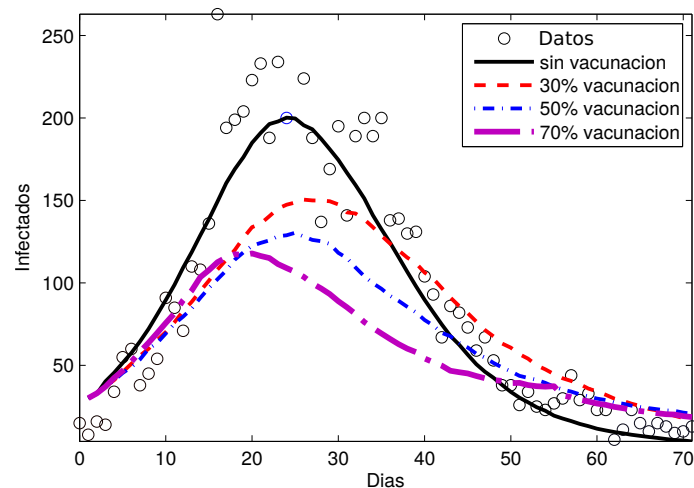


**Figura 5-10.:** Grafo de estados del modelo con vacunación

Si bien la vacuna contra la influenza es una muy buena manera de prevenir la enfermedad, la protección puede variar ampliamente según a quien se le administre la vacuna, además de qué tan bien se corresponde dicha vacuna contra la influenza con los virus en circulación. En general, la vacuna contra la influenza tiene mayor efecto entre adultos saludables y niños mayores. Algunas personas mayores y personas con ciertas enfermedades crónicas podrían desarrollar menos inmunidad que los niños y los adultos sanos después de la vacunación. No obstante, incluso para estas personas, la vacuna contra la influenza puede brindar cierta protección. A fin de contemplar estos factores en el modelo, se establece que los individuos vacunados tienen una probabilidad menor que los individuos susceptibles no vacunados de contraer la enfermedad. En la Figura 5-10 se puede observar el grafo de estados del autómata cuando se contempla la vacunación. Con el propósito de evaluar la variabilidad en la eficiencia de la vacuna, el parámetro  $\varphi$  se puede tomar valores en un intervalo  $[0,45, 0,65]$  siguiendo una distribución de pro-

bilidad uniforme. Para llevar a cabo el experimento se inicializa la población con un número variable de individuos vacunados.

Desde el punto de vista de las redes de contactos, vacunar a un individuo significa cambiar la influencia que un nodo ejerce sobre los nodos vecinos. Al aumentar el porcentaje de individuos vacunados en la población disminuye la influencia de que los nodos infecciosos tienen sobre los nodos susceptibles. En la Figura 5-11 pueden verse los resultados de cómo la dinámica temporal de la enfermedad se ve afectada cuando aumenta el porcentaje de individuos en la población que son vacunados. Como es intuitivo pensar, al vacunar individuos en la población, se reduce la probabilidad de contactos efectivos. Desde el punto de vista de una red de contactos, esta práctica es equivalente a decir que ciertos nodos de la red (los individuos vacunados) representan obstáculos por los cuales el patógeno no puede circular y esparcirse por el resto de la red.



**Figura 5-11.:** Evolución de la epidemia para diferentes porcentajes de individuos vacunados. Respuesta promedio para 500 realizaciones por caso.

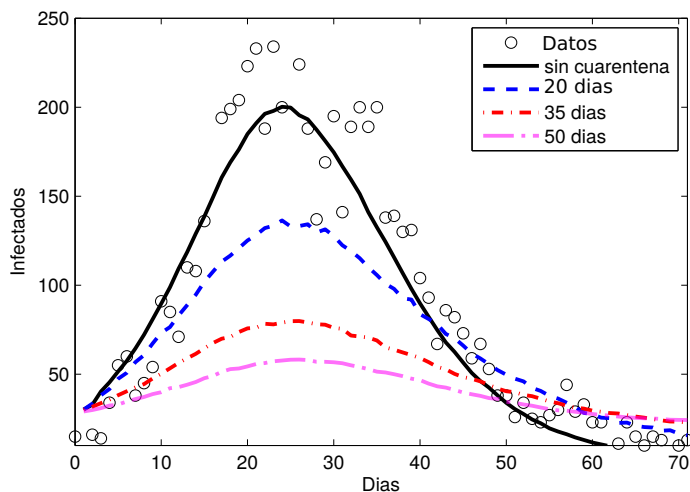
Como es de esperar al aplicar una política de control sobre la población, en este caso vacunación, puede observarse como el pico de la epidemia disminuye a medida que se aumenta el porcentaje de población que es vacunada.

### 5.3.3. Cuarentena

El aislamiento y la cuarentena ayudan a proteger al público mediante la prevención de la exposición a las personas que tienen o pueden tener una enfermedad. El aislamiento separa a las personas enfermas con una enfermedad contagiosa de las personas que no están enfermas. La cuarentena separa y restringe el movimiento de las personas que estuvieron expuestas a una enfermedad contagiosa para ver si se enferman.

Existen diferentes formas de modelar este tipo de prácticas, según sea la herramienta utilizada. En Este trabajo se modela esta práctica limitando el movimiento de los individuos en el espacio de los contactos. Debido a que en la práctica resulta imposible detectar todos los casos de enfermedad se añade un componente estocástico al modelo, es decir, del total de individuos infectados se supone que solo una porción de los mismos son detectados por los agentes de salud pública y puestos en cuarentena.

En los modelos poblacionales clásicos se asume que la la práctica de esta política sanitaria no es 100 % efectiva, por ello se contempla que la población puesta en cuarentena puede contagiar a la población susceptible a una tasa menor que la población infecciosa que no está puesta en cuarentena. En el caso del modelo basado de redes de contactos propuesto esta hipótesis es también cierta. Si bien no se contempla explícitamente como un estado en el modelo del estado epidémico, aquellos individuos puestos en cuarentena dejan de moverse en el espacio de contactos, limitando su radio de acción, pero no dejan de ejercer influencia sobre los individuos que se encuentran dentro de su radio de acción. De esta manera, si bien se limita la circulación de agentes infecciosos en la red global, a nivel local el patógeno puede seguir circulando siempre que haya susceptibles al alcance de un infectado.



**Figura 5-12.:** Evolución de la epidemia para diferentes períodos de duración de cuarentena. Respuesta promedio para 500

En la Figura 5-12 se puede observar la evolución de la dinámica temporal para diferentes porcentajes de población infectada puesta en cuarentena. Cuando los individuos son puestos en cuarentena no pueden moverse libremente a través de la red de contactos, esto equivale a que la topología de la red permanece constante en ciertas zonas a partir del instante de tiempo en el que esta práctica es puesta en práctica, como resultado, los individuos infectados no propagan la enfermedad ya que el número de contactos que tie-



ne disminuye cuando se reduce el movimiento. Los parámetros utilizados en este caso corresponden a los presentados en el capítulo 3 con un  $\beta = 8,3$  constante.

## 5.4. Conclusiones

Es importante resaltar la importancia que tienen las políticas sanitarias en el desarrollo de una epidemia. Desde el punto de vista epidemiológico, puede apreciarse que la dinámica epidemiológica cambia de forma drástica cuando se interviene el proceso natural con prácticas como la vacunación o la cuarentena. Por otro lado, también puede verse que la circulación de múltiples cepas representa un riesgo natural para el surgimiento de nuevos brotes y es claramente el factor determinante de lo que se conoce como olas epidémicas. Sin embargo, como ya se ha mencionado antes, desde el punto de vista del marco propuesto es más interesante resaltar que lo que en este trabajo se presenta no queda limitado al caso práctico aquí estudiado. El objetivo de este capítulo era mostrar como realizando modificaciones mínimas en ciertos módulos de nuestro marco resultan en modelos nuevos que se adaptan a la situación que se desea modelar.

