
UNIVERSIDAD DE GUADALAJARA

FACULTAD DE CIENCIAS BIOLÓGICAS



SIMULACION POR COMPUTADORA DE PROCESOS BIOLÓGICOS
MEDIANTE LA SOLUCION DE SUS MODELOS MATEMATICOS.

TESIS PROFESIONAL

QUE PARA OBTENER EL TITULO DE
LICENCIADO EN BIOLOGIA

P R E S E N T A :

GERARDO VALENCIA MAGALLANES

GUADALAJARA, JAL. NOVIEMBRE 1993



UNIVERSIDAD DE GUADALAJARA
FACULTAD DE CIENCIAS BIOLÓGICAS

Expediente

Número

Sección

C. GERARDO VALENCIA MAGALLANES

P R E S E N T E . -

Manifestamos a usted, que con esta fecha ha sido aprobado el tema de tesis "SIMULACION POR COMPUTADORA DE PROCESOS BIOLÓGICOS MEDIANTE LA SOLUCION DE SUS MODELOS MATEMÁTICOS" para obtener la Licenciatura en Biología.

Al mismo tiempo le informamos que ha sido aceptado como Director de dicha Tesis el Dr. Victor González Alvarez.

A T E N T A M E N T E
"PIENSA Y TRABAJA"

Guadalajara, Jal., 14 de Septiembre de 1993

EL DIRECTOR

DR. EULOGIO PIMIENTA BARRIOS



FACULTAD DE
CIENCIAS BIOLÓGICAS

EL SECRETARIO

M. EN C. MA. GEORGINA GUZMAN GODINEZ

c.c.p.- El Dr. Victor González Alvarez, Directora de Tesis.-pte.
c.c.p.- El expediente del alumno

EPB/MGGG/cglr.

Agradecimientos

Este trabajo de tesis fué posible gracias a la intervención de varias personas que coadyuvaron en diferentes aspectos de su realización. Agradezco de manera muy particular al Dr. Víctor González Álvarez por su tutoría y porque con su dirección fué factible esta incursión multidisciplinaria en cuestiones biológicas. Agradezco a Gabriela Coronado Quirarte su colaboración en la preparación de los elementos esquemáticos del documento y en el diseño de pantallas del programa. Agradezco, finalmente, a todas las personas e instituciones que hicieron posible al autor llegar al grado académico para cuya otorgación se realizó este trabajo de tesis.

**Simulación por computadora
de procesos biológicos mediante
la solución de sus modelos matemáticos**

Gerardo Valencia Magallanes

Facultad de Ciencias Biológicas
Universidad de Guadalajara

TABLA DE CONTENIDOS

	pág.
Resumen	4
1. Introducción	5
2. Antecedentes	6
2.1 Modelos	6
2.1.1 Modelos conceptuales	6
2.1.2 Modelos matemáticos	9
2.2 Simulación	11
2.2.1 Simulación por computadora	11
2.2.2 Desarrollo de sistemas multicomponentes	13
2.3 Relación de los modelos y la simulación con el proceso de la investigación	16
3. Justificación	19
4. Objetivo	22
5. Metodología	23
5.1 Modelos analíticos basados en ecuaciones diferenciales	25
5.1.1 Modelo de crecimiento microbial de Monod	27
5.1.1.1 Modelos biomatemáticos y fenomenológicos para el estudio de poblaciones de microorganismos	27
5.1.1.2 Modelo de Monod	28
5.1.2 Difusión pasiva a través de una membrana	33
5.1.3 Ley del enfriamiento de cuerpos de Newton	34
5.1.4 Modelo de crecimiento de peces de Von Bertalanffy	35
5.1.4.1 Teoría del crecimiento animal	35
5.1.4.2 Modelo simple del crecimiento de peces	38
5.1.5 Modelo compartamental simple	39
5.1.6 Modelo de poblaciones de Verhulst-Pearl	41
5.1.6.1 Ecuación de Verhulst-Pearl	41
5.1.7 Modelo de depredación de Lotka-Volterra	44
5.1.8 Modelo de epidemias de Bailey	47
5.1.8.1 Modelos de epidemias	47
5.1.8.2 Modelo general de epidemias	48
5.1.8.3 La curva epidémica	50
5.1.8.4 Modelo de epidemia simple	50
5.2 Métodos numéricos	52
5.2.1 Método de Euler	52

5.2.2 Método de Heun	53
5.2.3 Métodos Runge-Kutta	53
5.2.3.1 Método Runge-Kutta de tercer orden	54
5.2.3.2 Método Runge-Kutta de cuarto orden	54
5.3 Aspectos generales de la programación estructurada	57
5.3.1 Lenguaje de programación C++	57
5.3.2 Modularidad	60
5.3.3 Bloques básicos para la construcción de un programa	61
5.3.4 Pruebas y mantenimiento	64
6. Resultados	65
7. Discusión	67
8. Conclusiones	70
Literatura citada	71

INDICE DE FIGURAS

Figura	pág.
1. Representación de sistemas multicomponentes	15
2. Papel de la simulación en el proceso de la investigación científica	18
3. Solución de problemas por computadora	21
4. Graficación de variables	25
5. Comportamiento de un cultivo de microorganismos	28
6. Sistema de un sólo compartimiento	39
7. Líneas isoclinas de las variables presa y depredador	46
8. Relación de la función <i>main()</i> con el programa estructurado	59

RESUMEN

En el actual contexto de creciente introducción del equipo de cómputo electrónico en las actividades de investigación científica, ha sido posible la realización de simulaciones de fenómenos biológicos basadas en el procesamiento de sus modelos cuantitativos. La construcción del modelo matemático conlleva una postura sumamente crítica que permite la identificación de los elementos fundamentales que conforman al sistema biológico y la manera en que éstos se encuentran interrelacionados para producir su comportamiento. Este trabajo tuvo por objetivo simular procesos biológicos de interés científico mediante la utilización de técnicas de programación estructurada, con la finalidad de apoyar los procesos de investigación y docencia de esta área del conocimiento. Para la realización de simulaciones por computadora se requiere de implementar el modelo matemático del fenómeno de interés en un lenguaje de programación. Con este fin, se utilizó el lenguaje Turbo C++. El programa de cómputo debe de incluir los métodos numéricos adecuados para resolver al modelo, dado un conjunto de condiciones iniciales y parámetros particulares. El método numérico empleado fue el de Runge-Kutta de cuarto orden, que está diseñado para resolver problemas de valor inicial. Debe de incluir, además, medios de representación gráfica de los resultados que permitan la observación del comportamiento manifestado por el sistema bajo las condiciones especificadas. El programa diseñado realiza gráficas bidimensionales que relacionan la variable dependiente contra el tiempo, o dos variables dependientes de un sistema de ecuaciones. Los fenómenos abordados por el programa para su simulación oscilan entre algunos de carácter fisicoquímico (difusión transmembranal y pérdida de calor) hasta otros de interés ecológico (poblaciones e interacciones tróficas) y pretenden representar a la amplitud de aplicaciones de las técnicas utilizadas. La simulación por computadora permite manipular teóricamente al proceso estudiado con lo que pueden ser obtenidas predicciones racionales acerca de su comportamiento y, por lo tanto, apoya la toma de decisiones en el direccionamiento de la investigación y la conservación de recursos.

1. Introducción

En los enfoques modernos de las ciencias, los investigadores se caracterizan cada vez más por su habilidad para el uso de equipo de cómputo electrónico para el análisis de datos y simulación (Keen y Spain, 1992). En las ciencias de la vida, esta corriente ha aportado avances insospechados y aún queda mucho por explorar. Esto se debe a que en la biología aun existen sistemas cuyo estudio cuenta con un gran potencial de desarrollo matemático (López, 1985).

En términos sencillos, un sistema puede ser definido como una colección de entidades, que actúan e interactúan juntos hacia la realización de un fin lógico (Law y Kelton, 1991). Los sistemas vivientes están compuestos por interacciones de procesos químicos y físicos que pueden ser descritos en términos matemáticos; sin embargo, debido a la inmensa complejidad intrínseca de tales interacciones, los procesos biológicos se han resistido a arrojar resultados exitosos a la luz de los análisis matemáticos tradicionales usados por los químicos y físicos. Fué sólo hasta la segunda mitad de este siglo, con el advenimiento de las computadoras digitales, que los estudiosos de la biología tuvieron ante sí la posibilidad de abordar formalmente estos complejos sistemas (Karplus y Petsko, 1990).

Las computadoras son ampliamente utilizadas en la actualidad en la investigación biológica en la planeación de experimentos, en el procesamiento y análisis de datos, en la búsqueda de bibliografía y la preparación de documentos para la publicación de resultados experimentales. Estos usos han llegado a ser asequibles para la mayoría de los investigadores tras la introducción generalizada de las microcomputadoras personales. Ahora, en una naciente y constante expansión se encuentra la simulación de procesos biológicos basada en modelos cuantitativos. Tales simulaciones son importantes para la formulación y mejoramiento de los modelos conceptuales extraídos del análisis empírico de los fenómenos biológicos (Keen y Spain, 1992).

2. Antecedentes

2.1 Modelos

2.1.1 Modelos conceptuales

Los modelos y otros tipos de analogías han jugado siempre un papel importante en el pensamiento científico. Un modelo es cualquier representación de un sistema real y puede estar configurado con respecto a la estructura o a la función de tal sistema. El modelo puede involucrar el uso de palabras, diagramas, notación matemática o estructuras físicas que representen al sistema. El término "modelo" tiene un significado similar al de "concepto", "hipótesis" o "analogía"; debido a que ningún modelo puede representar totalmente al sistema real en todos sus detalles, éste involucra siempre cierto grado de simplificación o abstracción (Naylor *et al.*, 1986).

Un modelo considera siempre los elementos que de manera general sean concernientes al fenómeno estudiado y pretende siempre evitar los detalles particulares. Los modelos se convierten, por lo tanto, en representaciones abstractas del sistema real. Los modelos son, adicionalmente, medios más rápidos y baratos para abordar el estudio de un proceso. Otra ventaja ganada es la generalidad, pues con una representación directa sólo se abarca una situación particular del problema (López, 1985).

A lo largo del universo de estudio de la biología, el medio esencial para desentrañar los principios que rigen a estos sistemas ha sido la formulación de modelos conceptuales, desde el nivel molecular hasta el ecológico (Keen y Spain, 1992).

El átomo y la molécula han sido ampliamente utilizados como modelos de unidad estructural elemental y unidad de reactividad química, respectivamente. El estudio estructural de las biomoléculas parte de patrones bidimensionales obtenidos con la técnica de difracción de rayos X; a través del análisis de los datos provistos por estos patrones gráficos (en representación de la molécula real) fueron descubiertas las estructuras de la ubicua α -hélice proteica, por Linus Pauling, y la

doble hélice del ADN, por Watson y Crick; los modelos se vuelven exitosos al poder explicar las características observadas en el sistema real o predecir comportamientos específicos de éste (Morrison y Boyd, 1987).

En bioquímica, se encuentra una gran variedad de modelos conceptuales, tales como los que describen la acción de las enzimas, que incluyen tanto la forma en que interactúan con su sustrato así como sus respuestas a la temperatura y pH. Las vías metabólicas que se conocen como glucólisis, ciclo de Krebs y el ciclo del carbono en la fotosíntesis, son realmente modelos conceptuales que son permanentemente actualizados en base a los últimos hallazgos experimentales. Estos modelos dan lugar a cuestionamientos acerca del sistema real, lo que permite el direccionamiento de la investigación y la concatenante planeación de experimentos. Estos modelos son significativamente importantes para la toma de decisiones acerca de tratamientos médicos, planeación y predicción de la acción de fármacos y efectos de sustancias tóxicas, y métodos para la nutrición de cualquier tipo de organismo (Lehninger, 1975).

Tras los modelos conceptuales del gen hay una larga historia de teorías propuestas, descartadas, modificadas y refinadas. El gen es considerado actualmente como una porción de ADN cromosómico que codifica para una sola cadena polipeptídica y, pese a los grandes avances en genética, su estudio está dominado por modelos conceptuales basados principalmente sobre la observación de efectos genéticos (Solomon *et al.*, 1987).

La obtención de los modelos tridimensionales de células, tejidos y órganos se debe al procesamiento de resultados obtenidos por diversas técnicas observacionales, que al ser acoplados, producen una síntesis compleja del sistema (Ibid.).

En ecología, los estudios de poblaciones y depredación, por ejemplo, producen modelos que coadyuvan a la comprensión del transporte de energía y materia dentro de los ecosistemas, su equilibrio y, en general toda su dinámica (Hutchinson, 1978).

Resulta clara la importancia del desarrollo de modelos en biología. Al aplicarse a su diseño, el investigador descubre detalles antes inadvertidos, aparecen ante su razón nuevas e interesantes relaciones entre los elementos componentes del

sistema real; por tanto, el modelo aporta su más significativo beneficio en el momento en que lleva al investigador a la realización del trabajo analítico requerido para desarrollarlo (Keen y Spain, 1992).

Los modelos conceptuales generalmente carecen de rigor. Pueden ser imprecisos e interpretados de diferentes maneras por diferentes personas. Para evitar este problema, el modelo conceptual debe ser expresado en una forma más precisa que le reste ambigüedad y pueda ser evaluada y validada. Una forma de modelo conceptual que provee estas ventajas es el modelo matemático (López, 1985).

