

I.1 HERRAMIENTAS DE APOYO A LA TOMA DE DECISIONES EN LA GESTIÓN DEL MEDIO NATURAL

Comenzaremos esbozando algunas ideas generales sobre la teoría de la decisión, haciendo hincapié no tanto en los métodos estadísticos que la soportan, sino en los conceptos que pueden resultar de interés para nuestros objetivos. A continuación definiremos el concepto de Sistema de apoyo a la toma de decisiones (en adelante **SAD**), como generalización y sistematización de la aplicación de la teoría de la decisión a un ámbito concreto. Describiremos cómo los SAD suministran, entre otras cosas, una aproximación común a la hora de abordar problemas complejos relacionados con la toma de decisiones. El análisis de esta aproximación basada en la teoría de la decisión, nos permitirá asemejar lo anterior a la teoría de modelos. Es decir, veremos cómo la resolución de un problema determinado usando los preceptos de la teoría de la decisión, implica la elaboración de modelos más o menos complejos que permitan explicar cómo funciona el proceso que está implicado en el problema decisional inicial. Describiremos entonces los distintos tipos modelos espaciales aplicados a la gestión del medio natural que han sido evaluados durante este trabajo.

I.1.1 La teoría de la decisión

Según el Diccionario de la Real Academia de la Lengua, **decisión** es una resolución que se toma o se da en una cosa dudosa. Cuando se plantea un problema en cualquier ámbito, casi siempre nos vemos obligados a tomar una decisión para tratar de resolverlo. Sin embargo, en el ámbito que más nos interesa aquí (toma de decisiones en la gestión del territorio), el concepto de decisión suele implicar una cierta distribución de recursos en un sentido determinado (Spradlin 1997). Si ante un problema de riesgo de extinción de una especie, decidimos diseñar un plan de cría en cautividad, es necesario dotar con una serie de recursos dicha decisión. La persona o personas capaces de tomar decisiones suele denominarse **centro decisor**. Al aumentar el nivel jerárquico del centro decisor lo hacen también los riesgos asumidos al tomar la decisión y la cantidad de recursos puestos en juego en

la misma. Este centro decisor suele tomar decisiones en función de una serie de **objetivos**, que no son más que lo que se pretende alcanzar mediante la distribución de recursos que implica la toma de decisiones.

Hasta ahora podemos comparar dos conceptos que suelen confundirse. Una cosa es una decisión y otra muy diferente un objetivo. Por ejemplo, evitar la extinción del lince, no es una decisión, sino un objetivo. Para lograr dicho objetivo, habrá que tomar muchas decisiones. Es posible que las decisiones sean correctas o que no lo sean. Para intentar que se dé la primera situación, los centros decisores se valen de una serie de **criterios** (que pueden ser propios o bien pueden venir dados desde fuera), definiéndose éstos como los aspectos que son considerados importantes para tomar la decisión. Y aquí es donde empieza a complicarse la situación. Los criterios pueden variar en función de la personalidad de cada centro decisor, o en función de su formación, intereses personales, obligaciones políticas o sociales, etc. Son los criterios los que incorporan un mayor grado de discrecionalidad al proceso de toma de decisiones, y esto complica enormemente la modelización del proceso.

El proceso mediante el cual, el centro decisor combina sus objetivos, con los criterios que tiene para lograrlos, y que conduce a una decisión final, se denomina **análisis de la decisión**. De forma más o menos consciente y estructurada todos llevamos a cabo un análisis de la situación antes de tomar una decisión determinada. Evaluamos las ventajas e inconvenientes de una decisión, analizamos las posibles consecuencias, los costes que ocasionará, etc. En definitiva, **simulamos** mentalmente el proceso de toma de decisiones, generando un **modelo** que nos permite evaluar *a priori* el resultado de la decisión. Volviendo al Diccionario de la Real Academia, observamos (entre otras acepciones) que un modelo es una representación en pequeño de alguna cosa. En efecto, un modelo no es más que una simplificación de una realidad compleja que resulta difícil estudiar en su conjunto. Mediante modelos lógicos o matemáticos se pueden representar las relaciones entre los distintos factores que afectan a la decisión que debemos tomar. Así podemos aprender cómo funcionan estas relaciones, al mismo tiempo que podemos estimar las posibles implicaciones de la toma de decisiones.