De esta forma, cuando se modela cuarentena, vacunación o múltiples cepas, puede verse como un cambio en el módulo de compartimiento epidémico de unidad de modelado básica. A su vez la cuarentena, también puede ser vista como una modificación en las reglas que rigen el comportamiento individual.



# 6. Conclusiones

## 6.1. Conclusiones

En esta tesis se abordó el problema del modelado de epidemias a través de un enfoque basado en individuos. Estos modelos representan un marco de modelado significativamente diferente del propuesto por los modelos basados en individuos tradicionales. La forma clásica de implementar este tipo de modelos se basa en reproducir el comportamiento de los individuos a través de un conjunto de reglas. En este trabajo proponemos modelar el comportamiento de los individuos a través de la agregación de tres bloques conductuales. De esta forma, las interacciones e infecciones individualizadas definen la dispersión de la enfermedad a nivel de la población.

Si bien, hasta la fecha han sido implementados un gran número de trabajos que utilizan el modelado basado en individuos para modelar diferentes enfermedades. Los diferentes enfoques varían considerablemente y no existe un marco conceptual unificado para la implementación de modelos basados en individuos. Por otro lado, existe un gran número de herramientas de software, plataformas y librerías que permiten implementar todo tipo de modelos pero se trata de aplicaciones y no de un marco de trabajo formal.

El marco conceptual presentado en esta tesis propone la construcción de un modelo basado en individuos a través de un enfoque modular. La dinámica del individuo es consecuencia del comportamiento del mismo. El comportamiento se construye a partir de tres unidades: *i)* comportamiento individual; *ii)* comportamiento social y *iii)* el estado epidémico. De esta forma, los módulos por separado pueden ser modelados utilizando diferentes técnicas de modelado. La dinámica global del sistema surge como resultado de la agregación de las características individuales.

Para validar el modelo propuesto, se abordó el problema del modelado de una epidemia de gripe utilizando diferentes técnicas para modelar el comportamiento social, individual y epidemiológico. Por un lado las redes de contactos que permiten obtener buenas aproximaciones de las dinámicas temporales que caracterizan a este tipo de epidemias y posibilitan ver la dimensión social del fenómeno a través del contacto entre los actores de la red. Por otro lado, los modelos espacialmente explícitos permiten contemplar al espacio de forma explícita y de esta forma contempla la dinámica espacial además de la dinámica temporal.

Las principales conclusiones de esta tesis se resumen en los siguientes puntos:

- Al tratar con poblaciones, la interacción social es la piedra basal de la epidemiología. La interacción tiene influencia en la transmisión de las enfermedades, siendo una variable fundamental a la hora de modelar este tipo de fenómenos.
- El estudio de los cambios que caracterizan las dinámicas poblacionales ayudan a controlar y comprender los factores que influyen dichos cambios.
- Los modelos clásicos se basan en el cálculo del  $R_0$  para determinar la permanencia en forma endémica de la epidemia o su extinción. Este parámetro depende de la tasa de infección, la tasa de paso de expuesto a infeccioso, la tasa de recuperación y la tasa de mortandad, y determina un equilibrio endémico para un valor mayor a 1, mientras que predice la extinción de la epidemia si es menor o igual a 1. Para que el modelo funcione y este parámetro prediga el curso de la epidemia deben cumplirse razonablemente bien las hipótesis de distribución uniforme e instantánea, así como ser constantes los parámetros en todos los individuos. Esta hipótesis generalmente no se cumple, lo que deja a estos modelos restringidos a un universo de aplicaciones relativamente pequeño.
- El enfoque que brindan los modelos basados en individuos resulta especialmente útil cuando se intenta modelar un sistema como una colección masiva de objetos que en su totalidad afectan el comportamiento de todo el sistema. Los sistemas biológicos son no lineales por naturaleza y si bien los modelos clásicos pueden representar correctamente muchos aspectos de los mismos, la validez de sus resultados está fuertemente atada a que se cumplan ciertas hipótesis que sólo se garantizan bajo la ley de los grandes números.
- Los patrones de contacto interpersonal que subyacen a la transmisión de enfermedades de forma natural pueden ser considerados para formar una red. En un brote, la enfermedad se extiende a lo largo de los enlaces de la misma, es decir, a través de los contactos que se establecen en los individuos de la red. Es por ello que se puede pensar en utilizar modelos basados en redes de contactos para simular este tipo de fenómenos. Como puede observarse en los resultados del capítulo 3 de la tesis, utilizando las mismas hipótesis que en los modelos clásicos (mezcla homogénea de los individuos y contactos homogéneos), pero limitando el número máximo de contactos que un individuo puede tener en la red, se pueden obtener dinámicas realistas modelando la conducta social como una red aleatoria.
- En el contexto de las redes de contacto (capítulo 3), el movimiento de los individuos contribuye a la reconfiguración en la topología de la red de contactos, incrementando la probabilidad de contacto entre determinados actores en un instante de tiempo dado, favoreciendo contactos entre susceptibles e infecciosos y por lo tanto la aumenta la probabilidad de contagio. En el contexto del espacio explícito (capítulo 4)

el movimiento de los individuos contribuye a la propagación del patógeno ya que si los individuos permanecen inmóviles en el ambiente, luego de un tiempo, no existe la probabilidad de contacto efectivo entre susceptibles e infecciosos ya que a medida que los infectados se recuperan estos actúan de barrera natural para contener la propagación de la enfermedad.

- El tamaño de la vecindad contribuye a determinar en nivel de heterogeneidad global. En el capítulo 3 puede verse como los resultados del modelo cuando se asume un tamaño de vecindad mas grande se aproximan a los resultados del modelo clásico. Esto puede interpretarse como una forma de llevar las hipótesis del modelo hacia las hipótesis que gobiernan a los modelos clásicos (heterogeneidad de contactos).
- En el modelo de espacio explícito (capítulo 4) los agentes permiten representar al sistema como un conjunto de individuos que interactúan en un ambiente determinado. En este modelo el espacio físico se contempla de forma explícita y ya no se puede suponer homogeneidad de contactos entre los individuos. Este aspecto es reflejado en los parámetros que gobiernan al estado epidémico, que si bien son similares a los obtenidos para la red de contactos, puede observarse que el rol de los individuos asintomáticos en la propagación de la epidemia es considerablemente mas importante. El algoritmo de flocking utilizado para simular el comportamiento social de los agentes contribuye a la propagación del patógeno en la red ya que la formación de grupos sumada a la distribución inicial homogénea de los individuos en el ambiente contribuyen al surgimiento de brotes distribuidos en el espacio y a una incremento en la probabilidad de los contactos efectivos.
- Si nos alejamos de la condición de homogeneidad perfecta se puede ver un comportamiento que está lejos de ser el ideal; esto se debe a que las variaciones en las interacciones ocasionan una dinámica anormal, prolongando o reduciendo ciertos períodos durante la epidemia. Este comportamiento puede apreciarse cuando se relajan las hipótesis de los modelos clásicos tanto para el modelo basado en redes de contactos como para el modelo de espacio explícito.
- En el capítulo 5 es importante destacar la versatilidad y la facilidad que permite este marco para contemplar diferentes situaciones epidemiológicas. Pueden de esta forma contemplarse la hipótesis de una epidemia con múltiples cepas simplemente cambiando el modelo que gobierna el estado epidémico. Esta forma de modelar el problema permite abordar diferentes situaciones sin la necesidad de ajustar nuevamente todos los parámetros del modelo.
- Los modelos espacialmente explícitos basados en individuos permiten incorporar heterogeneidad al sistema y observar comportamientos que se asemejan mas al del

sistema real. Por otro lado, al ser capaces de incorporar características inherentes al comportamiento de los individuos que componen la población, permiten crear sistemas más realistas.