2.1.2 Modelos matemáticos

Las distintas ciencias y ramas del conocimiento comienzan generalmente por una etapa intuitiva durante la cual el pensamiento y el lenguaje se utilizan para formar imágenes vagas, conceptos imprecisos y analogías más o menos superficiales de los procesos que se busca comprender. Es la etapa en la cual no se conocen aún los elementos esenciales de una situación. Es una etapa de búsqueda, es el proceso de construcción de una abstracción adecuada del problema con la que se pueda empezar a trabajar. La vaguedad de las imágenes y la imprecisión de los lenguajes humanos parecen estar adaptados a la flexibilidad que se requiere para explorar en direcciones muy diversas, para no caer en construcciones teóricas prematuras que no reflejen adecuadamente la realidad. De aquí se deriva que a medida que se precisan los conceptos y se hacen resaltar los elementos fundamentales de la situación, mediante la exploración conceptual ligada a la experimentación, el lenguaje se precisa también. Se llega a la etapa de los lenguajes especializados de las ciencias. Se crean nuevas palabras para designar conceptos y relaciones, más o menos abstractas, y se llega a modificar el significado de muchas palabras existentes e incluso a alterar ciertas formas gramaticales. En ocasiones se suele complementar éstas con la introducción de algunos símbolos especiales. Estos lenguajes pueden llegar a tener una gran precisión, sin salirse de la estructura básica del lenguaje común (López, 1973).

Al haber desglosado los elementos esenciales se puede empezar a construir modelos. Estos consisten generalmente en una simulación de la situación real en condiciones más sencillas y favorables. Esto permite desarrollar una observación y experimentación más racionales y sistemáticas, que a su vez permiten desarrollar nuevos conceptos y relaciones y precisar los antiguos. Permite también empezar a descubrir las relaciones entre las distintas ciencias, y a intercambiar conceptos y relaciones entre una ciencia y otra (Keen y Spain, 1992).

Finalmente puede llegarse a una abstracción mayor, en la cual se olvida la interpretación concreta de los conceptos, símbolos y relaciones, para fijarse en su estructura. Es la etapa en la cual se pueden encontrar modelos simbólicos más o menos operativos y en la cual se puede desarrollar un lenguaje simbólico propio. La comparación de estas estructuras y lenguajes de ciencias que estudian fenómenos aparentemente distintos, permite encontrar analogías profundas y sugerir nuevos métodos, ideas y puntos de vista. Las matemáticas aparecen generalmente

en las etapas avanzadas de este proceso. Sin que se pueda definir exactamente lo que son las matemáticas, se puede observar que se ocupan del estudio y desarrollo de los modelos, lenguajes y teorías más abstractos y generales (López, 1973).

Un modelo matemático de un sistema biológico puede ser tan simple como una sola ecuación relacionando una variable con otra, o puede ser un modelo multicomponente involucrando la interacción de múltiples variables mutuamente dependientes. Los modelos matemáticos más simples pueden derivarse directamente de los modelos conceptuales o pueden obtenerse empíricamente del análisis de datos experimentales. Una ecuación o grupo de ecuaciones por sí mismo puede no contribuir mucho a la comprensión de un fenómeno particular, sino hasta que la información que puede aportar sea desplegada de alguna manera. Por esta razón, usualmente es necesario resolver la ecuación para ciertos valores representativos de la variable independiente (por ejemplo tiempo, pH, temperatura) y presentar la información resultante en forma gráfica. Debido a la cantidad de cálculos que deben realizarse, resulta sumamente práctico hacer esto mediante la implementación del modelo matemático en una computadora y por medio de la examinación de la presentación gráfica de resultados que el programa puede generar (Keen y Spain, 1992).

Según Von Bertalanffy (1986), la aproximación de datos empíricos no implica la verificación de las particulares expresiones matemáticas utilizadas. Sólo se puede hablar de verificación y de ecuaciones que representan una teoría si:

- a) los parámetros presentes son confirmables por experimentación independiente, y si
- b) de la teoría pueden derivarse predicciones de hechos aún no observados.

Debe decirse, además, que el enfoque matemático adoptado para el estudio de los sistemas biológicos no es el único posible ni el más general. Existe otra serie de enfoques modernos afines, tales como la teoría de la información, las teorías de los juegos, la decisión y las redes, los modelos estocásticos, la investigación de operaciones -por sólo mencionar los más importantes-; sin embargo, el hecho de que las ecuaciones diferenciales cubran vastas áreas de la biología y de otras ciencias naturales y sociales, las hace vía apropiada de acceso al estudio de los sistemas generalizados (Von Bertalanffy, 1986).

2.2 Simulación

2.2.1 Simulación por computadora

Dicho de una manera simple, la realización de una simulación por computadora incluye la conversión de un modelo matemático a un lenguaje de computadora e igualmente la conversión de un método numérico para su resolución, con el fin de producir datos simulados (Keen y Spain, 1992). En otras palabras, la simulación de un sistema es la puesta en operación de un modelo que represente a dicho sistema con el fin de estudiar las propiedades concernientes a su comportamiento (Naylor *et al.*, 1986). De esta manera, los datos de salida del modelo matemático pueden ser comparados con los datos experimentales para la evaluación del modelo y obtener información no asequible experimentalmente (Law y Kelton, 1991).

Según West Churchman (sin año) (cit. por Naylor *et al.*, 1986) una definición estructuralmente formal que admite las ambigüedades e inconsistencias inherentes al uso común del término simulación es la siguiente:

"x simula a y" si y sólo si: (a), x y y son sistemas formales; (b), y se considera como el sistema real; (c), x se toma como una aproximación del sistema real; (d), las reglas de validez en x no están exentas de error.

La simulación de un sistema biológico puede dar solamente una aproximación al comportamiento del sistema real y la medida en que los datos simulados se acerquen a la realidad depende de la complejidad del diseño del modelo que esté siendo procesado. Sin embargo, es posible simular sistemas que serían imposibles de estudiar experimentalmente, debido a la cantidad de tiempo y espacio implicados. Por ejemplo, la investigación de un sistema real presa—

depredador podría involucrar estimaciones de población hechas para un período de 10 a 50 años, y un área de 10 a 100 km². Incluso experimentos de población en sistemas pequeños y cerrados pueden requerir de estudios intensos durante meses para redituar resultados significantes. La simulación de este tipo de sistemas podría ser ejecutada en unos cuantos segundos con el uso de una computadora (Keen y Spain, 1992).

Los modelos matemáticos utilizados para simulación han sido divididos por Naylor *et al.* (1986) en cuatro tipos (más los modelos de tipo discreto y continuo considerados por Law y Kelton (1991)):

1. Modelos determinísticos: en estos, ni las variables exógenas ni las endógenas se les permite un comportamiento al azar, en tanto que se suponen relaciones exactas para las características de operación.
2. Modelos estocásticos: son aquellos modelos en los que por lo menos una de las características de operación está dada por una función de probabilidad.
3. Modelos estáticos: son aquellos modelos que no toman en cuenta, explícitamente, la variable tiempo.
4. Modelos dinámicos: los modelos matemáticos que tratan de las interacciones que varían con el tiempo se denominan modelos dinámicos.
5. Modelos discretos y continuos: Un modelo discreto es aquel que representa un sistema para el cual las variables de estado cambian instantáneamente a puntos separados en el tiempo. Un modelo continuo es el que describe a un sistema cuyas variables de estado cambian continuamente con respecto al tiempo.

2.2.2 Desarrollo de sistemas multicomponentes

De primera intención puede parecer muy difícil la producción de modelos de sistemas complicados, así como su evaluación. El problema se atenúa al mirar más de cerca y percatarse de que los modelos multicomponentes están a su vez contruidos de submodelos con múltiples dependencias. Cada submodelo puede ser desarrollado independientemente y validado al comparar sus resultados simulados con su respectivo subsistema real. Los submodelos pueden ser mezclados en uno solo con el poder para simular al sistema total. Esta aproximación reduccionista hace manejables las simulaciones muy complejas (Keen y Spain, 1992).

La esencia de este acercamiento es la idea de que los sistemas tienen de hecho jerarquías de organización (Walters, 1971; cit. por Keen y Spain, 1992). Cada componente del sistema incorpora sub-componentes que a su vez contienen sub-sub-componentes, y así sucesivamente (Figura 1). Afortunadamente, no es necesario tener una comprensión conceptual y mecánica de todos los subcomponentes con el fin de producir simulaciones funcionales. Algunos de los componentes y subcomponentes pueden ser incluidos en uno o más submodelos que definan el efecto neto de estos subcomponentes sobre el sistema como un todo. Este proceso de conjunción de componentes reduce la resolución del modelo, pero puede tener el efecto de enfocar la atención del modelador en factores primarios de control. La simulación por computadora es particularmente valiosa para sistemas que involucran interacciones no lineales múltiples. Esto es especialmente cierto en áreas de la biología tales como la fisiología y la ecología en las que a menudo implican control o manejo de sistemas multicomponentes. Tales sistemas contienen usualmente numerosas variables interrelacionadas por ciclos de realimentación positiva y negativa. Es esencial que las decisiones de manejo se sustenten en algo más que modelos conceptuales (Ibid.).

Al modelar un nivel intermedio de jerarquía, los niveles superiores son usualmente considerados como externos al sistema de interés. El comportamiento de una variable externa es generalmente aproximado y después incorporado al modelo como una entrada (*input*) del ambiente. Si posteriormente se vuelve evidente que la entrada aproximada es importante para el comportamiento del modelo componente, el modelo puede ser expandido para incluir este otro nivel (Ibid.).

Es un error pensar que los datos obtenidos en la simulación deben duplicar los datos obtenidos mediante la experimentación con el sistema real. Por definición un modelo simulado no es el sistema real, y no es razonable esperar datos simulados que coincidan con el sistema real punto por punto (Meadows *et al.*, 1972; cit. por Keen y Spain, 1992). Generalmente es suficiente que el modelo sea capaz de demostrar la correspondencia general con la conducta del sistema real. La medida de la desviación permisible de los datos generados por la simulación y por tanto la confianza en el modelo, depende del sistema modelado y de la función del modelo en el particular proceso de investigación en el que se desarrolle. Es posible que la simulación llegue a ser tan compleja que contribuya muy poco o nada al entendimiento del sistema modelado. Los principales factores que controlan al sistema pueden permanecer ocultos si el modelo es demasiado complejo (Ibid).

Los resultados de simulaciones obtenidos de modelos multicomponentes de sistemas complejos son a menudo controvertidos. Estos modelos son más dependientes de la capacidad de la computadora para manejar grandes cantidades de datos en forma precisa. Ellos se prestan además para dar lugar a datos inesperados. Es sólo a través de la experiencia en el trabajo con modelos simples que es posible entender las capacidades y limitaciones de las simulaciones por computadora. Con la experiencia se gana confianza en los resultados de simulaciones complejas aún cuando muestren un comportamiento contrario a lo que pudiera dictar la intuición (Naylor *et al.*, 1986; Keen y Spain, 1992).

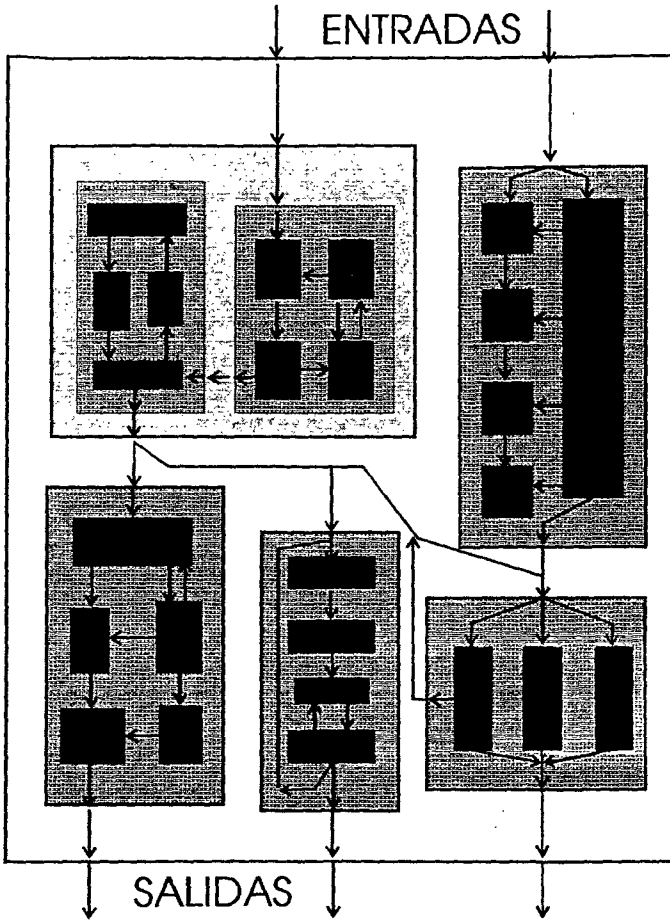


FIGURA 1. Representación de sistemas multicomponentes. Diagrama que muestra la jerarquía organizacional en cualquier sistema viviente (tomado de Keen y Spain, 1992).

2.3 Relación de los modelos y la simulación con el proceso de investigación

Los conceptos presentados anteriormente pueden ser representados en un diagrama que relacione la modelización y simulación con el proceso global de la investigación y el manejo (Figura 2). El foco de interés del investigador es el sistema real. Todo sistema biológico permanece siempre como una "caja negra". No importa cuanta información se tenga acerca de una entidad biológica particular, el investigador permanece siempre en su exterior, observándolo. Por esta razón, el sistema real se distingue de los otros componentes del diagrama de la figura 2 como un círculo. Todos los cuadros representan medios de información que son derivados, de una forma u otra, del sistema real; y como tales, son reflejos o percepciones del sistema real. Las flechas del diagrama son procesos mediante los cuales la información es obtenida y manipulada (Keen y Spain, 1992).