Hay otros conceptos muy importantes a la hora de describir el proceso de toma de decisiones. En primer lugar necesitamos contar con varias **alternativas** para compararlas y decidir cuál es la que mejor se ajusta a los objetivos. El proceso de análisis consiste en comparar varias alternativas que se diferencian en la forma en la que se combinan los criterios, la importancia relativa de cada uno, etc. Una vez que nos decidimos por una alternativa, surge uno de los grandes problemas del proceso de toma de decisiones: la **incertidumbre**. A veces se le llama suerte, pero en realidad la incertidumbre es un concepto inherente a todo sistema complejo (Schrödinger 1988, Manson 2001). La incertidumbre surge de la incapacidad que tenemos de controlar todas las variables que están implicadas en un proceso concreto. Aunque es imposible evitarla, podemos evaluar la incertidumbre mediante métodos estadísticos de complejidad variable. Lofti Zadeh plasmó la incertidumbre en su principio de incompatibilidad (Zadeh 1973): “conforme aumenta la complejidad de un sistema, se reduce nuestra capacidad de hacer afirmaciones precisas y suficientemente significativas sobre su comportamiento, hasta llegar a un umbral sobre el cual la precisión y la relevancia son conceptos mutuamente excluyentes”. La combinación del proceso de análisis de la decisión, con la incertidumbre inherente al sistema, dan lugar a un **resultado**. El hecho de que la incertidumbre forme parte del resultado hace que éste también sea incierto. En otras palabras, no podemos asegurar que el resultado se ajuste al objetivo por muy elaborado que sea el proceso de análisis de la situación (Gough & Ward 1996). Por tanto, cuando el centro decisor opta por una alternativa, asume un cierto **riesgo** de que el resultado no se ajuste a sus objetivos.

La descripción del proceso de toma de decisiones que hemos realizado es la base de la **Teoría de la decisión** (Romero 1996). Esta rama científica pretende sistematizar el proceso de toma de decisiones en los seres humanos para poder simularlo y aplicarlo a la resolución de problemas concretos (Dale & English 1999). De manera algo más formal, la teoría de la decisión es una teoría axiomática que permite tomar decisiones en condiciones de incertidumbre (Gough & Ward 1996). Fue a partir de la década de 1960 cuando empezaron a realizarse aplicaciones de la teoría de la decisión, sobre todo en cuestiones económicas relacionadas con la maximización del rendimiento de inversiones y otras situaciones similares. Desde el

punto de vista metodológico, la teoría de la decisión se basa en buena medida en las técnicas estadísticas y en la teoría económica.

La simulación del proceso de análisis decisional requiere una simplificación determinada del sistema. Como hemos visto antes, se trata de modelizar el proceso de toma de decisiones. De hecho, el análisis de la dinámica de los sistemas complejos es similar al análisis de los procesos de toma de decisiones (Gough & Ward 1996). En este sentido, surgen dos "escuelas" que tratan de analizar el proceso de toma de decisiones desde perspectivas casi opuestas.

Por un lado tendríamos la aproximación "dura", que se basa en la existencia de dos estados o posiciones en el sistema analizado: el estado actual y el estado deseable. Esta aproximación propone que la forma de llegar del primer al segundo estado consiste en la toma de una serie de decisiones que maximicen la eficiencia de manera individual. La conjunción de los resultados de estas decisiones conducirá al sistema al estado deseado. Esta aproximación se aplica sobre todo a cuestiones económicas donde la eficiencia y la rentabilidad son los aspectos más importantes. La aproximación "blanda" es más asimilable a sistemas naturales: contempla la existencia de objetivos ambiguos y a veces contrapuestos. La mayoría de los sistemas ecológicos deben ser analizados bajo esta perspectiva. Aquí no tenemos un único objetivo claro, sino que éstos pueden cambiar conforme vayamos enriqueciendo el modelo y vayamos aprendiendo más sobre su dinámica. Es decir, los resultados del proceso de análisis de la decisión son comparados con la realidad para redefinir los objetivos. Se trata de conseguir una mejora continua tendente a la resolución de un problema decisional concreto (Gough & Ward 1996).