- En el contexto del trabajo realizado, el modelo implementado permite modelar de forma sencilla sistemas complejos a través de la agregación de tres módulos principales. La fortaleza del marco propuesto permite abarcar diversos problemas simplemente cambiando una parte del modelo sin la necesidad de ajustar nuevamente todos los grados de libertad.
- El uso de mapas cognitivos difusos para representar el comportamiento individual permite representar de manera sencilla la complejidad que gobierna la conducta de los individuos. Este aspecto no siempre puede ser representado a través de reglas debido a la complejidad natural de este proceso.
- Finalmente es importante resaltar que el marco propuesto en esta tesis permite la construcción de modelos complejos basados en individuos a partir de módulos simples. Explotando las ventajas de los sistemas auto-organizados para modelar fenómenos no lineales de manera precisa.

## 6.2. Publicaciones resultantes

Como resultado de esta tesis doctoral se han publicado diferentes trabajos a congresos y revistas que se listan a continuación:

### Trabajos publicados en revistas

- [1] Leonardo López; Germán Burguener; Leonardo Giovanini. Addressing population heterogeneity and distribution in epidemics models using a cellular automata approach. BMC Research Notes.: Biomed Central. 2014 vol. n°. p - . issn 1756 – 0500.
- [2] Leonardo López; Agustín Izquierdo; Darío Manzoli; Pablo Beldoménico; Leonardo Giovanini. A myiasis model for *Philornis torquans* (Diptera: Muscidae) and *Pitangus sulphuratus* (Passeriformes: Tyrannidae). ECOLOGICAL MODELLING. Amsterdam: ELSEVIER SCIENCE BV. 2016 vol.328 n°. p62 - 71. issn 0304 – 3800.

### Trabajos en eventos científicos-tecnológicos publicados

- [1] Leonardo López, Burguener Germán, Giovanini Leonardo. Modelling population heterogeneity in epidemics using cellular automata. Argentina. Rosario. 2011. Revista. Artículo Completo. Congreso. ENIEF 2011. Asociación Argentina de Mecánica Computacional

- [2] Leonardo López, Germán Burguener, Leonardo Giovanini. Estudio comparativo de modelos epidemiológicos utilizando técnicas convencionales y modelos basados en individuos. Argentina. Buenos Aires. 2011. Revista. Otro. Congreso. II Congreso Argentino de Bioinformática y Biología Computacional. Asociación Argentina de Bioinformática y Biología Computacional, Universidad Católica de Córdoba, Agencia
- [3] Leonardo López, Burguener Germán, Giovanini Leonardo. A cellular automata to model epidemics. Argentina. Buenos Aires. 2013. Revista. Artículo Completo. Congreso. Congreso Argentino de Informática y Salud. SADIO

### **Trabajos en eventos científicos-tecnológicos no publicados**

- [1] Leonardo López, Germán Burguener, Leonardo Giovanini. Addressing Population Heterogeneity in Epidemic Models Using Agent-Based Models. Argentina. Córdoba. 2013. Congreso. Argentine Symposium on Artificial Intelligence. SADIO
- [2] Agustín Izquierdo; Leonardo López; Leonardo Giovanini. Host-parasite model for philornis larvae and pitangus sulphuratus interactions using a compartmental approach. Argentina. Rosario. 2013. Congreso. 4to congreso argentino de bioinformática y biología computacional. A2B2C
- [3] Leonardo López; Leonardo Giovanini. An agent based model to address population heterogeneity in epidemics. Argentina. Buenos Aires. 2015. Workshop. La matemática como herramienta para entender la biología la biología como fuente de problemas matemáticos. MINCyT

## **6.3. Trabajo Futuro**

El marco conceptual propuesto en esta tesis permite modelar el comportamiento de los individuos través de un diseño modular. Esto permite una gran flexibilidad y adaptabilidad haciendo al modelo especialmente útil en problemas donde la estructura espacial y social es un factor determinante ya que los individuos son considerados como entidades simples que interactúan de forma dinámica. Por otro lado, el comportamiento emergente hace posible representar la dinámica compleja del sistema a través de la agregación de las dinámicas individuales de las unidades de modelado.

A través del marco propuesto en esta tesis, en trabajos futuros se plantea analizar los efectos de los mecanismos de adaptación y aprendizaje de los individuos en la evolución espacio-temporal de sistemas epidemiológicos constituidos por poblaciones heterogéneas. La inclusión de los mecanismos de aprendizaje en los individuos permitirá evaluar los efectos de los factores socioculturales en el comportamiento dinámico de las epidemias. En este sentido, los objetivos específicos de futuros trabajos serán:

- Desarrollar y analizar mecanismos de aprendizaje para el modelo desarrollado en esta tesis;
- Analizar y evaluar los efectos del conocimiento y aprendizaje en el comportamiento emergente del sistema;
- Desarrollar herramientas matemáticas y computacionales para analizar el comportamiento dinámico de los sistemas epidemiológicos;
- Evaluar y validar las herramientas propuestas utilizando el caso de estudio de la bioecología de *Philornis* (*Diptera Muscidae*) y la evolución de la *tuberculosis*.
- Desarrollar herramientas que permitan realizar un análisis teórico del modelo utilizando teoría de juegos y mecánica estadística.
- Desarrollar un modelo de comportamiento individual que permita incorporar mecanismos de aprendizaje.

El uso de mecanismos de auto-organización para modelar el comportamiento de las poblaciones permite describir el comportamiento del sistema como la agregación de los comportamientos de las poblaciones individuales. Esta forma de modelar los comportamientos puede ser descrita desde el punto de vista matemático utilizando herramientas de sistemas multi-agentes como son la teoría de juegos y la mecánica estadística. Esto provee de herramientas matemáticas formales para analizar el comportamiento de las poblaciones y el sistema, cosa que es difícil de hacer en los modelos basados en individuos propuestos hasta el momento.

En este sentido se propone para trabajos futuros implementar sistemas en los cuales las interacciones dinámicas entre los componentes de un sistema auto-organizado sean modeladas utilizando un juego dinámico. Esta herramienta permite una representación precisa, simple y eficiente de los conflictos y las cooperaciones que emergen de la interacción entre los componentes de un sistema. Es decir, el juego dinámico resultante de las interacciones agente-agente y agente-ambiente puede ser implementado a través de mecanismos muy diferentes. El juego dinámico que describe las interacciones entre los agentes, sus objetivos y el ambiente puede transformarse en un problema de optimización distribuido explícito que puede resolverse con algoritmos convencionales. En este caso, el comportamiento de cada agente será descrito por la función objetivo y restricciones del problema de optimización de cada uno de ellos. Dependiendo de la estructura de las funciones objetivo, las restricciones y la información que intercambia cada agente, el problema de optimización distribuido tiene un problema de optimización centralizado equivalente. Esta propiedad permite un análisis preciso y detallado de las propiedades globales del comportamiento emergente del sistema (estabilidad, convergencia).