El primer objetivo de la ciencia es la producción de un modelo conceptual que reproduzca lo más cercanamente posible al sistema real (Bunge, 1959). El modelo conceptual es una imagen mental del sistema real basada en la información disponible interpretada a la luz de la experiencia del investigador con sistemas similares. Aun la observación más casual generalmente da lugar a un rudimentario modelo conceptual. Inicialmente éste toma la forma de analogía con sistemas conocidos. Un buen modelo conceptual puede sugerir hipótesis que pueden ser probadas experimentalmente para corroborar el modelo. Por medio de los experimentos se obtienen nuevos datos que arrojan nueva luz a las observaciones originales o a sus interpretaciones. Este proceso da lugar al mejoramiento del modelo conceptual. Esta secuencia describe el ciclo clásico de la investigación (López, 1985).

Un modelo matemático proviene de la formalización del modelo conceptual en términos cuantitativos, generalmente como una ecuación que describe la respuesta de un sistema a alguna variable o variables. El modelo matemático puede ser derivado también del análisis estadístico de datos experimentales, o por la combinación de estas dos vías. De cualquier forma, el esfuerzo analítico requerido para la formulación del modelo matemático generalmente mejora al modelo conceptual. Este modelo mejorado sugerirá nuevos experimentos. La incursión

repetida en el ciclo de la investigación cuantitativa de la figura 2 producirá mejoramientos sucesivos tanto al modelo matemático como al conceptual (Naylor *et al.*, 1986).

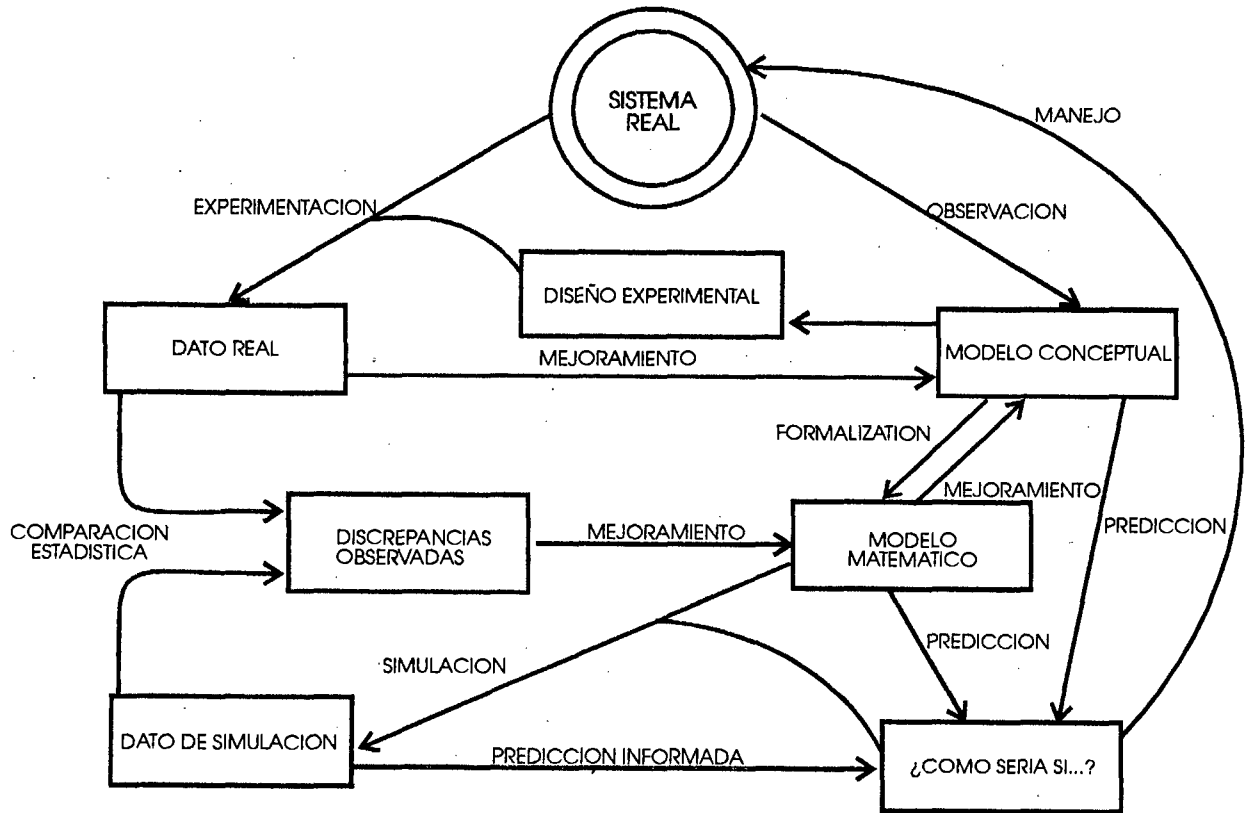


Figura 2. Papel de la simulación en el proceso de la investigación científica (tomado de Keen y Spain, 1992).

3. Justificación

Las teorías científicas tienen como último objetivo el entendimiento y predicción de los fenómenos naturales. En física y en química, la teoría y la experimentación han trabajado conjuntamente para lograr interpretar los resultados de los procesos de investigación. La biología ha abrazado con más lentitud y recelo los enfoques teóricos, en parte debido al escepticismo en cuanto a la validez de los supuestos y los métodos que son aplicables a sistemas grandes y complejos, tales como las macromoléculas de interés biológico, sin hablar de organelos, células o niveles de organización superiores. Las teorías más exitosas hacen uso de una combinación de enfoques analíticos y computacionales (Karplus y Petsko, 1990).

La formalización es la primera actividad en el proceso de solución de un problema científico y consiste en la realización de su descripción precisa a través de una cantidad finita y suficiente de información. El objetivo que se persigue con la formalización es crear un modelo adecuado del fenómeno (problema) de interés. Al aplicar la inducción y la deducción al modelo es posible inferir una ley, como consecuencia, o propiedad de una realidad formalizada. Finalmente, por medio de la igualdad se confronta la validez y utilidad de la ley obtenida (Díaz Barriga, 1992).

Mediante la sustitución de la noción de fenómeno por problema intuitivo, la de modelo por problema representado y la de ley por solución del problema y al considerar al sistema hombre-máquina como el encargado de obtener la respuesta, se puede enunciar la siguiente metodología de solución, representada gráficamente en la figura 3.

Por problema intuitivo se hace referencia al estado en el cual se cuenta con información indefinida y posiblemente ilimitada y que existe como una identidad independiente. A partir de él, el ser humano realiza una extracción de información finita, relevante y precisa, a través de un proceso de formalización, para obtener una representación del problema en términos de los datos presentes y el resultado que se espera obtener. Una vez definida la representación de la información se procede a desarrollar el algoritmo de solución, en el cual se especifican los procesos de transformación de los datos proporcionados y que posibilitan la

obtención de la respuesta buscada. Una vez terminado el algoritmo se procede a escribirlo en algún lenguaje de computación, obteniéndose así el denominado programa fuente. La computadora traduce dicho programa a instrucciones susceptibles de ser realizadas por ella misma (programa objeto), y finalmente, la computadora, al ejecutar el programa objeto, obtiene la respuesta (Ibid.).

La presentación en este documento de todos los aspectos teóricos que envuelven la solución de algunos tipos de problemas biológicos por ordenador, tiene la pretensión de mostrar al lector la trascendencia de la modelización conceptual y matemática y la simulación por medios computarizados en la biología moderna e introducirlo a las técnicas que la hacen posible.

Este trabajo se desarrolló, además, al tomar en cuenta la necesidad de despertar la atención del investigador o el estudiante hacia la elaboración de programas con aplicaciones particulares, en consideración al papel fundamental que esta tarea puede jugar en el avance local de la investigación y la docencia.

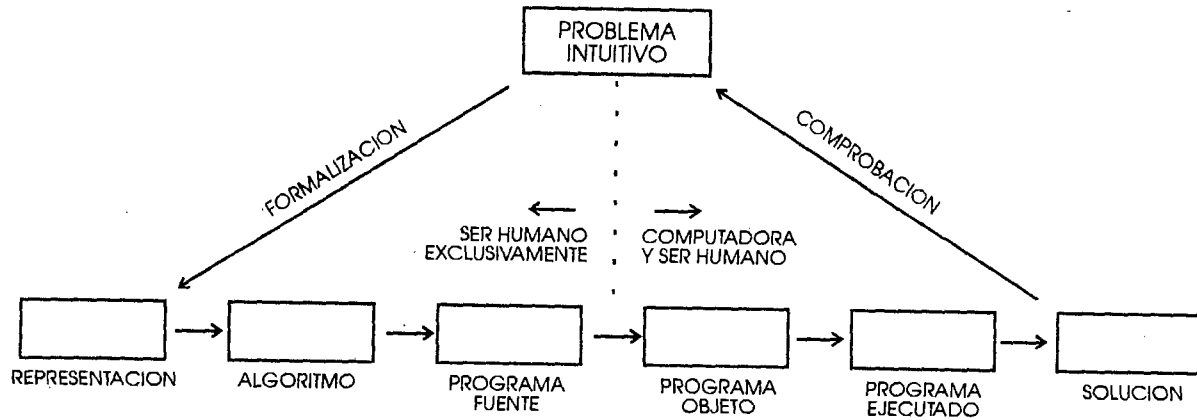


Figura 3. Solución de problemas por computadora (tomado de Díaz Barriga, 1992).

4. Objetivo

Simular procesos biológicos de interés científico mediante técnicas de programación estructurada por computadora, para aportar, de esta manera, *software* especializado en biología con la finalidad de apoyar a los procesos de investigación y docencia en esta área del conocimiento.

5. Metodología

Como primer paso para la realización del objetivo planteado se seleccionaron los sistemas biológicos cuyos modelos matemáticos se procesaron, así como los métodos numéricos requeridos para su solución. Los modelos matemáticos implementados en el programa, más que cubrir de manera amplia disciplinas específicas, pretendieron ser representativos de los diferentes niveles de organización de los sistemas biológicos. Abordaron, además, aspectos globales de estos sistemas, sin definir tendencias particulares de los fenómenos con el fin de ampliar su ámbito de aplicación.

Una vez definidos los elementos funcionales se procedió a decidir el esquema estructural bajo el cual se diseñó el programa de computadora. Este particular desarrollo de software exige la intervención de ciertas funciones de programación que tuvieron que ser estudiadas a profundidad para asegurar su aplicación adecuada. Como punto intermedio entre las entidades matemáticas utilizadas y su implementación en el lenguaje de programación fue necesario el desarrollo de algoritmos pseudocodificados con el fin de evitar ambigüedades en esta transición.

El resto de las tareas se concentraron en actividades de programación, las cuales comenzaron por la conversión de los algoritmos a código de lenguaje C++ de Borland bajo un esquema de modularización funcional. Un aspecto, particularmente complicado, que cubrió el programa fué la implementación de rutinas que permitan al usuario alimentar ciertos tipos de ecuaciones diferenciales que representen modelos no incluidos en el programa. Con el fin de dar fluidez al uso del programa, se implementaron ambientes de trabajo y cualidades funcionales como: secuencias de ventanas, activación de funciones por *mouse*, opciones de graficación y conservación de información en archivos. Finalmente, se realizaron pruebas de ejecución y optimizaciones al programa resultante.

Para la realización de las actividades de programación se empleó una microcomputadora (PC) compatible con IBM con procesador 80286. La salida gráfica puede adaptarse a cualquier tipo de controlador y modo de video. Sin embargo, el programa fue configurado para exhibirse gráficamente de la mejor manera en un monitor SuperVGA a color. La programación se llevó a cabo con

lenguaje Turbo C++ de Borland versión 1.01, en el compilador de esta misma marca. Las sentencias y comandos de programación se extrajeron de Wagner-Dobler (1987), Kernighan y Ritchie (1989), Ezzel (1990), Barkakati (1990), Schildt (1990) (2 publicaciones), y McCord (1991).

Este documento está diseñado para exponer las ideas subyacentes a todos los procedimientos utilizados, en el orden en el que éstos se expresan y con las citas a las referencias pertinentes. Debe considerarse que entre la teoría consultada y la construcción del software propuesto no existe una vía directa y única que los vincule como "receta y platillo", y, por el contrario, la teoría provee de los elementos que tienen que ser considerados para la toma de decisiones en la práctica de la realización de programas de cómputo.

5.1 Modelos analíticos basados en ecuaciones diferenciales

La presunción crítica detrás de toda simulación biológica es que una ecuación puede servir como una analogía o modelo de un proceso biológico simple. La presunción parece razonable porque casi cualquier proceso biológico puede ser descrito por una curva de causa-efecto o estímulo respuesta. La curva relaciona la intensidad o cantidad de un agente estimulante o causal y la intensidad o cantidad de una respuesta o efecto biológico. Tradicionalmente, la medida del agente, llamado variable independiente x , es trazada a lo largo del eje horizontal de una gráfica y la medida del efecto biológico o respuesta, llamado variable dependiente y , es expresada a lo largo del eje vertical. Es posible arreglar los ejes de otra manera, pero por convención es la presentación descrita la que se utilizará en este trabajo (Figura 4) (Keen y Spain, 1992).

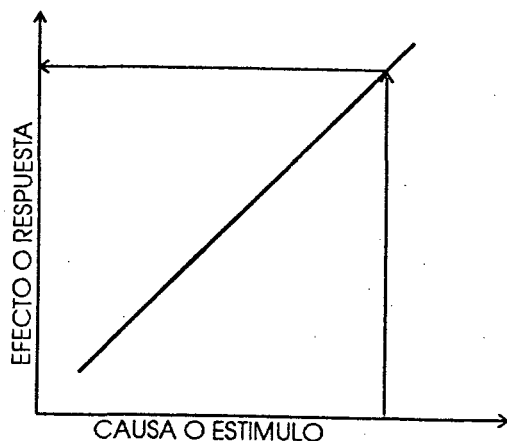


Figura 4. Graficación de variables.
Representación de una relación simple
causa-efecto (tomado de Keen y Spain, 1992).