I.1.2 Sistemas de apoyo a la toma de decisiones

Los Sistemas de Apoyo a la toma de Decisiones, pueden definirse como un conjunto de procedimientos o mecanismos que facilitan el proceso de toma de decisiones (Davis 1998) utilizando para ello métodos procedentes de la teoría de la decisión. De manera algo más detallada, podemos decir que los SAD combinan la metodología del análisis de la decisión y los procedimientos para el manejo y de la

información (procedentes de los Sistemas de Información). Es decir, la esencia de los SAD es la integración de la información mediante metodologías procedentes de diferentes fuentes, con el propósito de asistir al manejo de los sistemas complejos (Djokic 1996, Frysinger y col. 1996).

Este tipo de sistemas pueden estar basados en aplicaciones informáticas, o pueden ponerse en marcha en papel, o "simplemente" pueden implementarse en la forma de trabajar de un grupo multidisciplinar dado. Actualmente hay múltiples ejemplos de SAD aplicados a la gestión ambiental. En general predominan aquellos diseñados para ayudar al manejo de los recursos hídricos (Gough & Ward 1996, Giupponi y col. 2002).

Si nos centramos en los SAD basados en las nuevas tecnologías de la información, podemos definirlos como un conjunto de aplicaciones informáticas que utilizan métodos analíticos procedentes de la teoría de la decisión, más una serie de algoritmos de optimización, procedimientos de integración de información, con objeto de generar modelos que asisten a los decisores a formular alternativas, analizarlas e interpretarlas (Ostfeld 2001).

Además de la anterior interpretación de los SAD, hay autores que los definen de una manera más "logística". En este sentido, Rogers (Rogers 1995), considera que el SAD es el puente tecnológico necesario entre gestión e investigación. En este contexto, los resultados de la investigación no llegan directamente al gestor, sino que previamente pasan por el SAD, donde se procesan y se hacen accesibles a los técnicos responsables. Asimismo, las propuestas o las necesidades de los gestores son transmitidas a los científicos, mediante este sistema de apoyo a la toma de decisiones. En este sentido, podemos decir que los SAD podrían aportar lo siguiente:

- Suministran un marco operativo para establecer y evaluar objetivos de gestión aceptables para la sociedad y asequibles científicamente. Estos objetivos deben de ser definidos en función de los conceptos ecológicos actualmente vigentes.
- Aportan un modelo predictivo de la naturaleza, dirección y tasa de cambio del sistema. Al mismo tiempo pueden actuar para evaluar las consecuencias de la gestión del medio natural.

- Por último proporcionan un marco de gestión adaptativa en el que se evalúan de forma dinámica tanto la gestión como la presión ejercida por las actividades humanas sobre el medio.

Dentro de un SAD, la investigación es el proceso mediante el cual se van generando datos necesarios para la gestión. Conforme la calidad de estos datos se va incrementando, también lo hace la calidad de la gestión realizada. Por lo tanto, las necesidades de los gestores en cuanto a información científica se refiere, se van convirtiendo en objetivos concretos de los científicos.