Estos aspectos tendrán un impacto significativo en el estudio de los problemas epidemiológicos ya que hasta el momento los mismos no han sido abordados integralmente en el



estudio de sistemas complejos, permitiendo una mejor y más precisa descripción de los sistemas estudiados y una mayor y mejor capacidad de análisis de los mismos.



# A. Rutinas implementadas del modelo de redes de contactos

---

## Algoritmo 4 Algoritmo de Inicialización

---

$G = \mathcal{M}_{M \times N}(\mathbb{R})$  {Se define la red de contactos}  
 $N_s$  {Número de actores en la red de contactos susceptibles}  
 $N_i$  {Número de actores en la red de contactos infecciosos}  
 $N_e$  {Número de actores en la red de contactos expuestos}  
 $N_r$  {Número de actores en la red de contactos recuperados}  
**para**  $i \rightarrow N_s$  **hacer**  
     $G(x, y) = S, 1 \leq x \leq M, 1 \leq y \leq N$  {Se definen los actores susceptibles en la grilla}  
**fin para**  
**para**  $i \rightarrow N_i$  **hacer**  
     $G(x, y) = I, 1 \leq x \leq M, 1 \leq y \leq N$  {Se definen los actores infecciosos en la grilla}  
**fin para**  
**para**  $i \rightarrow N_e$  **hacer**  
     $G(x, y) = E, 1 \leq x \leq M, 1 \leq y \leq N$  {Se definen los actores expuestos en la grilla}  
**fin para**  
**para**  $i \rightarrow N_r$  **hacer**  
     $G(x, y) = R, 1 \leq x \leq M, 1 \leq y \leq N$  {Se definen los actores recuperados en la grilla}  
**fin para**

---

---

## Algoritmo 5 Algoritmo de extracción

---

$G_e = \mathcal{M}_{m \times n}(\mathbb{R}) \in G = \mathcal{M}_{M \times N}(\mathbb{R})$  {se define el fragmento a extraer de la red}  
**para**  $g_i \in G_e$  **hacer**  
    Algoritmo 7  
    Algoritmo 8  
    Algoritmo 9  
    Algoritmo 10  
**fin para**

---

**Algoritmo 6** Algoritmo reemplazo

---


$$G_r = \mathcal{M}_{m \times n}(\mathbb{R}) \in G = \mathcal{M}_{M \times N}(\mathbb{R}) \text{ \{se define el fragmento a reemplazar de la red\}}$$

$$G_r = G_e$$


---

**Algoritmo 7** Estado infeccioso

---

**si**  $Estado = I$  **entonces**
**si**  $Vecino = S$  **entonces**
 $Z \sim U[0, 1]$ 
**si**  $Z < \beta/\nu$  **entonces**
 $Estado = E$ 
**fin si**
**fin si**
**fin si**
**si**  $Estado = A$  **entonces**
**si**  $Vecino = S$  **entonces**
 $Z \sim U[0, 1]$ 
**si**  $Z < q\beta/\nu$  **entonces**
 $Estado = E$ 
**fin si**
**fin si**
**fin si**


---

**Algoritmo 8** Estado Expuesto

---

**si**  $Estado = E$  **entonces**
 $Z \sim U[0, 1]$ 
**si**  $Z < \varepsilon\rho$  **entonces**
 $Estado = I$ 
**si**  $Z < \varepsilon(1 - \rho)$  **entonces**
 $Estado = A$ 
**fin si**
**fin si**
**fin si**


---

**Algoritmo 9** Fase de recuperación

---

**si**  $Estado = I$  **o**  $Estado = A$  **entonces**
 $Z \sim U[0, 1]$ 
**si**  $Z < \gamma 1$  **entonces**
 $Estado = R$ 
**fin si**
**fin si**


---

---

**Algoritmo 10** Movimiento de individuos

---

$$Z_1, Z_2 \sim U[-r, r]$$

$$Aux = Estado(i, j)$$

$$Estado(i, j) = Estado(i + Z_1, j + Z_2)$$

$$Estado(i + Z_1, j + Z_2) = Aux$$

---



## B. Rutinas del modelo espacialmente explícito

---

**Algoritmo 11** Se setean los parámetros iniciales del modelo

---

$\mathcal{A} = \{a_i, i = 1, \dots, n\}$  {se crean  $n$  agentes}

$G = [M \times N]$ ,  $M, N \in \mathbb{R}$  {Se define el tamaño de la grilla}

$T = k$ ,  $k \in \mathbb{N}$  {Se define el tiempo de simulación}

---

---

**Algoritmo 12** Estado infeccioso

---

**si**  $a_i = I$  **entonces**

**si**  $a_j = S$ ,  $j \in \sigma_i$  **entonces**

$Z \sim U[0, 1]$  {Donde  $\sigma_i$  es la vecindad de del agente  $a_i$ }

**si**  $Z < \beta/\nu$  **entonces**

$a_i = E$

**fin si**

**fin si**

**fin si**

**si**  $a_i = A$  **entonces**

**si**  $a_j = S$ ,  $j \in \sigma_i$  **entonces**

$Z \sim U[0, 1]$

**si**  $Z < q\beta/\nu$  **entonces**

$a_i = E$

**fin si**

**fin si**

**fin si**

---

---

**Algoritmo 13** Estado Expuesto

---

**si**  $a_i = E$  **entonces**  
 $Z \sim U[0, 1]$   
**si**  $Z < \varepsilon r$  **entonces**  
 $a_i = I$   
**si**  $Z < r$  **entonces**  
 $a_i = A$   
**fin si**  
**fin si**  
**fin si**

---



---

**Algoritmo 14** Fase de recuperación

---

**si**  $a_i = I$  **or**  $a_i = A$  **entonces**  
 $Z \sim U[0, 1]$   
**si**  $Z < \gamma 1$  **entonces**  
 $a_i = R$   
**fin si**  
**fin si**

---



---

**Algoritmo 15** mecanismo de infección

---

**para**  $a_i, i = 1, \dots, n$  **hacer**  
 $U_i = \sum_{j \in N_i} U_j, N_i = \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$  {Se calcula la virulencia de cada individuo teniendo en cuenta la influencia de sus vecinos,  $U_i : \mathbb{R}^m \rightarrow \mathbb{R}$ }  
**fin para**

---



---

**Algoritmo 16** mecanismo de movimiento

---

**para**  $a_i, i = 1, \dots, n$  **hacer**  
 $\dot{q}_i = p_i$   
 $\dot{p}_i = u_i$  {Donde los valores de  $q_i$  y  $p_i$  se pueden determinar de forma aleatoria o por medio de algoritmos de flocking}  
**fin para**

---





## C. Estimación de la probabilidad de error de los modelos propuestos

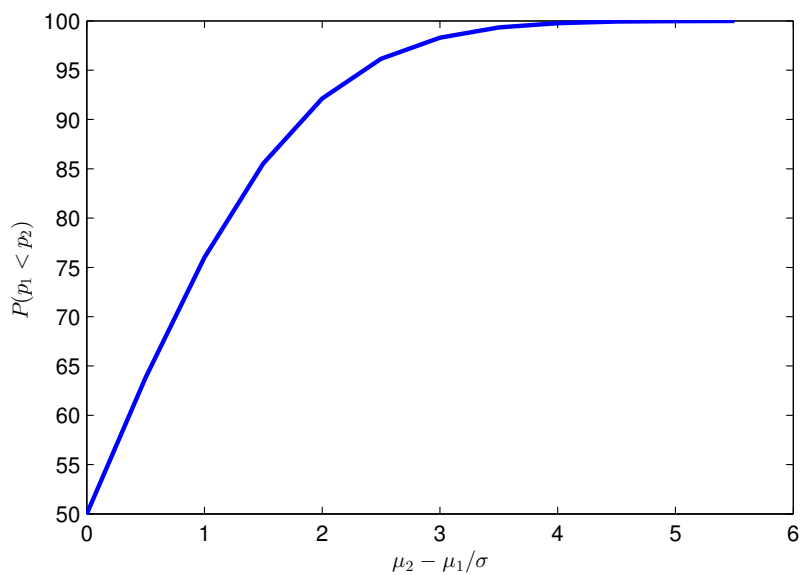
Uno de los aspectos fundamentales en la evaluación de los modelos implementados es estimar la probabilidad del error. Determinar la probabilidad del error es importante por dos razones:

- Para saber si el modelo implementado es útil para los fines propuestos.
- Para comparar su desempeño contra otros modelos.