La respuesta de un sistema biológico puede ser estudiada experimentalmente bajo diferentes condiciones, y los datos resultantes pueden ser usados para construir una curva de causa-efecto. La curva mostrará el comportamiento cuantitativo del sistema bajo cierto contexto circunstancial. Es por lo tanto un modelo diagramático del sistema. Las curvas son también utilizadas para mostrar la conducta de ecuaciones que relacionan una variable con otra. Así es posible, de alguna manera, encontrar una ecuación que genere datos de salida que constituyan una aproximación matemática de los datos producidos por el sistema biológico. La ecuación, por lo tanto, se convierte en un modelo del proceso biológico. La ecuación puede ser un sustituto del sistema biológico en términos de sus relaciones con otros componentes del sistema involucrado en el proceso (Ibid.)

En general, una ecuación es una relación que está definida entre los elementos de un conjunto y que se puede considerar que define un subconjunto cuyos elementos son los mismos que pertenecen al conjunto solución de la relación. En el lenguaje de la lógica, una ecuación es una proposición abierta que se convierte en una proposición verdadera cuando la variable se sustituye por un elemento del conjunto solución (The Open University, 1974).

Una *ecuación diferencial* es una ecuación que contiene una variable desconocida y sus derivadas. Una ecuación diferencial es una *ecuación diferencial ordinaria* si la variable desconocida depende solamente de una variable independiente. Si la variable desconocida depende de dos o más variables independientes, la ecuación diferencial es una *ecuación diferencial parcial*. El *orden* de una ecuación diferencial es el orden de la mayor derivada que aparece en la ecuación. El *grado* de una ecuación diferencial, que puede escribirse como un polinomio en la variable desconocida y sus derivadas es la potencia a la cual está elevada su derivada de mayor orden (Bronson, 1984).

Los modelos analíticos son frecuentemente expresados como ecuaciones diferenciales que definen una velocidad de cambio de una variable dependiente con respecto a una variable dependiente. En los modelos biológicos, la variable independiente es usualmente el tiempo, la distancia o la concentración (Keen y Spain, 1992).

El resto de este subcapítulo (5.1) estará dedicado a la exposición de los modelos incluidos en el programa de cómputo.

5.1.1 Modelo de crecimiento microbial de Monod

El crecimiento microbiano se puede considerar como un conjunto de reacciones químicas en cadena, que conducen a la síntesis de los constituyentes de la biomasa microbiana obtenida al final de la operación. Globalmente, el principio obedece al principio de la conservación de la materia (Leveau y Bouix, 1985).

5.1.1.1 Modelos biomatemáticos y fenomenológicos para el estudio de las poblaciones de microorganismos

Para la formulación de estos modelos de crecimiento microbiano, se pueden considerar los diferentes tipos de leyes que generalmente se aplican en mayor grado como son las de conservación de la materia y de la energía. No obstante, el crecimiento celular es un conjunto de fenómenos complejos en los cuales intervienen numerosas reacciones enzimáticas sometidas a mecanismos de regulación muy precisos. Aunque las vías metabólicas inherentes a los fenómenos de crecimiento así como sus mecanismos de regulación son conocidos, el modelo matemático se establece sobre la base de los principios que los rigen: ley de Michaelis-Menten para las reacciones enzimáticas sencillas y las leyes de King y Altman para las reacciones con enzimas alostéricas. Los modelos biomatemáticos que así se obtienen, aunque son complejos, convienen para el estudio de los fenómenos microbianos en los cuales entran en juego un número relativamente limitado de reacciones enzimáticas (Leveau y Bouix, 1985).

Cuando se desconocen los mecanismos bioquímicos o son muy numerosos, se realiza una descripción de la evolución del sistema con ayuda de una función matemática empírica. El tipo de modelo generado es llamado descriptivo o fenomenológico (Leveau y Bouix, 1985).

5.1.1.2 Modelo de Monod

Desde hace bastante tiempo los microbiólogos se han percatado de que los microorganismos responden a sus fuentes de nutrientes de manera muy parecida a la manera en que las enzimas actúan sobre sus sustratos. La similitud más sobresaliente es el modo en que la velocidad de crecimiento de los microorganismos responde al incremento en la concentración de un nutriente limitante. Jacques Monod notó que esta respuesta es hiperbólica, como la conducta de saturación de las enzimas (Figura 5) (Keen y Spain, 1992).

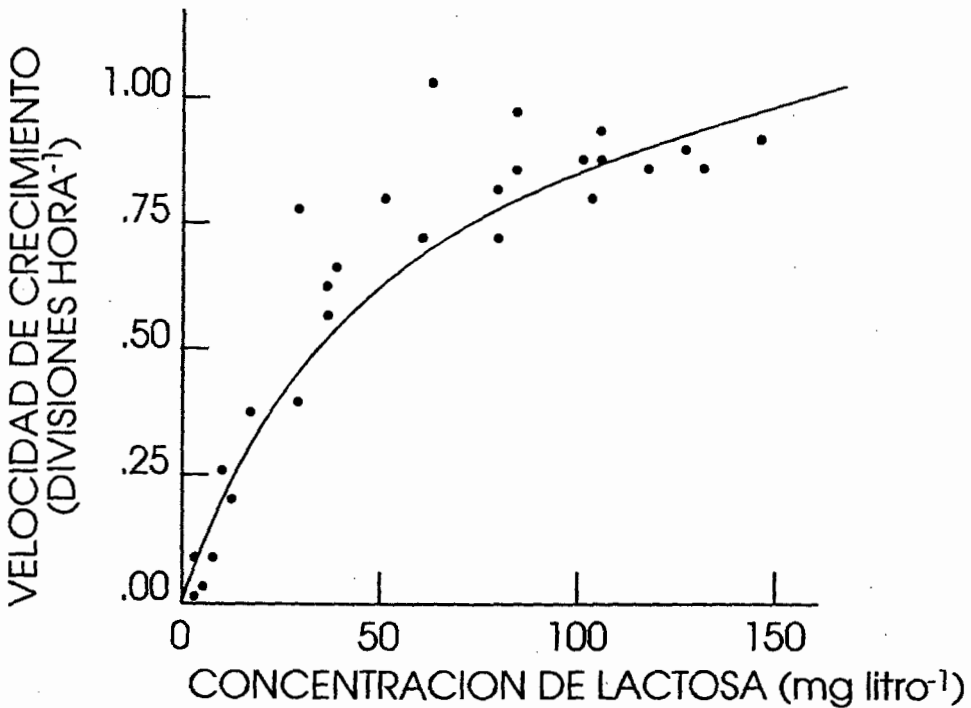


Figura 5. Comportamiento de un cultivo de microorganismos *Escherichia coli* (tomado de Leveau y Bouix, 1985).

Este comportamiento sugiere que las ecuaciones para las enzimas podrían servir como modelos para algunos tipos de crecimiento microbial, como se expresa en la siguiente ecuación:

$$\frac{dB}{dt} = \mu B = \mu_m \frac{[S]}{K_s + [S]} B,$$

1

donde B es la densidad celular, μ es la constante de velocidad de crecimiento, S es la concentración de nutrientes, y μ_m es la máxima velocidad de crecimiento cuando la concentración de nutrientes es completamente ilimitada. K_s es la constante de saturación media, que representa la concentración de sustrato que permite el crecimiento a la mitad de la velocidad máxima; así, $[S] = K_s$ cuando $\mu = \mu_m / 2$. En efecto, K_s determina qué tan rápidamente una curva hiperbólica como la de la figura 5 se acerca a la asíntota. B puede ser medida como biomasa (p.ejm. $\text{mg} \cdot \text{l}^{-1}$) o como una densidad de células (p.ejm. $\text{cantidad} \cdot \text{ml}^{-1}$) tomando el tamaño celular como uniforme. μ tiene unidades de biomasa: $(\text{unidad de biomasa})^{-1} \cdot \text{t}^{-1}$ (Ibid.).

Una derivación formal de esta ecuación posiblemente no sea posible, pero puede ser obtenida por analogía a la ecuación de Michaelis-Menten. Una interacción enzima-sustrato puede ser expresada gráficamente como

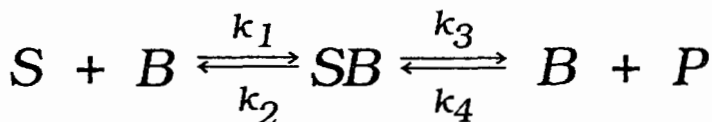


donde E es la enzima libre, S es el sustrato, ES es el complejo enzima-sustrato, P es el producto, y k_1 , k_2 , k_3 , y k_4 son las constantes de velocidad para las reacciones. Estas reacciones pueden ser modeladas con la siguiente ecuación:

$$v = V_{max} \frac{[S]}{K_m + [S]}$$

Esta ecuación se derivó suponiendo una concentración estacionaria para ES y suponiendo que v se acerca a V_{max} según ES se acerca a E_{total} .

Similarmente, el crecimiento en una población de células en cultivo puede ser descrito por una "reacción" como la siguiente:



En esta descripción hipotética y no muy precisa, B se refiere a las células vivas, S al nutriente limitante, SB a la combinación de células más nutriente absorbido pero no asimilado, y P es la nueva biomasa celular derivada de los nutrientes metabolizados. k_1 es la velocidad de toma de nutriente, k_2 es la velocidad de pérdida de nutriente a través de la excreción, y k_3 es la velocidad de formación de nuevo material celular. En este esquema k_4 es cero porque P es la nueva biomasa celular y no es distinguible de B (k_4 podría tomar un valor positivo en el caso de

la presencia de canibalismo). Debido a los requerimientos metabólicos, debe ser usado más S por las células de lo que aparece en la cantidad de nueva biomasa, P . El proceso puede exhibir las características de retroalimentación de una reacción enzimática autocatalítica (Ibid.).

Como sucede con las enzimas, el nuevo material celular se forma más rápidamente cuando la biomasa celular y sus sistemas asociados están saturados con el nutriente en la forma de SB . Se ha acumulado considerable evidencia de la validez del modelo de Monod desde que fué propuesto. En la práctica, los valores de μ son encontrados experimentalmente para diferentes concentraciones de nutriente S (Ibid.).

En muchos experimentos, la ocurrencia de respiración, mortalidad y autólisis ha sido ubicada como causa de decremento en la velocidad de crecimiento independientemente de la concentración del nutriente (Herbert, 1958; cit. por Keen y Spain, 1992). Para estos casos, la Ecuación 1 es modificada para incluir un término adicional que considere estas pérdidas:

$$\frac{dB}{dt} = \mu B = \left(\mu_m \frac{[S]}{K_s + [S]} - R \right) B,$$

3

donde R es una constante de velocidad que describe la pérdida de biomasa (Ibid.).

En 1936 Teissier (cit. por Leveau y Bouix, 1985) desarrolló una ley de crecimiento basada sobre el principio de la conservación de la materia para establecer el balance:

$$\frac{dB}{dt} = -C \frac{d[S]}{dt},$$

4

en la cual $d[S] / dt$ es la rapidez de desaparición del sustrato y C el coeficiente de conversión de sustrato a biomasa (g de biomasa / g de sustrato).

Mediante la utilización de esta ley, se deriva la ecuación para el cálculo de la cantidad sustrato presente en el tiempo:

$$\frac{d[S]}{dt} = - \frac{1}{C} \left(\mu_m \frac{[S]}{K_s + [S]} - R \right) B .$$

5

Mediante la solución del sistema formado por las ecuaciones 3 y 5 es posible analizar el comportamiento de una población microbiana en relación al consumo de nutrientes y con respecto al tiempo.

5.1.2 Difusión pasiva a través de una membrana

Una ecuación similar a la siguiente:

$$\frac{dC}{dt} = -k C,$$

6

que describe el decaimiento exponencial, puede ser utilizada para modelar el proceso de difusión pasiva. La velocidad de cambio de concentración del soluto interno de una célula, causado por la difusión pasiva hacia el ambiente con una concentración constante de soluto, está dada por:

$$\frac{dC}{dt} = -k (C - C_x),$$

7

donde C es la concentración interna para una célula de volumen y superficie unitarios, C_x es la concentración ambiental, y k es la constante de proporcionalidad para la velocidad de difusión (Keen y Spain, 1992).

5.1.3 Ley del enfriamiento de cuerpos Newton

La forma básica para el decaimiento exponencial dada por la Ecuación 6 ha sido modificada para dar las bases de numerosos modelos biológicos, entre los que se encuentra el anterior modelo descrito. Otro de estos modelos conforma a la ley del enfriamiento de Newton que describe la pérdida de calor de un cuerpo en enfriamiento. Esta ley establece que la temperatura de un objeto baja a una velocidad proporcional a la diferencia entre la temperatura del objeto y la temperatura del medio. La velocidad del cambio de temperatura con respecto al tiempo está dada por:

$$\frac{dT}{dt} = -k (T - T_m),$$

8

donde T_m es la temperatura del medio y k es la constante de velocidad de enfriamiento (Keen y Spain, 1992).

5.1.4 Modelo de crecimiento de peces de Von Bertalanffy

5.1.4.1 Teoría del crecimiento animal

Las ideas fundamentales sobre el crecimiento animal se remontan al fisiólogo alemán Pütter (1920) (cit. por Von Bertalanffy, 1986), y fueron desarrolladas posteriormente por Ludwig Von Bertalanffy. Si un organismo es un sistema abierto, su incremento o tasa de crecimiento (*T.C.*) puede expresarse, muy generalmente, por una ecuación de balance de la forma:

$$\frac{dw}{dt} = T.C. = \textit{síntesis} - \textit{degeneración} + \dots,$$

9

es decir, que el incremento en peso es representado por la diferencia entre procesos de síntesis y degeneración de sus materiales constituyentes, más cualquier número de factores indeterminados que influyan sobre el proceso. Sin pérdida de generalidad, puede suponerse también que los términos son algunas funciones indefinidas de las variables en cuestión:

$$T.C. = f_1(w, t) - f_2(w, t) + \dots$$

10

Se observa ahora, que el tiempo *t*, no debe entrar en la ecuación. Pues algunos procesos de crecimiento son equifinales, es decir, que alcanzan los mismos valores finales en diferentes tiempos. Aún sin prueba matemática estricta, se ve intuitivamente que esto no sería posible si la tasa de crecimiento dependiera directamente del tiempo, pues, de ser este el caso, no podrían darse tasas diferentes en tiempos dados, como sucede en la realidad (Von Bertalanffy, 1986).

En consecuencia, los términos considerados serán funciones de la masa corporal presente:

$$T.C. = f_1(w) - f_2(w) , \quad 11$$

si provisionalmente se limita la consideración al más sencillo esquema de sistema abierto. El supuesto más simple posible es que los términos sean funciones tipo potencia de la masa corporal. Y de hecho se sabe empíricamente que, con gran generalidad, la dependencia de procesos fisiológicos con respecto al tamaño es susceptible de buena aproximación por medio de expresiones alométricas (ecuaciones de crecimiento que relacionan los incrementos de dos variables) . Se tiene entonces:

$$\frac{dw}{dt} = \sigma w^n - \kappa w^m , \quad 12$$

donde σ y κ son constantes de anabolismo y catabolismo, respectivamente, correspondiendo a la estructura general de las ecuaciones alométricas (Ibid.).

La síntesis de materiales de construcción requiere energía que, en los animales aerobios, es suministrada por procesos de respiración celular y, a fin de cuentas, el sistema del ATP. Puede suponerse que hay correlaciones entre el metabolismo energético de un animal y sus procesos anabólicos. Esto es plausible en la medida en que el metabolismo energético debe, de uno u otro modo, suministrar las energías requeridas para la síntesis de componentes del cuerpo. Se inserta, pues, como dependencia del anabolismo con respecto al tamaño, la de las velocidades metabólicas ($n = \alpha$) y se llega a la sencilla ecuación (Von Bertalanffy, 1986):

$$\frac{dw}{dt} = \sigma w^\alpha - \kappa w .$$

13

5.1.4.2 Modelo simple del crecimiento de peces

Este modelo clásico describe la longitud de un pez como una función de la edad, basándose en que el crecimiento de éste es proporcional a la diferencia entre su longitud y una longitud máxima teórica. Esto es, que el pez crece más rápido cuando es más pequeño, con una velocidad de crecimiento que decrece según el tamaño se acerca a su máximo valor. La ecuación diferencial que describe este proceso es:

$$\frac{dL}{dt} = k (L_m - L),$$

14

donde L es la longitud del pez, L_m es la longitud máxima teórica, y k es la constante de velocidad de crecimiento (Keen y Spain, 1992).

5.1.5 Modelo compartamental simple

Un proceso físico o biológico complejo puede ser dividido a veces en varios estadios distintos. El proceso total puede describirse por la interacción entre los estados individuales. Cada estado se llama *compartimiento* o estanque y se supone que el contenido de cada compartimiento está bien mezclado. Se transfiere material de un compartimiento a otro que se incorpora inmediatamente al segundo. El proceso total se denomina *sistemas de compartimientos*. Un sistema es *abierto* si entra o sale material del mismo a través de uno o más de los compartimientos. Un sistema que no es abierto se denomina *cerrado* (Derrick y Grossman, 1984).

A continuación se describe el sistema más sencillo: el formado por un sólo compartimiento.

La figura 6 representa un sistema de un compartimiento constituido por una cantidad $x(t)$ de material en el compartimiento, una velocidad de flujo $i(t)$ a la cual se introduce material al sistema, y un *coeficiente de transferencia fraccional* k que indica la cantidad de material del compartimiento que se extrae del sistema por unidad de tiempo.

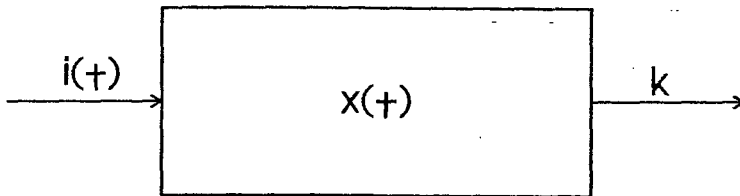


Figura 6. Sistema de un solo compartimiento (tomado de Derrick y Grossman, 1984).

Es claro que la tasa a la cual se cambia la cantidad x depende de la diferencia entre la entrada y la salida de material en cualquier tiempo t , lo cual conduce a la ecuación diferencial (Derrick y Grossman, 1984):

$$\frac{dx}{dt} = i(t) - k x(t) .$$

15

5.1.6 Modelo de poblaciones de Verhulst-Pearl

Al estudiar la dinámica de las poblaciones de organismos se hace la suposición de que éstas son homogéneas, es decir, que están compuestas por un solo tipo de organismo y suelen ser descritas por una sola variable, la densidad. Las diferencias en edad, sexo, genotipo, fenotipo, etc. son ignoradas o tomadas como irrelevantes para el modelo. La densidad es generalmente expresada como cantidad por unidad de área (p.ejm. venados * ha⁻¹) o cantidad por unidad de volumen (p.ejm. *Paramecium* * ml⁻¹). Los modelos para poblaciones homogéneas suponen que las interacciones (depredación, parasitismo, competencia, etc.) entre poblaciones procede a velocidades directamente proporcionales al producto de las densidades de poblaciones (Keen y Spain, 1992).

La mayoría de las poblaciones de la mayoría de los organismos no son homogéneas, por lo que las simplificaciones aquí efectuadas se traducen en imprecisiones al comparar los datos simulados con datos reales. Sin embargo, el modelo que a continuación se describirá permite acercarse al comportamiento de las poblaciones sin detallar los rasgos de una especie en particular, además de que este modelo y otros que de él se derivan son importantes en la actual investigación científica de poblaciones (Ibid.).

5.1.6.1 Ecuación de Verhulst-Pearl

El modelo más simple de crecimiento de poblaciones biológicas en un medio ambiente ilimitado y con una tasa constante de crecimiento está dado por:

$$\frac{dN}{dt} = k N .$$

16

Las observaciones de las poblaciones reales sugieren que esta ecuación no es un buen modelo para el crecimiento biológico. La modificación más simple para adaptar la ecuación al contexto del crecimiento limitado es suponer que k en la

Ecuación 16 no es una constante, sino que decrece según N se incrementa. La suposición más simple es que k decrece linealmente con respecto a N , de manera que:

$$k = a - b N,$$

17

Si los organismos se encuentran en tal cantidad de manera que los recursos puedan ser tomados como abundantes, la tasa de crecimiento se acercará a a (Ibid.). Pero la tasa de mortalidad es proporcional al tamaño de la población, debido a los efectos del hacinamiento y la competencia creciente por el alimento disponible; supóngase que esta última constante de proporcionalidad es $b > 0$ (Derrick y Grossman, 1984). Hay también una densidad limitante donde la tasa de crecimiento es cero, y a densidades más grandes que ésta, la tasa de crecimiento se vuelve negativa. Si esta expresión es sustituida por k en la Ecuación 16, se genera la siguiente ecuación:

$$\frac{dN}{dt} = (a - b N) N = a N - b N^2.$$

18

Esta ecuación es una forma de la logística de Verhulst-Pearl, propuesta por primera vez por Verhulst en 1838 (cit. por Von Bertalanffy, 1986) para el estudio de poblaciones humanas y redescubierta después por Pearl y Reed en 1920 (cit. por Keen y Spain, 1992). El tipo de crecimiento representado por esta ecuación se llama crecimiento logístico (Derrick y Grossman, 1984). Otra forma en que se presenta es:

$$\frac{dN}{dt} = c N (L - N) .$$

19

Esta ecuación es usualmente expresada con la siguiente terminología:

$$\frac{dN}{dt} = r N (1 - N / K) .$$

20

Todas estas expresiones son formalmente idénticas, con $a = r$, $L = K$, y $b = c = r/K$. Aquí, r es el término que describe la velocidad de crecimiento *per capita* a densidad mínima, y K describe la densidad de población a la cual la velocidad de crecimiento es cero (Keen y Spain, 1992).

La Ecuación 20 es quizás la forma más fácilmente interpretada de la logística. Cuando N es relativamente más pequeña que K , $(1 - N / K)$ se acerca a 1 y la velocidad de crecimiento es casi exponencial a velocidad r . Cuando $N = K$, la tasa de crecimiento de la población es cero. Además, una población crecerá asintóticamente al valor K . Si N excede K , la tasa de crecimiento es negativa y la población declina hacia K (Ibid.).

La forma de la logística en la Ecuación 18 muestra claramente la relación de la logística con la ley de la acción de masas. El primer término aN describe el crecimiento ilimitado, y representa el crecimiento potencial de la población. El segundo término es negativo, y es una función de $N * N$. Puede ser visto como una pérdida del potencial de población, proviniendo de los efectos negativos de la interacción de los organismos. Con el tiempo, la densidad de población tenderá a un estado estacionario en el cual el potencial de incremento está exactamente balanceado por la pérdida de potencial. El segundo término constituye un buen ejemplo de la aplicación de la ley de masas a la dinámica de poblaciones (Ibid.).

5.1.7 Modelo de depredación de Lotka-Volterra

Lotka en 1925 y Volterra en 1926 desarrollaron independientemente un modelo de interacción depredador-presa (Hutchinson, 1978). Este modelo ha servido como base para muchos modelos de depredación y constituye un punto de partida útil para cualquier investigación de depredación (Keen y Spain, 1992).

El modelo está constituido por dos ecuaciones. La primera describe los cambios en la densidad de población de la presa:

$$\frac{dN}{dt} = rN - gNP .$$

21

En esta ecuación, N es la densidad de presa y r es la velocidad de crecimiento de la población de presa como en la Ecuación 20 del anterior modelo tratado. P es la densidad de depredador y g es una velocidad constante que representa la eficiencia de depredación. Nótese que la población de presa puede crecer sin límites, excepto por la pérdida ocurrida por depredación. Si los depredadores están ausentes ($P = 0$), entonces la ecuación anterior se convierte en la ecuación de crecimiento exponencial (Ecuación 16) (Ibid.).

La ecuación para el cambio en la densidad de poblaciones de depredador es:

$$\frac{dP}{dt} = hNP - mP .$$

22

En esta ecuación el crecimiento de la población de depredador depende solamente de la existencia de presa (hNP), de manera que cuando N sea cero, la población de depredador declinará exponencialmente con una velocidad m , la constante de velocidad de muerte del depredador. La constante h incluye la velocidad de captura (g en la Ecuación 21) multiplicada por un factor por la eficiencia de la conversión de presa capturada en depredador. Ambas ecuaciones incluyen el producto NP que es típico de los modelos basados en la ley de la acción de masas (Ibid.).

Como una ayuda en la visualización de la conducta de los dos modelos para dos especies interactuantes, es útil mostrar las densidades de población graficadas una contra la otra, formando un trazado de fase. La dimensión temporal no aparece directamente en estos trazados. Es también valioso para estos trazados de fase incluir líneas, llamadas isoclinas, que indican los valores de equilibrio para cada especie. Una isoclina efectivamente divide el trazado de fase en dos regiones, una en la que la población se incrementará, y otra que la que decrecerá. Con los modelos depredador-presa, la densidad de presa es convencionalmente dada en el eje de las x , y la densidad de depredador en el eje de las y . La isoclina de presa es una línea que comprende todos los puntos para los cuales $dN/dt = 0$. La ecuación para la línea se encuentra al hacer la Ecuación 21 igual a cero y al resolver para encontrar:

$$P = \frac{r}{g}.$$

23

De esta manera se describe una línea que corre paralela al eje de las x . Así, con la combinación de las densidades de depredador y presa por encima de la línea, las poblaciones de presa decrecerán; por debajo de ella, se incrementarán. Una isoclina para cantidades de depredador puede ser similarmente construída al hacer la Ecuación 22 igual a cero y al resolver:

$$N = \frac{m}{h}$$

24

La isoclina de depredador es una línea vertical, a cuya derecha la población de depredador se incrementará, y a la izquierda, decrecerá. Un trazado de fase para las interacciones presa- depredador basándose en el modelo de Lotka-Volterra se muestra en la figura 7 (Ibid.).

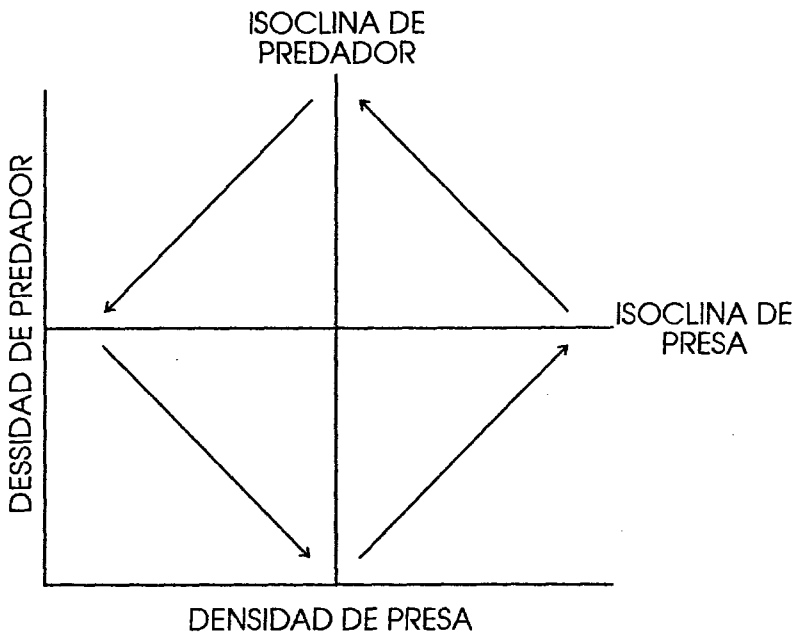


Figura 7. Líneas isoclinas de las variables presa y depredador. Un trazado de fase basado en el modelo simple de Lotka-Volterra (tomado de Keen y Spain, 1992).