En los capítulos 3 y 4 se detalla la forma en la que fue llevado adelante el análisis de robustez para los modelos propuestos. En la tabla C-1 se muestra la probabilidad de que el modelo propuesto tenga un mejor desempeño que el modelo clásico.

**Tabla C-1.:** Probabilidad de la significancia estadística del error

$\mu_1 - \mu_2$	$P(p_1 < p_2)$
0	50
0,5	63,81
1	76,02
1,5	85,54
2,0	92,13
2,5	96,16
3,0	98,30
3,5	99,33
4,0	99,77
4,5	99,93
5,0	99,98
5,5	99,99



**Figura C-1.:** Probabilidad de que el error del modelo implementado sea menor al error del modelo clásico



# Bibliografía

- Adams, A. L., Barth-Jones, D. C., Chick, S. E., & Koopman, J. S. (1998). Simulations to evaluate HIV vaccine trial designs. *Simulation*, 71(4), 228-241.
- Albert, R. & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1), 47.
- Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M., & Kadmon, R. (2012). Area-heterogeneity tradeoff and the diversity of ecological communities. *Proceedings of the National Academy of Sciences*, 109(43), 17495-17500.
- Amrhein, C. G. & MacKinnon, R. D. (1988). A micro-simulation model of a spatial labor market. *Annals of the Association of American Geographers*, 78(1), 112-131.
- Anderson, R. M., May, R. M., & Anderson, B. (1992). *Infectious diseases of humans: dynamics and control*. Wiley Online Library.
- Aparicio, J. P. & Pascual, M. (2007). Building epidemiological models from R0: an implicit treatment of transmission in networks. *Proceedings of the Royal Society of London B: Biological Sciences*, 274(1609), 505-512.
- Armano, G. & Martin, A. (2006). The Second International Workshop on Multi-Agent Systems for Medicine, Computational Biology, and Bioinformatics. *Future*.
- Athithan, S., Shukla, V. P., & Biradar, S. R. (2014). Dynamic Cellular Automata Based Epidemic Spread Model for Population in Patches with Movement.
- Auchincloss, A. & Roux, A. D. (2008). A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *American journal of epidemiology*, 168(1), 1.
- Bansal, S., Grenfell, B. T., & Meyers, L. A. (2007). When individual behaviour matters: homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, 4(16), 879-891.
- Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439), 509-512.
- Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11), 3747-3752.
- Barrett, C., Eubank, S., & Smith, J. (2005). If smallpox strikes portland... *Scientific American*, 292(3), 54-61.
- Bauer, A., Beauchemin, C., & Perelson, A. (2009). Agent-based modeling of host-pathogen systems: The successes and challenges. *Information sciences*, 179(10), 1379-1389.

- Berkman, L. F., Kawachi, I., & Glymour, M. (2014). *Social epidemiology*. Oxford University Press.
- Bhopal, R. S. y col. (2002). *Concepts of epidemiology: an integrated introduction to the ideas, theories, principles, and methods of epidemiology*. Oxford University Press Oxford.
- Bian, L. (2004). A conceptual framework for an individual-based spatially explicit epidemiological model. *Environment and Planning B: Planning and Design*, 31(3), 381-395.
- Bonita, R., Beaglehole, R., & Kjellström, T. (2006). *Basic epidemiology*. World Health Organization.
- Botkin, D. (1993). *Jabowa-ii: a computer model of forest growth*. Oxford University Press, NY.
- Byrd, R. H., Gilbert, J. C., & Nocedal, J. (2000). A trust region method based on interior point techniques for nonlinear programming. *Mathematical Programming*, 89(1), 149-185.
- Cao, H. & Zhou, Y. (2012). The discrete age-structured seit model with application to tuberculosis transmission in China. *Mathematical and Computer Modelling*, 55(3), 385-395.
- Carley, K. M., Fridsma, D. B., Casman, E., Yahja, A., Altman, N., Chen, L.-C., ... Nave, D. (2006). BioWar: scalable agent-based model of bioattacks. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 36(2), 252-265.
- Casagrandi, R., Bolzoni, L., Levin, S. A., & Andreasen, V. (2006). The SIRC model and influenza A. *Mathematical biosciences*, 200(2), 152-169.
- Chao, D. L., Halloran, M. E., Obenchain, V. J., & Longini Jr, I. M. (2010). FluTE, a publicly available stochastic influenza epidemic simulation model. *PLoS Comput Biol*, 6(1), e1000656.
- Chierichetti, F., Lattanzi, S., & Panconesi, A. (2011). Rumor spreading in social networks. *Theoretical Computer Science*, 412(24), 2602-2610.
- Chowell, Ammon, Hengartner, & Hyman. (2006). Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: assessing the effects of hypothetical interventions. *Journal of theoretical biology*, 241(2), 193-204.
- Chowell, Nishiura, & Bettencourt. (2007). Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. *Journal of the Royal Society Interface*, 4(12), 155.
- Chowell, Miller, & Viboud. (2008). Seasonal influenza in the United States, France, and Australia: transmission and prospects for control. *Epidemiology and infection*, 136(06), 852-864.
- Cissé, B., El Yacoubi, S., & Tridane, A. (2013). Impact of Neighborhood Structure on Epidemic Spreading by Means of Cellular Automata Approach. *Procedia Computer Science*, 18, 2603-2606.
- Coburn, B., Wagner, B., & Blower, S. (2009). Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1). *BMC medicine*, 7(1), 30.