5.1.8 Modelo de epidemias de Bailey

5.1.8.1 Modelos de epidemias

La simulación de la propagación de una enfermedad entre los individuos de una población es una de las aplicaciones realmente prácticas de las simulaciones por computadora basadas en modelos biológicos (Keen y Spain, 1992). Los modelos de epidemias han sido de un valor evidente como parte de los estudios médicos humanos y veterinarios durante varias décadas (Derrick y Grossman, 1984).

El vocabulario de la epidemiología involucra algunas definiciones estrictas para su uso en el planteamiento de los modelos. Cuando un patógeno es transmitido a un nuevo huésped, puede haber un período latente entre el momento de la infección y el comienzo del período en el que el huésped se vuelve infeccioso. El período infeccioso dura por todo el tiempo en que el huésped es capaz de transmitir la enfermedad a otros individuos. El período de incubación es el tiempo transcurrido entre la transmisión y la aparición de síntomas de la enfermedad. El período de incubación puede ser independiente de los períodos latente e infeccioso, dependiendo de la patología del padecimiento. Para algunas enfermedades existe un período inmune después de que ha transcurrido la infección. Un individuo inmune no puede ser infectado, pero puede ser capaz de transmitir la enfermedad (Keen y Spain, 1992).

La biología de la transmisión de enfermedades da lugar a varias clases posibles de organismos huéspedes, que son importantes en la configuración de los modelos de epidemias. Un individuo susceptible no está infectado por un patógeno; la enfermedad puede ser transmitida a los susceptibles. Un portador se encuentra usualmente en un estadio de la infección y puede transmitir la enfermedad a individuos susceptibles. Durante una epidemia, los individuos pueden ser removidos del grupo de infectados para evitar que transmitan la enfermedad a los susceptibles. La recuperación y subsecuente ganancia de inmunidad también removerá individuos de la población epidémica, así como lo harán las muertes que se presenten. La vacunación contra el patógeno transferirá a un individuo de la clase de los susceptibles a la clase de los removidos sin pasar por la clase de los portadores infectados (Derrick y Grossman, 1984; Keen y Spain, 1992).

5.1.8.2 Modelo general de epidemias

La mayor parte de los conceptos utilizados en la formulación de los modelos de epidemias fueron desarrollados inicialmente por Kermack y McKendrick (sin año) (cit. por Keen y Spain, 1992). En 1957, N.T.J. Bailey (cit. por Derrick y Grossman, 1984) publicó su libro *The Mathematical Theory of Epidemics* en el cual describe varios modelos que han sido usados para estudiar la transmisión de enfermedades.

Las características biológicas de la enfermedad epidémica se han simplificado para su aplicación al modelo que a continuación se describe. Los huéspedes pueden existir en sólo tres clases: susceptibles, portadores infectados, y removidos. Para simplificar, el período latente es tomado como cero, de manera que todos los individuos infectados son también portadores que infectan susceptibles. Los individuos removidos han tenido la enfermedad, pero ya no participan en la propagación de la enfermedad debido a la inmunidad adquirida, al aislamiento, o a la muerte (Keen y Spain, 1992).

El modelo general toma el movimiento de los individuos a través de las clases sucesivas de susceptibles, portadores y removidos. El tamaño de la población (o densidad) es N , el número de individuos susceptibles a la enfermedad es S , la cantidad de portadores infectados es C , y la cantidad de individuos removidos es R . La población total N en cualquier momento durante la epidemia es la suma de las tres clases de huéspedes (Derrick y Grossman, 1984):

$$N = S + C + R .$$

25

Nótese que en modelo general, las muertes debidas a la enfermedad epidémica no decrementan el valor de N .

La velocidad de contacto entre portadores y susceptibles es proporcional al producto de sus cantidades. Si la constante de transmisión de la enfermedad es k , entonces la velocidad a la que los susceptibles son infectados y se convierten en portadores será:

$$\frac{dS}{dt} = -k_1 S C .$$

26

El cambio en la cantidad de portadores está determinado por la velocidad a la cual los susceptibles se convierten en portadores, y por la velocidad a la cual los portadores son removidos:

$$\frac{dC}{dt} = k_1 S C - K_2 C ,$$

27

donde k_2 es el coeficiente de remoción (Keen y Spain, 1992).

El cambio en la cantidad de individuos removidos está dado por:

$$\frac{dR}{dt} = k_2 C .$$

28

Una simulación epidémica comienza con una población N consistente en su mayor parte de individuos susceptibles S , y una pequeña cantidad C de organismos infectados. La cantidad R de individuos removidos es tomada como cero inicialmente (Ibid.).

5.1.8.3 La curva epidémica

Además del decremento en susceptibles y el correspondiente incremento de portadores, los epidemiólogos frecuentemente se interesan en la cantidad de individuos infectados que aparecen en el transcurso del tiempo. La información acerca de nuevos casos por unidad de tiempo producirá la "curva epidémica" cuando es graficada con respecto al tiempo. Para algunos casos, la curva epidémica puede también ser obtenida trazando muertes u hospitalizaciones por unidad de tiempo. Para una epidemia simple, la curva se eleva hacia un pico único y después declina simétricamente. La simulación de una curva epidémica puede basarse en el modelo general graficando $-dS/dt$ de la Ecuación 26 con respecto al tiempo. El signo negativo se debe a que S siempre decrece durante el curso de una epidemia simple (Ibid.).

5.1.8.4 Modelo de epidemia simple

Bailey (sin año) (cit. por Keen y Spain, 1992) desarrolló un modelo epidémico elemental basado en simplificaciones del modelo general antes descrito. Las características biológicas de la enfermedad epidémica son simplificadas tomando solamente dos clases de huéspedes: individuos susceptibles y portadores infectados. El período latente es cero, y la velocidad de remoción es también tomada como cero, de manera que los infectados son tomados como portadores durante toda la epidemia, sin muertes, recuperación, o inmunidad. Estas suposiciones son tomadas debido a que cubren razonablemente algunas enfermedades conocidas de plantas y animales; coinciden aproximadamente también con las características de la expansión del resfriado común en poblaciones pequeñas y cerradas de humanos.

Como antes, el tamaño de la población es N , la cantidad de individuos susceptibles S , y el número de portadores infectados es C . Antes de que la epidemia comience, todos los individuos son tomados como susceptibles. En cualquier momento durante la epidemia,

$$N = S + C .$$

Si la velocidad de transmisión de la enfermedad es k , entonces la velocidad a la cual los susceptibles son infectados y se vuelven portadores puede ser descrita como en la Ecuación 26:

$$\frac{dS}{dt} = -k S C .$$

30

Para este modelo simple, el tamaño N de la población se mantiene constante y la ecuación puede ser escrita como:

$$\frac{dS}{dt} = -k S (N - S) .$$

31

La literatura biomédica contiene muchos ejemplos de simulaciones de epidemias de varias enfermedades. La mayor parte de éstas provienen de modelos complejos de enfermedades específicas, pero todos están basados en los elementos generales que han sido expuestos aquí. Los modelos estocásticos han sido importantes en el desarrollo de la teoría de epidemias, pero su papel en la simulación parece ser menos trascendente que el de los modelos determinísticos. Los trabajos recientes en la dinámica caótica de las epidemias conducen actualmente a una nueva reorientación en la modelización epidemiológica (Ibid.).

5.2 Métodos numéricos

Una ecuación diferencial junto con las condiciones complementarias de la variable desconocida y sus derivadas, todas dadas para el mismo valor de la variable independiente, constituye un problema de valor inicial (Bronson, 1984).

Un método numérico para resolver un problema de valor inicial es un procedimiento que conduce a soluciones aproximadas en puntos especiales, al usar solamente las operaciones de adición, sustracción, multiplicación y división y cálculos de funciones (Bronson, 1984).

A continuación se hace una exposición resumida de los métodos numéricos utilizados para la resolución de los problemas de valor inicial de primer orden de la forma:

$$y' = f(x, y); \quad y(X_0) = y_0.$$

32

Todos los métodos numéricos implican encontrar soluciones aproximadas en x_0, x_1, x_2, \dots , donde la diferencia entre dos valores sucesivos cualesquiera de x es una constante h ; es decir $x_{n+1} - x_n = h$ ($n = 0, 1, 2, \dots$). El valor de h se escoge arbitrariamente; en general, entre más pequeño es h , la solución aproximada es más exacta (Bronson, 1984).

La solución aproximada en x_n se designará por $y(x_n)$, o simplemente y_n .

5.2.1 Método de Euler

Este Método fué ideado por Euler(1707-1783) hace más de 200 años. Es fácil de entender y de usar, pero no es tan preciso como otros métodos que serán expuestos más adelante. Si embargo el método de Euler sirve como punto de partida hacia técnicas alternativas (Smith, 1989).

Este método está dado por:

$$y_{n+1} = y_n + h y'_n, \quad 33$$

o por:

$$y_{n+1} = y_n + h f(x_n, y_n). \quad 34$$

5.2.2 Método de Heun

Este es una mejora del método de Euler (es también llamado método Euler mejorado (Keen y Spain, 1992) y está dado por

$$y_{n+1} = y_n + \frac{h}{2} [y'_n + f(x_n + h, y_n + h y'_n)]. \quad 35$$

5.2.3 Métodos Runge-Kutta

Los métodos Runge-Kutta son un conjunto de métodos numéricos que se inician por sí mismos. El método de Euler es un método de Runge-Kutta de primer orden; el método de Heun es un método de Runge-Kutta de orden dos (Bronson, 1984).

Los métodos Runge-Kutta pueden utilizarse para encontrar soluciones completas.

5.2.3.1 Método Runge-Kutta de tercer orden

$$y_{n+1} = y_n + 1/6 (k_1 + 4 k_2 + k_3),$$

36

donde

$$k_1 = h f(x_n, y_n),$$

$$k_2 = h f(x_n + 1/2 h, y_n + 1/2 k_1),$$

$$k_3 = h f(x_n + h, y_n - k_1 + 2 k_2).$$

5.2.3.2 Método Runge-Kutta de cuarto orden

El método Runge-Kutta de orden tres es raramente utilizado. En éste, como sería de esperar, se evalúa f en tres puntos seleccionados (por longitud de paso h). En el caso de los métodos de orden cuatro evaluamos f en cuatro puntos seleccionados y usamos $k_1, k_2, k_3,$ y k_4 (Smith, 1989). Es decir, se hace:

$$y_{n+1} = y_n + 1/6 (k_1 + 2 k_2 + 2 k_3 + k_4),$$

37

donde

$$k_1 = h f(x_n, y_n),$$

$$k_2 = h f(x_n + 1/2 h, y_n + 1/2 k_1),$$

$$k_3 = h f(x_n + 1/2 h, y_n + 1/2 k_2),$$

$$k_4 = h f(x_n + h, y_n + k_3).$$

Aunque los métodos de Runge-Kutta pueden ser desarrollados a casi cualquier orden, el de 4o. orden resulta el más conveniente debido a que ofrece un balance de ganancias y gastos en su ejecución, es decir, los resultados que estima son bastante aproximados a las soluciones reales y el tiempo empleado en el cálculo no es desmedido (Keen y Spain, 1992).

Finalmente, se presenta en la próxima página el método Runge-Kutta de orden cuatro expresado como un algoritmo pseudocodificado (extraído de Burden y Faires, 1985) que constituye un paso intermedio hacia su implementación en el programa de computadora.

Algoritmo de Runge-Kutta de 4o. orden
para aproximar la solución del problema de valor inicial

$$y' = f(t, y), \quad a \leq t \leq b, \quad y(a) = \alpha,$$

en $(N + 1)$ números igualmente espaciados en el intervalo $[a, b]$:

ENTRADA puntos extremos a, b ; entero N ; condición inicial α .

SALIDA aproximación w de y en los $(N + 1)$ valores de t .

Paso 1 Tomar $h = (b - a) / N$;

$$t = a;$$

$$w = \alpha;$$

SALIDA (t, w) .

Paso 2 Para $i = 1, 2, \dots, N$ seguir los Pasos 3-5

Paso 3 Tomar $K_1 = h f(t, w)$;

$$K_2 = h f(t + h/2, w + K_1/2);$$

$$K_3 = h f(t + h/2, w + K_2/2);$$

$$K_4 = h f(t + h, w + K_3).$$

Paso 4 Tomar $w = w + (K_1 + 2 K_2 + 2 K_3 + K_4) / 6$;

(Calcular w_i)

$$t = a + i h.$$

(Calcular t_i)

Paso 5 SALIDA (t, w) .

Paso 6 PARAR.

5.3 Aspectos generales de la programación estructurada

5.3.1 Lenguaje de programación C++

C es un lenguaje de programación de empleo general, caracterizado por su concisión y por poseer un moderno flujo de control y estructuras de datos, así como un rico conjunto de operadores. C es un lenguaje de relativo "bajo nivel". Esta tipificación no es peyorativa: simplemente significa que C trabaja con la misma clase de objetos que la mayoría de las computadoras: caracteres, números y direcciones, que pueden ser combinados con los operadores aritméticos y lógicos, utilizados normalmente en las máquinas. C no contiene operaciones para trabajar directamente con objetos compuestos, tales como cadenas de caracteres, conjuntos, listas, arreglos o vectores, considerados como un todo. El lenguaje no tiene definida ninguna posibilidad de realizar asignación de memoria aparte de las definiciones estáticas y el manejo de pilas para las variables locales de las funciones. C no cuenta con operaciones de entrada-salida: no existen proposiciones tipo READ o WRITE, ni métodos propios para el acceso a archivos. Todos estos mecanismos de alto nivel deben ser aportados por funciones llamadas explícitamente. De la misma manera, C sólo ofrece proposiciones (sentencias) de control de flujo sencillas, secuenciales, de selección, de iteración, bloques y subprogramas pero no multiprogramación, paralelismo, sincronización o corrutinas (Kernigham y Ritchie, 1989).