- Colizza, V., Barrat, A., Barthelemy, M., Valleron, A.-J., & Vespignani, A. (2007). Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Med*, 4(1), e13.
- Control, C. f. D. & (US), P. (2000). *School health index for physical activity and healthy eating: a self-assessment and planning guide: middle school/high school*. The Centers.
- Cooley, P., Ganapathi, L., Ghneim, G., Holmberg, S., Wheaton, W., & Hollingsworth, C. R. (2008). Using influenza-like illness data to reconstruct an influenza outbreak. *Mathematical and computer modelling*, 48(5), 929-939.
- Cordero, A. F., Koopman, H., & van der Helm, F. C. (2004). Mechanical model of the recovery from stumbling. *Biological Cybernetics*, 91(4), 212-220.
- Cristiani, E., Priuli, F. S., & Tosin, A. (2015). Modeling rationality to control self-organization of crowds: an environmental approach. *SIAM Journal on Applied Mathematics*, 75(2), 605-629.
- Crompton, R. (2008). *Class and stratification*. Polity.
- Cucker, F. & Smale, S. (2005). Emergent behaviour in flocks. preliminary version.
- Damasio, A. (1994). *Descartes' error: emotion, reason, and the human brain*. New York: Putnam.
- Davis, S., Trapman, P., Leirs, H., Begon, M., & Heesterbeek, J. (2008). The abundance threshold for plague as a critical percolation phenomenon. *Nature*, 454(7204), 634-637.
- Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: nsga-ii. *IEEE transactions on evolutionary computation*, 6(2), 182-197.
- Dibble, C. & Feldman, P. G. (2004). The GeoGraph 3D computational laboratory: Network and terrain landscapes for RePast. *Journal of artificial societies and social simulation*, 7(1).
- Doran, R. J. & Laffan, S. W. (2005). Simulating the spatial dynamics of foot and mouth disease outbreaks in feral pigs and livestock in Queensland, Australia, using a susceptible-infected-recovered cellular automata model. *Preventive Veterinary Medicine*, 70(1), 133-152.
- Dorogovtsev, S. N. (2010). *Lectures on complex networks*. Oxford University Press Oxford.
- Dunham, J. (2005). An agent-based spatially explicit epidemiological model in MASON. *Journal of Artificial Societies and Social Simulation*, 9(1), 3.
- Dye, C. & Gay, N. (2003). Modeling the SARS epidemic. *Science*, 300(5627), 1884-1885.
- Ellemers, N., Wilke, H., & Van Knippenberg, A. (1993). Effects of the legitimacy of low group or individual status on individual and collective status-enhancement strategies. *Journal of personality and social psychology*, 64(5), 766.
- Epstein, J. (2006). *Generative social science: Studies in agent-based computational modeling*. Princeton Univ Pr.
- Epstein, J., Institution, B., on Social, C., & Dynamics, E. (2002). *Toward a containment strategy for smallpox bioterror: an individual-based computational approach*. Citeseer.

- Erdős, P. & Rényi, A. (1959). On random graphs, I. *Publicationes Mathematicae (Debrecen)*, 6, 290-297.
- Eubank, S., Guclu, H., Kumar, V. A., Marathe, M. V., Srinivasan, A., Toroczkai, Z., & Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988), 180-184.
- Eyob, T., Tshilidzi, M., & Conrad, M. (2005). Agent-based modeling: A case study in HIV epidemic. En *Proceedings of the Fourth International Conference on Hybrid Intelligent Systems*.
- Fang, Y. (2011). Asymptotic equivalence between cross-validations and akaike information criteria in mixed-effects models. *Journal of data science*, 9(1), 15-21.
- Ferguson, Cummings, Cauchemez, Fraser, Riley, Meeyai, ... Burke. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature*, 437(7056), 209-214.
- Ferguson, Cummings, Fraser, Cajka, Cooley, & Burke. (2006). Strategies for mitigating an influenza pandemic. *Nature*, 442(7101), 448-452.
- Ferrari, A., Meyer, J., Scardaci, V., Casiraghi, C., Lazzeri, M., Mauri, F., ... Roth, S., y col. (2006). Raman spectrum of graphene and graphene layers. *Physical review letters*, 97(18), 187401.
- Firestone, S. M., Christley, R. M., Ward, M. P., & Dhand, N. K. (2012). Adding the spatial dimension to the social network analysis of an epidemic: Investigation of the 2007 outbreak of equine influenza in Australia. *Preventive Veterinary Medicine*, 106(2), 123-135. Special Issue: {SVEPM} 2011 - Current Advances in Understanding the Spread and Control of Animal Diseases 2011 Society of Veterinary Epidemiology and Preventive Medicine conference. doi:10.1016/j.prevetmed.2012.01.020
- Fu, S. (2002). Modelling epidemic spread using cellular automata. *University of Western Australia*.
- Fuks, H. & Lawniczak, A. T. (2001). Individual-based lattice model for spatial spread of epidemics. *Discrete Dynamics in Nature and Society*, 6(3), 191-200.
- Galea, S. (2013). An argument for a consequentialist epidemiology. *American journal of epidemiology*, kwt172.
- Galvani, Ferrini, & Ottani. (2004). Natriuretic peptides for risk stratification of patients with acute coronary syndromes. *European journal of heart failure*, 6(3), 327-333.
- Galvani & May. (2005a). Epidemiology: dimensions of superspreading. *Nature*, 438(7066), 293-295.
- Galvani & May. (2005b). Epidemiology: dimensions of superspreading. *Nature*, 438(7066), 293-295.
- Ghani, A. C., Swinton, J., & Garnett, G. P. (1997). The role of sexual partnership networks in the epidemiology of gonorrhoea. *Sexually transmitted diseases*, 24(1), 45-56.
- Gornick, J. C. & Meyers, M. K. (2003). *Families that work: Policies for reconciling parenthood and employment*. Russell Sage Foundation.



- Grenfell, B. & Harwood, J. (1997). (Meta) population dynamics of infectious diseases. *Trends in ecology & evolution*, 12(10), 395-399.
- Grimm, V. & Railsback, S. F. (2013). *Individual-based modeling and ecology*. Princeton university press.
- Halloran, E., Ferguson, N., Eubank, S., Longini, I. M., Cummings, D. A., Lewis, B., ... Germann, T. C., y col. (2008). Modeling targeted layered containment of an influenza pandemic in the United States. *Proceedings of the National Academy of Sciences*, 105(12), 4639-4644.
- Haynes, R. B. (2012). *Clinical epidemiology: how to do clinical practice research*. Lippincott williams & wilkins.
- Helbing, D., Yu, W., & Rauhut, H. (2011). Self-organization and emergence in social systems: modeling the coevolution of social environments and cooperative behavior. *The Journal of Mathematical Sociology*, 35(1-3), 177-208.
- Hethcote. (2000). The mathematics of infectious diseases. *SIAM review*, 599-653.
- Hethcote & Yorke. (1984). Lecture notes in biomathematics.
- Hethcote, Zhien, & Shengbing. (2002). Effects of quarantine in six endemic models for infectious diseases. *Mathematical Biosciences*, 180(1), 141-160.
- Holmes, E. E. (1997). Basic epidemiological concepts in a spatial context. *Spatial ecology*, 111-136.
- Hong, S., Yang, H., Zhao, T., & Ma, X. (2016). Epidemic spreading model of complex dynamical network with the heterogeneity of nodes. *International Journal of Systems Science*, 47(11), 2745-2752.
- Houck, C. R., Joines, J., & Kay, M. G. (1995). A genetic algorithm for function optimization: a matlab implementation. *NCSU-IE TR*, 95(09).
- Hoya y col. (2007). Modeling epidemics using cellular automata. *Applied mathematics and computation*, 186(1), 193-202.
- Huang, C.-Y., Sun, C.-T., Hsieh, J.-L., & Lin, H. (2004). Simulating SARS: Small-world epidemiological modeling and public health policy assessments. *Journal of Artificial Societies and Social Simulation*, 7(4).
- Iantovics, B. (2010). Cognitive Medical Multiagent Systems. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 1(1), pp-12.
- Keeling. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1421), 859-867.
- Keeling. (2000). Metapopulation moments: coupling, stochasticity and persistence. *Journal of Animal Ecology*, 69(5), 725-736.
- Keeling, Woolhouse, Shaw, Matthews, Chase-Topping, Haydon, ... Grenfell. (2001). Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*, 294(5543), 813-817.