Aunque las posibilidades de C corresponden a las de muchas computadoras, el lenguaje es independiente de cualquier arquitectura de máquina en particular y así, con ciertas precauciones, es fácil escribir programas portátiles, es decir, programas que pueden ejecutarse sin cambios en una amplia variedad de computadoras (Ibid.).

En C los objetos fundamentales son caracteres, enteros de varios tamaños y números en punto flotante. Además existe una jerarquía de tipos de datos derivados, creados con apuntadores, arreglos, estructuras, uniones y funciones. C posee las construcciones fundamentales de control de flujo sin las cuales no es

posible escribir programas bien estructurados: agrupamiento de sentencias, toma de decisiones (if), ciclos (bucles), con la comprobación de la condición de terminación al principio (for, while) o al final (do) y selección entre un conjunto de casos posibles (switch) (Ibid.).

C incluye apuntadores y capacidad de aritmética de direcciones. Los argumentos de las funciones se pasan copiando su valor, siendo imposible que la función llamada altere el valor del argumento del llamante. Cuando se desea realizar una "llamada por referencia", se puede pasar un apuntador explícitamente, pudiendo la función cambiar el valor del objeto al que se apunta. Los nombres de arreglos se pasan como la dirección de origen del arreglo mismo. Por tanto, los argumentos de tipo arreglo se transmiten por referencia (Ibid.).

Cualquier función puede ser llamada recursivamente y sus variables locales son normalmente automáticas o creadas de nuevo en cada invocación. Las definiciones de función no pueden anidarse, pero las variables pueden ser declaradas como una estructura de bloques. Las funciones de un programa en C pueden compilarse por separado. Las variables pueden ser internas a una función, externas (pero conocidas únicamente a lo largo de un solo archivo fuente) y completamente globales. Las variables internas son automáticas o estáticas. Las variables automáticas pueden colocarse en un registro para incrementar su eficiencia; pero su declaración como registro es únicamente una indicación al compilador y no se refiere a registros específicos de la máquina (Ibid.).

El lenguaje C++ de Borland es una extensión y mejoramiento del lenguaje C. C++ provee además funciones para programación orientada a objetos (McCord, 1991).

Debido a que los lenguajes C y C++ tienen las características de nivel medio y bajo, el programador cuenta con la flexibilidad de la optimización de la estructura de código relativa a los requerimientos de la aplicación. Los lenguajes C y C++ producen un código ejecutable rápido. Los lenguajes C y C++ son muy versátiles; estos proveen de las funciones necesarias que ofrecen la flexibilidad en el diseño e implementación de sistemas de software. El espectro de funciones dadas por C y C++, al ser combinadas con las librerías incluidas en el compilador C++ de Borland, proporcionan un ambiente excelente para el desarrollo de casi cualquier

tipo de software de aplicación (Ibid.).

Al utilizar el lenguaje C++ de Borland, la estructura de todo programa será generalmente la misma. Al igual que con otros lenguajes modulares, existe una función principal que es reconocida por el sistema operativo como la función controladora. Para el lenguaje C++ esta función de control es referida como la función *main()* (Ibid.).

La función *main()*, en operaciones normales, es el punto de inicio y terminación de todo programa. Otras funciones, referidas como subrutinas en otros lenguajes, son llamadas por *main()*. Estas funciones, a su vez, pueden llamar a otras, y así sucesivamente. Cuando cada función ha completado su tarea, el control es regresado a la función llamante. De esta manera, el control siempre regresa a *main()* (Ibid.).

La figura 8 ilustra un ejemplo de relación entre *main()* y las funciones de soporte.

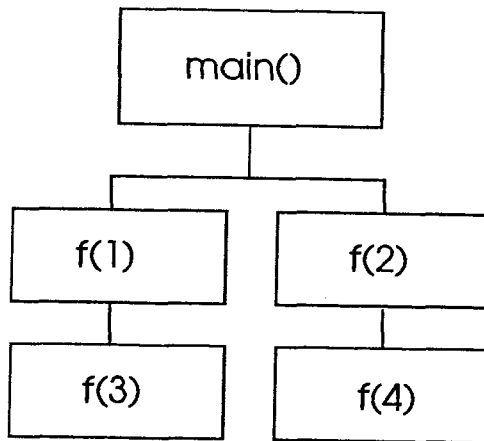


Figura 8. Relación de la función *main()* con el programa estructurado (tomado de McCord, 1991).

El lenguaje C++ está diseñado para ser un lenguaje modular; esto le aporta numerosas ventajas. El diseño modular elimina la redundancia de código. Es más fácil y eficiente producir una función que pueda ser llamada para realizar una tarea particular que añadir las líneas requeridas cada vez que se le necesite. El código modularizado es más fácil de leer y entender, lo que facilita la aplicación de mantenimiento de programas complicados (Ibid.).

5.3.2 Modularidad

Los sistemas modulares constan de unidades claramente definidas y manejables con las interfases claramente definidas entre los diversos módulos. Un sistema modular cumple con los siguientes criterios:

-
- Cada diseño de un proceso es un subsistema claramente definido y con el potencial de ser útil para otras aplicaciones.
 - Cada función en el diseño tiene un propósito específico claramente definido.
 - Cada función maneja no más de una estructura de datos principal del sistema.
 - Las funciones comparten datos globales en forma selectiva; es ciertamente fácil de identificar todas las rutinas que comparten una estructura de datos principal.
-

La modularidad mejora la claridad del diseño, que a su vez facilita la instrumentación, la depuración, la documentación y el mantenimiento del sistema (Guerrero, 1992).

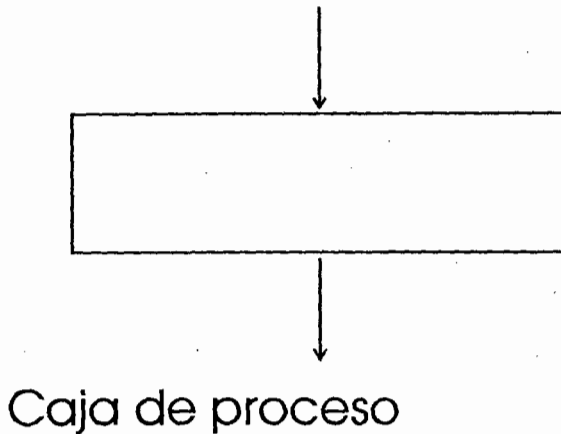
5.3.3 Bloques básicos para la construcción de un programa

Bohm y Jacopini (sin año) (cit. por Guerrero, 1992) afirman que todo diseño necesita sólo de tres bloques básicos para la construcción de un programa y son:

1. Caja de proceso.
2. Mecanismo de ciclo.
3. Mecanismo de decisión binaria

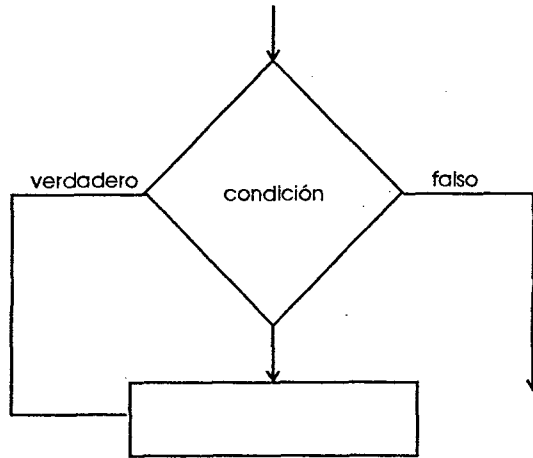
1. Caja de proceso.

Una caja de proceso es cualquier instrucción, la cual realiza o ejecuta alguna operación e inmediatamente procede a la próxima instrucción; su representación gráfica es la siguiente:



2. Mecanismo de ciclo.

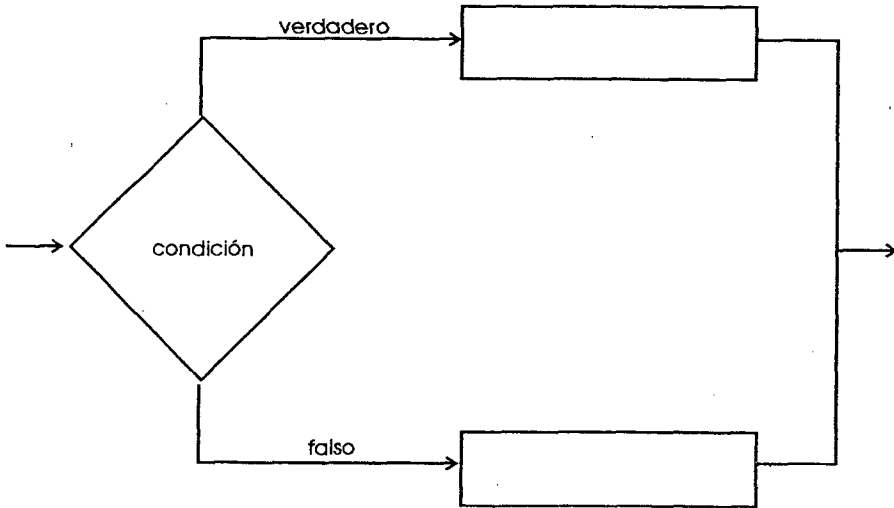
Repite un conjunto de acciones, mientras una proposición condicional es verdadera. Es con frecuencia referenciado como un mecanismo DO{conjunto de acciones}WHILE{expresión condicional} o FOR(valor inicial; sentencia de evaluación; sentencia de incremento){conjunto de acciones}, siendo su representación gráfica la siguiente:



Mecanismo de ciclo

3. Mecanismo de decisión binaria.

Ejecuta una de dos tareas, en dependencia de la verdad o de la falsedad de una proposición condicional. Con frecuencia es referenciado como un mecanismo IF(proposición condicional) {conjunto de acciones} ELSE {conjunto de acciones}, siendo su representación gráfica la siguiente:



Mecanismo de decisión binaria

Los dos mecanismos anteriores se pueden pensar como una caja de proceso, ya que solamente tienen una entrada y una salida. También se puede transformar cualquier secuencia lineal de cajas de procesos en una sola caja de proceso, simplemente más compleja.

5.3.4 Pruebas y mantenimiento

Según Guerrero (1992) hay cuatro pruebas que un sistema debe de satisfacer: pruebas funcionales, pruebas de desempeño, pruebas de tensión y pruebas estructurales. Las pruebas funcionales y de desempeño se basan en las especificaciones de requisitos. Los casos de pruebas funcionales especifican las condiciones operativas normales, los valores comunes de entrada y los resultados esperados. Las pruebas funcionales también deben de contemplarse para probar condiciones de frontera. También deben probarse archivos y arreglos que contengan valores idénticos, la matriz cero, etc.

Las pruebas de desempeño deben de contemplarse para verificar el tiempo de respuesta, tiempo de ejecución, capacidad de procesamiento, utilización de memoria y transferencias de datos.

Cada prueba funcional y de desempeño debe de especificar la configuración de la máquina, las suposiciones concernientes al estado del sistema para el caso de prueba, los requisitos que están sometidos a prueba, las entradas de prueba y los resultados esperados. Es particularmente importante que los resultados esperados de cada prueba se especifiquen previamente a la instrumentación del sistema.

El propósito de las pruebas de tensión es determinar las limitaciones del sistema y cuando el sistema falla, determinar la manera en que se manifiesta la falla. Las pruebas de tensión pueden proporcionar un conocimiento valioso relacionado con la fortaleza y debilidad de un sistema. Las pruebas de tensión se obtienen de los requisitos, el diseño y de la intuición del diseñador.

Las pruebas estructurales están relacionadas con el examen de la lógica del procesamiento interna de un sistema de programación. El objetivo de las pruebas estructurales es recorrer un número específico de caminos lógicos a través de cada subrutina en el sistema, para establecer la profundidad de la prueba.

Por otra parte, las actividades de mantenimiento implican mejorar el sistema, adaptarlo a nuevos ambientes y corregir errores. La mejora del sistema puede dar como resultado proporcionar nuevas capacidades funcionales, mejorar los despliegues al usuario y los modos de interacción, revalorar los documentos externos y la documentación interna (Guerrero, 1992).

6. Resultados

A continuación se exponen los resultados obtenidos, referidos a los mismos puntos planteados en la metodología de este trabajo.

El método numérico empleado fué el de 4o. orden de Runge-Kutta. Los modelos matemáticos empleados para la simulación fueron:

- Modelo de crecimiento microbial de Monod
- Modelo para la difusión pasiva a través de una membrana
- Ley del enfriamiento de cuerpos de Newton
- Modelo de crecimiento de peces de Von Bertalanffy
- Modelo compartamental simple
- Modelo de poblaciones de Verhulst-Pearl
- Modelo de depredación de Lotka-Volterra
- Modelo de epidemias de Bailey

Una última opción en las funciones del programa fué la del ingreso de ecuaciones diferenciales no incluidas en el menú ofrecido. Estas ecuaciones tendrán que ser del tipo que resuelve el método Runge-Kutta, y tendrán que ser alimentadas conforme a las reglas establecidas en el menú de ayuda de esa sección.

La estructura del programa se dividió en "subrutinas" (en el sentido que este término toma en programación estructurada) con identidad funcional, conforme a los principios de la modularidad expuestos en este trabajo. Estos módulos pueden ser identificados en el listado del programa, ya que todos están etiquetados.