- Kikuchi, S., Tominaga, D., Arita, M., Takahashi, K., & Tomita, M. (2003). Dynamic modeling of genetic networks using genetic algorithm and S-system. *Bioinformatics*, 19(5), 643-650.
- Klov Dahl, A. S., Dhofier, Z., Oddy, G., O'Hara, J., Stoutjesdijk, S., & Whish, A. (1977). Social Networks in an Urban Area: First Canberra Study1. *Journal of Sociology*, 13(2), 169-172.
- Koopman, J. S. & Lynch, J. W. (1999). Individual causal models and population system models in epidemiology. *American Journal of Public Health*, 89(8), 1170-1174.
- Kothapalli, K., Kirkness, E., Natale, L., & Meyers-Wallen, V. (2003). Exclusion of PISRT1 as a candidate locus for canine Sry-negative XX sex reversal. *Animal genetics*, 34(6), 467-469.
- Kretzschmar, M. & Morris, M. (1996). Measures of concurrency in networks and the spread of infectious disease. *Mathematical biosciences*, 133(2), 165-195.
- Lambert, J.-C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., ... Beecham, G. W., y col. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature genetics*, 45(12), 1452-1458.
- Lee. (2011). Bounded rationality and the emergence of simplicity amidst complexity. *Journal of Economic Surveys*, 25(3), 507-526.
- Lee, Kim, & Cho. (2012). Design of activation functions for inference of fuzzy cognitive maps: application to clinical decision making in diagnosis of pulmonary infection. *Healthcare informatics research*, 18(2), 105-114.
- Liljeros, F., Edling, C. R., Amaral, L. A. N., Stanley, H. E., & Åberg, Y. (2001). The web of human sexual contacts. *Nature*, 411(6840), 907-908.
- Li & Muldowney. (1995). Global stability for the SEIR model in epidemiology. *Mathematical Biosciences*, 125(2), 155-164.
- Li, Graef, Wang, & Karsai. (1999). Global dynamics of a SEIR model with varying total population size. *Mathematical biosciences*, 160(2), 191-213.
- Li & Jin. (2005). Global stability of a SEIR epidemic model with infectious force in latent, infected and immune period. *Chaos, Solitons & Fractals*, 25(5), 1177-1184.
- Li, Wang, & Jin. (2006). Global stability of an SEIR epidemic model with constant immigration. *Chaos, Solitons & Fractals*, 30(4), 1012-1019.
- Li & Guo. (2012). Using Cellular Automata to study the effect of competition for epidemic diseases. *Procedia Environmental Sciences*, 13, 1010-1018.
- Lloyd-Smith, J., Schreiber, S., Kopp, P., & Getz, W. (2005). Superspreading and the effect of individual variation on disease emergence. *NATURE-LONDON-*, 438(7066), 355.
- Longini, I. M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D. A., & Halloran, M. E. (2005). Containing pandemic influenza at the source. *Science*, 309(5737), 1083-1087.

- López, L., Burguener, G., & Giovanini, L. (2014). Addressing population heterogeneity and distribution in epidemics models using a cellular automata approach. *BMC research notes*, 7(1), 1.
- López, L., Izquierdo, A., Manzoli, D., Beldoménico, P., & Giovanini, L. (2016). A myiasis model for *philornis torquans* (diptera: muscidae) and *pitangus sulphuratus* (passeriformes: tyrannidae). *Ecological Modelling*, 328, 62-71.
- Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., & Balan, G. (2005). Mason: A multiagent simulation environment. *Simulation*, 81(7), 517-527.
- Luscombe, N. M., Babu, M. M., Yu, H., Snyder, M., Teichmann, S. A., & Gerstein, M. (2004). Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*, 431(7006), 308-312.
- Malcovati, P., Brigati, S., Francesconi, F., Maloberti, F., Cusinato, P., & Baschiroto, A. (2003). Behavioral modeling of switched-capacitor sigma-delta modulators. *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, 50(3), 352-364.
- Mao, L. & Bian, L. (2010). A dynamic network with individual mobility for designing vaccination strategies. *Transactions in GIS*, 14(4), 533-545.
- McLean, H. Q., Thompson, M. G., Sundaram, M. E., Kieke, B. A., Gaglani, M., Murthy, K., ... Raviotta, J. M., y col. (2014). Influenza vaccine effectiveness in the United States during 2012–2013: variable protection by age and virus type. *Journal of Infectious Diseases*, jiu647.
- Mei, S., Zhu, Y., Qiu, X., Zhou, X., Zu, Z., Boukhanovsky, A., & Sloat, P. M. (2014). Individual decision making can drive epidemics: a fuzzy cognitive map study. *IEEE Transactions on Fuzzy Systems*, 22(2), 264-273.
- Mikler, A., Venkatachalam, S., & Abbas, K. (2005). Modeling infectious diseases using global stochastic cellular automata. *Journal of Biological Systems*, 13(4), 421-439.
- Mniszewski, S. M., Del Valle, S. Y., Stroud, P. D., Riese, J. M., & Sydoriak, S. J. (2008). Epi-SimS simulation of a multi-component strategy for pandemic influenza. En *Proceedings of the 2008 Spring simulation multiconference* (pp. 556-563). Society for Computer Simulation International.
- Mogilner, A., Edelstein-Keshet, L., Bent, L., & Spiros, A. (2003). Mutual interactions, potentials, and individual distance in a social aggregation. *Journal of mathematical biology*, 47(4), 353-389.
- Monasson, R. (1999). Diffusion, localization and dispersion relations on “small-world” lattices. *The European Physical Journal B-Condensed Matter and Complex Systems*, 12(4), 555-567.
- Mooij, W. M., Bennetts, R. E., Kitchens, W. M., & DeAngelis, D. L. (2002). Exploring the effect of drought extent and interval on the florida snail kite: interplay between spatial and temporal scales. *Ecological Modelling*, 149(1), 25-39.

- Moreno, Y., Pastor-Satorras, R., & Vespignani, A. (2002). Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B-Condensed Matter and Complex Systems*, 26(4), 521-529.
- Moreno, Y., Nekovee, M., & Pacheco, A. F. (2004). Dynamics of rumor spreading in complex networks. *Physical Review E*, 69(6), 066130.
- Murillo, L. N., Murillo, M. S., & Perelson, A. S. (2013). Towards multiscale modeling of influenza infection. *Journal of theoretical biology*, 332, 267-290.
- Nishiura, Castillo-Chavez, Safan, Chowell, y col. (2009). Transmission potential of the new influenza A (H1N1) virus and its age-specificity in Japan. *Euro Surveill*, 14(22), 19227.
- Nuno, M., Feng, Z., Martcheva, M., & Castillo-Chavez, C. (2005). Dynamics of two-strain influenza with isolation and partial cross-immunity. *SIAM Journal on Applied Mathematics*, 65(3), 964-982.
- Ohmit, S. E., Thompson, M. G., Petrie, J. G., Thaker, S. N., Jackson, M. L., Belongia, E. A., ... Spencer, S. M., y col. (2014). Influenza vaccine effectiveness in the 2011–2012 season: protection against each circulating virus and the effect of prior vaccination on estimates. *Clinical infectious diseases*, 58(3), 319-327.
- Olfati-Saber, R. (2006). Flocking for multi-agent dynamic systems: Algorithms and theory. *IEEE Transactions on automatic control*, 51(3), 401-420.
- Ovaskainen, O., Finkelshtein, D., Kutoviy, O., Cornell, S., Bolker, B., & Kondratiev, Y. (2014). A general mathematical framework for the analysis of spatiotemporal point processes. *Theoretical Ecology*, 7(1), 101-113.
- Pastor-Satorras, R., Vázquez, A., & Vespignani, A. (2001). Dynamical and correlation properties of the Internet. *Physical review letters*, 87(25), 258701.
- Picard, R. W. & Picard, R. (1997). *Affective computing*. MIT press Cambridge.
- Pires, B. & Crooks, A. T. (2017). Modeling the emergence of riots: a geosimulation approach. *Computers, Environment and Urban Systems*, 61, 66-80.
- Porta, M. S., Greenland, S., Hernán, M., dos Santos Silva, I., & Last, J. M. (2014). *A dictionary of epidemiology*. Oxford University Press.
- Poutanen, S. M., Low, D. E., Henry, B., Finkelstein, S., Rose, D., Green, K., ... Ayers, M., y col. (2003). Identification of severe acute respiratory syndrome in Canada. *New England Journal of Medicine*, 348(20), 1995-2005.
- Powell, M. J. (1978). A fast algorithm for nonlinearly constrained optimization calculations. En *Numerical analysis* (pp. 144-157). Springer.
- Radivojac, P., Clark, W. T., Oron, T. R., Schnoes, A. M., Wittkop, T., Sokolov, A., ... Ben-Hur, A., y col. (2013). A large-scale evaluation of computational protein function prediction. *Nature methods*, 10(3), 221-227.
- Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH computer graphics*, 21(4), 25-34.

- Reynolds, C. W. (1999). Steering behaviors for autonomous characters. En *Game developers conference* (Vol. 1999, pp. 763-782).
- Reynolds, C. W. (2000). Interaction with groups of autonomous characters. En *Game developers conference* (Vol. 2000, pp. 449-460).
- Rhodes, C. & Anderson, R. M. (1997). Epidemic thresholds and vaccination in a lattice model of disease spread. *Theoretical Population Biology*, 52(2), 101-118.
- Rothman, K. J. (2012). *Epidemiology: an introduction*. Oxford University Press.
- Samsuzzoha, M., Singh, M., & Lucy, D. (2012). A numerical study on an influenza epidemic model with vaccination and diffusion. *Applied Mathematics and Computation*, 219(1), 122-141.
- Satsuma, J., Willox, R., Ramani, A., Grammaticos, B., & Carstea, A. (2004). Extending the SIR epidemic model. *Physica A: Statistical and Theoretical Physics*, 336(3-4), 369-375.
- Smallman-Raynor, M. & Cliff, A. D. (2001). Epidemic diffusion processes in a system of US military camps: Transfer diffusion and the spread of typhoid fever in the Spanish-American War, 1898. *Annals of the Association of American Geographers*, 91(1), 71-91.
- Solomonoff, R. & Rapoport, A. (1951). Connectivity of random nets. *The bulletin of mathematical biophysics*, 13(2), 107-117.
- Song, B., Jiang, G. P., & Song, Y. R. (2014). Epidemic Dynamics in Weighted Adaptive Networks. En *P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2014 Ninth International Conference on* (pp. 69-73). doi:10.1109/3PGCIC.2014.56
- Stein, R. A. (2011). Super-spreaders in infectious diseases. *International Journal of Infectious Diseases*, 15(8), e510-e513.
- Streicker, D. G., Fenton, A., & Pedersen, A. B. (2013). Differential sources of host species heterogeneity influence the transmission and control of multihost parasites. *Ecology letters*, 16(8), 975-984.
- Symonds, M. R. & Moussalli, A. (2011). A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using akaike's information criterion. *Behavioral Ecology and Sociobiology*, 65(1), 13-21.
- Szklo, M. & Nieto, F. J. (2014). *Epidemiology: beyond the basics*. Jones & Bartlett Publishers.
- Teweldemedhin, E., Marwala, T., & Mueller, C. (2004). Agent-based modelling: a case study in HIV epidemic. En *Hybrid Intelligent Systems, 2004. HIS'04. Fourth International Conference on* (pp. 154-159). IEEE.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic geography*, 46, 234-240.
- Turner, M. G., Wu, Y., Romme, W. H., & Wallace, L. L. (1993). A landscape simulation model of winter foraging by large ungulates. *Ecological Modelling*, 69(3), 163-184.
- Van der Ploeg, C. P., Van Vliet, C., De Vlas, S. J., Ndinya-Achola, J. O., Fransen, L., Van Oortmarssen, G. J., & Habbema, J. D. F. (1998). STDSIM: A microsimulation model for decision support in STD control. *Interfaces*, 28(3), 84-100.

- Wang, Birol, & Hatzimanikatis. (2004). Metabolic control analysis under uncertainty: framework development and case studies. *Biophysical journal*, 87(6), 3750-3763.
- Wang, Zeng, Zheng, & Wang. (2008). Analyzing online media as complex network. *Complex Systems and Complexity Science*, 6(3), 11-21.
- Wang & Chen. (2013). A virus spread model based on cellular automata in weighted scale-free networks. En *Proceedings of the 2012 International Conference on Communication, Electronics and Automation Engineering* (pp. 1169-1175). Springer.
- Watts & Strogatz. (1998). Collective dynamics of 'small-world' networks. *nature*, 393(6684), 440-442.
- Watts, Muhamad, Medina, & Dodds. (2005). Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proceedings of the National Academy of Sciences of the United States of America*, 102(32), 11157-11162.
- Welch, G., Chick, S. E., & Koopman, J. (1998). Effect of concurrent partnerships and sex-act rate on gonorrhea prevalence. *Simulation*, 71(4), 242-249.
- White, S., del Rey, A., & Sánchez, G. (2007). Modeling epidemics using cellular automata. *Applied Mathematics and Computation*, 186(1), 193-202.
- Wiegand, T., Naves, J., Stephan, T., & Fernandez, A. (1998). Assessing the risk of extinction for the brown bear (*ursus arctos*) in the cordillera cantabrica, Spain. *Ecological Monographs*, 68(4), 539-570.
- Wolff, W. (1994). An individual-oriented model of a wading bird nesting colony. *Ecological Modelling*, 72(1), 75-114.
- Woodward, M. (2013). *Epidemiology: study design and data analysis*. CRC Press.
- Worm, B. & Myers, R. A. (2003). Meta-analysis of cod-shrimp interactions reveals top-down control in oceanic food webs. *Ecology*, 84(1), 162-173.
- Xiao-Juan, J. D.-P. B. & Wei-Ren, Y. Y.-X. S. (2007). Research on emotion theory and the decision models based on emotion. *Computer Science*, 4, 042.
- Yang, J., Qiu, Z., & Li, X. (2014). Global stability of an age-structured cholera model. *Mathematical biosciences and engineering: MBE*, 11(3), 641-665.
- Yergens, D., Hiner, J., Denzinger, J., & Noseworthy, T. (2006). Multi agent simulation system for rapidly developing infectious disease models in developing countries. En *The Second International Workshop on Multi-Agent Systems for Medicine, Computational Biology, and Bioinformatics* (p. 104). Citeseer.
- Zhang, X.-B., Huo, H.-F., Xiang, H., & Meng, X.-Y. (2014). Dynamics of the deterministic and stochastic SIQS epidemic model with non-linear incidence. *Applied Mathematics and Computation*, 243, 546-558.
- Zhang & Teng. (2008). Global asymptotic stability of a delayed SEIRS epidemic model with saturation incidence. *Chaos, Solitons & Fractals*, 37(5), 1456-1468.
- Zheng, X., Li, H., & Sun, A. (2008). Exploring social dynamics in online bookmarking systems. En *International Conference on Intelligence and Security Informatics* (pp. 390-391). Springer.

**Doctorado en Ingeniería**  
**mención inteligencia computacional, señales y sistemas**

Título de la obra:

**Modelado de epidemias  
utilizando  
sistemas auto-organizados**

Autor: Leonardo López

Lugar: Santa Fe, Argentina

Palabras Claves:

Modelos epidémicos, autómata, agente  
sistemas auto-organizados, red de contactos  
autómata celular, flocking, comportamiento  
gripe, epidemia