Las funciones de programación empleadas pueden ser observadas en el listado del programa, ya que su descripción, por su amplitud y profundidad, escapan a los objetivos de este texto. En capítulo "Aspectos generales de la programación estructurada" se hacen sinopsis de las características del flujo de información en un programa estructurado y en particular del lenguaje C++.

Al final del subcapítulo dedicado a los métodos numéricos se muestra el algoritmo del método Runge-Kutta de 4o. orden en pseudocódigo de programación. Los modelos matemáticos fueron implementados directamente a lenguaje C++.

Características del programa

El programa es ejecutable desde sistema operativo; el archivo en el que se encuentra tiene como nombre SIMBIO con extensión .EXE; va acompañado por varios archivos más en los que se incluyen elementos necesarios para su ejecución, por lo que se recomienda incluirlos a todos ellos en un directorio propio.

El programa funciona en modo gráfico, para lo cual realiza una detección del tipo de controlador para gráficos instalado en la computadora en que se esté ejecutando. De existir un problema de inicialización del modo gráfico, en la pantalla se mostrará un mensaje comunicándolo. Al comenzar la ejecución, una pantalla de presentación muestra el nombre del programa, su versión y otros datos (el programa es susceptible de mantenimiento, según el significado dado para este término en el documento, por lo que pueden darse posteriores versiones del mismo). Aparece después un menú principal en el que se muestran todas las opciones arriba mencionadas; éstas pueden ser activadas por medio del tablero o del *mouse*. Una vez activada una opción, se piden los datos iniciales implicados en ésta. De ahí, se pasa a la pantalla de graficación en la que se mostrarán los valores generados paralelamente al trazado gráfico de la variable de estado en cuestión contra el tiempo. Tras la graficación se ofrecen opciones de impresión; de no existir total compatibilidad entre los controladores para impresoras incluidos en el programa y la impresora del usuario, se puede emplear con este fin la tecla "Print Screen" ("Impr Pant") del tablero, que puede ofrecer una resolución aceptable en tonos de gris. En todas las pantallas se cuenta con la posibilidad de "cajas de ayuda", en las que se explicarán de forma resumida las funciones de la pantalla activa.

7. Discusión

No resulta del todo fácil incursionar en el papel de pionero local en las tan promisorias tareas de la simulación por computadora en el área biológica; este enfoque no ha sido muy considerado por ser prácticamente desconocido por la tradicional mayoría de investigadores y docentes; este hecho proviene, posiblemente, de la deficiencia de bibliografía, escasez de equipo de cómputo apropiado, pobreza curricular en ciencias exactas y de la computación en la formación de biólogos, entre otros factores. Tal como lo afirman Keen y Spain (1992), los biólogos que intenten llevar a cabo cualquier investigación cuantitativa deben de contar sólidos antecedentes en matemáticas, estadística y técnicas de simulación, y la simulación debe jugar un papel activo en cualquier programa de investigación cuantitativa, como herramienta fundamental en el desarrollo y validación de modelos de la realidad. Un primer paso para acercarse a tal contexto es la investigación interdisciplinaria en la que participen biólogos y profesionistas de las áreas matemáticas, físicas, químicas e informáticas. Otro medio indispensable es la incorporación local de investigadores formados en centros donde se cuente ya con firmes antecedentes en la aplicación de las matemáticas y computación electrónica al estudio de los sistemas biológicos.

Debe hacerse notar que según numerosos autores (Law y Kelton, 1991; Naylor *et al.*, 1986; Karplus y Petsko, 1990; Keen y Spain, 1992) una de las principales ventajas extraídas de las actividades de modelado y simulación es la de desmembrar al fenómeno en estudio, descubrir las relaciones formales entre sus componentes y, por ende, penetrar en sus esencias funcionales, acercándose cada vez más y más a su comprensión más clara. Este papel de la modelización y la simulación es particularmente valioso para los biólogos en formación, ya que es capaz de vincularlos de la manera más precisa a sus objetos de estudio.

En los primeros capítulos de este trabajo se expone la génesis, desarrollo y trascendencia de los modelos conceptuales, matemáticos y la simulación por computadora. El proceso de la simulación expande el poder del ciclo de investigación cuantitativa mediante el uso de del modelo matemático para generar

datos simulados que pueden ser comparados con los datos reales para encontrar diferencias. Asumiendo que el modelo fué propiamente programado para la computadora, estas diferencias indican problemas con la validez de las presunciones usadas en la formulación del modelo matemático, y esto repercute a su vez en el modelo conceptual. Los errores pueden ser tan grandes que requieran de un modelo conceptual completamente nuevo, o de una magnitud tal que se requieran solamente ligeras modificaciones en los coeficientes utilizados en el modelo matemático. En cualquier caso, siempre se generan nuevos cuestionamientos y se necesita de nuevos experimentos. La afinación de una simulación puede producir mejoramientos consecuentes en los modelos conceptual y matemático.

Así, la simulación permite tomar la mayor ventaja del ciclo de investigación cuantitativa. Su más prominente ventaja reside en la provisión de una prueba clara e imparcial del pensamiento involucrado en la formulación del modelo matemático, y por lo tanto del modelo conceptual. La simulación puede por lo tanto jugar un papel clave en la comprensión de los sistemas biológicos.

Si este fuera el único valor de la simulación y la modelización en la investigación biológica, habría ya de por sí suficiente justificación para su uso. Sin embargo, la simulación aporta otro importante beneficio: permite la realización de "experimentos" de sistemas con condiciones que se ubican fuera del rango de lo posible. Permite hacerse preguntas como "¿qué pasaría si ...?". Mediante la simulación de varias estrategias de manejo del modelo matemático, se es capaz de hacer predicciones inteligentes acerca de su éxito o falla en la representación del sistema real.

Para la realización de simulaciones por computadora puede ser utilizado prácticamente cualquier lenguaje de programación. Existen lenguajes diseñados para ser aplicados específicamente a la simulación de ciertos tipos específicos de sistemas (Naylor *et al.*, 1986), entre los que no se encuentran los biológicos. Es frecuente la observación del uso del lenguaje de alto nivel BASIC, en simulaciones ingenieriles, por parte de investigadores de esta Universidad; igualmente Keen y Spain (1992) sugieren su empleo para la simulación en biología, debido, ante todo, por su presencia universal (BASIC se incluye en el sistema operativo de las PC compatibles con IBM). Law y Kelton (1991) recomiendan el uso de lenguajes estructurados como Pascal, C o Fortran, y ofrecen ejemplos de su aplicación a la simulación. C, al ser un lenguaje de nivel medio o bajo permite un acercamiento

mayor a la utilización de todas las posibilidades del ordenador. Turbo C++ es una versión relativamente reciente de C, equipada con funciones para programación orientada a objetos y con bibliotecas bien dotadas. Entre éstas, se encuentra la biblioteca de funciones matemáticas (math.h), que facilita sustancialmente el desarrollo de programas para simulación. Schildt (1990), aplica Turbo C++ a la simulación de sistemas estocásticos y considera que, dada la riqueza de las bibliotecas presentes en el compilador de Borland (entre muchas otras características de C, expuestas en este documento), "no se justifica el uso de cualquier otro lenguaje con prácticamente cualquier fin".

Los modelos que se implementaron en el programa son un reflejo de la generalidad de aspectos de los sistemas vivientes que pueden ser abordados para su simulación. Se escogieron entre los ofrecidos por la bibliografía consultada de acuerdo a su aplicabilidad genérica. Se descartaron modelos que pudieran describir aspectos muy particulares de los fenómenos representados. Los modelos utilizados se implementaron para proporcionar al usuario la comodidad de la inmediatez, pero la novena opción del programa es capaz de cubrir cualquier planteamiento particular de estos modelos que quisiera ser resuelto o cualquier modelo distinto a los incluídos que tenga la forma de los solucionables por el método utilizado. Por lo tanto, pese a la breve representatividad que tales modelos hacen de la vastedad de los posibles, el programa tiene una aplicación muy extensa.

El programa está diseñado para ser manejado fácilmente; sus funciones, ante el usuario, se activan y fluyen de manera estandarizada al *software* comunmente observado en cualquier microcomputadora. El programa no cuenta con mecanismos que garanticen en cien por ciento que no existirán detenimientos repentinos de la ejecución o disfunciones debido a un manejo no adecuado o ingreso de datos fuera del rango de los normalmente pertinentes a los modelos incluídos.

8. Conclusiones

- Es factible el desarrollo de medios computarizados para el estudio de variados aspectos y niveles de organización de los sistemas biológicos.
- Las tareas de construcción de modelos y su conversión a un lenguaje de programación exigen una posición sumamente analítica ante la naturaleza de los fenómenos modelados; estas actividades conducen a una comprensión más aguda del fenómeno.
- El desarrollo de programas de computadora capaces de producir datos simulados ofrece la posibilidad de manipulación teórica del proceso estudiado.
- La manipulación teórica permite al investigador o estudiante hacer predicciones justificadas acerca del comportamiento del fenómeno de interés y, por lo tanto, puede apoyar la toma de decisiones en el direccionamiento de la investigación o el manejo de recursos naturales.
- Resultó satisfactoria la implementación de los módulos funcionales debido a que facilitan las tareas de programación, permiten al programa fluir sin ambigüedades y dan la posibilidad de mantenimiento posterior.
- La utilización de un lenguaje de relativo "bajo nivel" con fines de simulación permite eficientar los cálculos implicados, y proveer de características de operación avanzadas.

LITERATURA CITADA

- Barkakati, N.; 1990; The Waite Group's Turbo C++ Bible; 1a. ed.; Ed. SAMS; U.S.A; 1084 pp.
- Bronson, R.; 1984; Ecuaciones diferenciales modernas; 1a. ed.; Ed. McGraw-Hill; México; 310 pp.
- Bunge, M.; 1959; La ciencia, su método y su filosofía; Ed. Quinto Sol; México; 100 pp.
- Burden, R.L. y J.D. Faires; 1985; Análisis numérico; 3a. ed.; Grupo editorial Iberoamérica; México; 721 pp.
- Derrick, W.R. y S.I. Grossman; 1984; Ecuaciones diferenciales con aplicaciones; 2a. ed.; Ed. Fondo Educativo Interamericano; México; 642 pp.
- Díaz Barriga Arceo, J.; 1992; Elementos de la solución de problemas por computadora. En: J. Díaz Barriga Arceo y M. Guerrero Zarco (compiladores); Computación. Temas selectos; 1a. ed.; UNAM; México; 204 pp.
- Ezzel, B.; 1990; Graphics programming in Turbo C++. An object-oriented approach; 1a. ed.; Ed. Addison-Wesley; U.S.A.; 544 pp.
- Guerrero Zarco, M.; 1992; Programación y diseño estructurado. En: J. Díaz Barriga Arceo y M. Guerrero Zarco (compiladores); Computación. Temas selectos; 1a. ed.; UNAM; México; 204 pp.
- Hutchinson, G.E.; 1978; An introduction to population ecology; Yale University Press; U.S.A.
- Karplus, M. y G.A. Petsko; 1990; Molecular dynamics simulations in biology; Nature 347(6294): 631 - 639.
-

- Keen, R.E. y J.D. Spain; 1992; Computer simulation in biology: a BASIC introduction; 2a. ed.; Ed. Wiley-Liss; U.S.A.; 497 pp.
- Kernigham, B.W. y D.M. Ritchie; 1989; El lenguaje de programación C; 1a. ed.; Ed. Prentice Hall; México; 235 pp.
- Law, A.M. y W.D. Kelton; 1991; Simulation modeling and analysis; 2a. ed.; Ed. MacGraw-Hill International; Singapore; 759 pp.
- Lehninger, A.L.; 1975; Biochemistry; 1a. ed.; Worth Publishers; U.S.A.; 1117 pp.
- Leveau, J.I. y M. Bouix; 1985; Cinéticas microbianas; En: René Scriban (compilador); Biotecnología; 2a. ed.; Editorial el manual moderno; México; 669 pp.
- López de Medrano, S.; 1973; Lenguajes Simbólicos; 1a. ed.; ANUIES; México; 63 pp.
- López de Medrano, S.; 1985; Modelos matemáticos; 2a. ed.; Ed. Trillas; México; 47 pp.
- McCord, J.; 1991; Borland C++ programmer's guide to graphics; 1a. ed.; Ed. SAMS; U.S.A.; 519 pp.
- Morrison, R.T. y R.N. Boyd; 1987; Química orgánica; 2a. ed.; Ed. Addison-Wesley Iberoamericana; México; 1375 pp.
- Naylor, T.H., J.L. Balintfy, D.S. Burdick y K. Chu; 1986; Técnicas de simulación en computadoras; 1a. ed.; Ed. LIMUSA; México; 381 pp.
- Schildt, H.; 1990; Turbo C: programación avanzada; 2a. ed.; Ed. Borland Osborne/McGraw-Hill; España; 353 pp.
- Schildt, H.; 1990; Using Turbo C++; 1a. ed.; Ed. Borland-Osborne/McGraw-Hill; U.S.A.; 755 pp.
- Smith, W.A.; 1989; Análisis numérico; 1a. ed.; Ed. Prentice Hall; México; 608 pp.
-

-Solomon, E.P., C.A. Villet y P.W. Davis; 1987; Biología; 1a. ed.; Ed. Interamericana; México; 1342 pp.

-The Open University; 1974; Ecuaciones diferenciales I; 1a. ed.; Ed. McGraw-Hill Latinoamericana; Colombia; 45 pp.

-Von Bertalanffy, L.; 1986; Teoría general de los sistemas; 1a. ed.; Ed. Fondo de Cultura Económica; México; 311 pp.

-Wagner-Dobler, F.; 1987; Lenguaje C; 1a. ed.; Ed. Addison-Wesley Iberoamericana; México; 116 pp.