Antes de continuar describiendo los aspectos generales de los SAD, es necesario aclarar que estos sistemas no pretenden en ningún momento suplantar la decisión humana ni siquiera automatizarla. Su objetivo es poner a disposición de los gestores una serie de herramientas que pueden facilitar el complejo proceso de la toma de decisiones. Pero, ¿son realmente necesarios los SAD?. Su uso se justifica en buena medida porque es prácticamente imposible que una única persona o un grupo multidisciplinar sean capaces de llevar a cabo el proceso de integración de las grandes cantidades de datos, los objetivos complejos y contrapuestos, y la inherente complejidad del sistema, necesario para tomar una decisión con posibilidades de ser correcta (Rauscher 1999, Vacik & Lexer 1999). Además de los impedimentos cognitivos anteriores (no somos capaces de integrar tanta información), hay otra serie de problemas que hacen necesario el uso de los SAD. Por un lado están los problemas económicos. Normalmente no disponemos de los suficientes recursos como para llevar a cabo caracterizaciones del medio lo suficientemente detalladas como para obviar el uso de los SAD. Tampoco hay recursos humanos suficientes como para crear múltiples equipos multidisciplinarios que asistan a la toma de decisiones. Por último, los SAD permiten incorporar el factor tiempo en el proceso de modelización de la decisión. Podemos evaluar cómo responderá el sistema en distintas situaciones temporales. Esto no es posible sin un SAD. En general los SAD mejoran considerablemente los resultados de la toma de decisiones incidiendo en los siguientes aspectos (Rauscher 1999, Vacik & Lexer 1999):

1. Facilitan el proceso de identificación de los factores implicados en el problema a resolver.

2. Asisten a la hora de elegir entre las posibles alternativas antes de tomar la decisión.
3. La sistematización del proceso de toma de decisiones hace que éste sea reproducible con facilidad. El proceso está documentado, es transparente y puede ser evaluado. Esto reduce notablemente la arbitrariedad y subjetividad del proceso de toma de decisiones.
4. Por último y muy importante, aumentan considerablemente la capacidad de explicar y justificar las decisiones que se han tomado. Esto se debe a la capacidad de análisis, despliegue y procesamiento de la información con la que cuentan. De esta manera su uso nos permite conocer mejor el proceso que estamos analizando, lo cual redundará en una mejora de la decisión tomada.

En cualquiera de las dos interpretaciones comentadas, una de las principales aportaciones de los SAD, es la generación de un protocolo de trabajo común a la hora de resolver un problema decisional concreto. Este protocolo se basa en los siguientes pasos (Rogers 1995, Rauscher 1999, Vacik & Lexer 1999, Reynolds 2001):

- 1. Identificación del problema.** Es fundamental conocer de la forma más clara posible cuál es el problema que debemos abordar. Aunque parezca obvio, no siempre es tan fácil identificar los orígenes de los problemas de índole ambiental. Por ejemplo, un problema sería: la Administración ambiental tiene la obligación de proponer una serie de montes públicos para que en ellos se lleven a cabo asentamientos apícolas. Sin embargo, no cuenta con la información necesaria para ello.
- 2. Determinación de objetivos.** Una vez identificado el problema, tenemos una idea aproximada de su importancia y de las posibilidades que tenemos de abordar su resolución. La plasmación de una serie de objetivos realizables es un ejercicio que nos muestra el grado real de comprensión que tenemos del problema. Los objetivos deben de ser claros y concisos. En el ejemplo anterior, uno de los objetivos que nos podemos plantear sería: generación de un modelo que nos permita determinar la potencialidad de los montes para albergar asentamientos apícolas.
- 3. Identificación de los factores que intervienen en el proceso.** El siguiente paso implica la identificación de todos los factores bióticos, abióticos, socioeconómicos, etc., que puedan intervenir en el proceso que estamos

modelando. Este paso es uno de los más importantes, ya que estos factores serán los que se integren para generar un resultado final. Esto suele requerir la consulta a expertos, así como una intensa búsqueda bibliográfica que nos permita documentarnos sobre el funcionamiento del sistema que estamos modelizando. Volviendo al ejemplo anterior, los factores a considerar podrían ser: vegetación, insolación, cercanía a puntos de agua, accesibilidad, temperaturas, precipitaciones, propiedad del suelo, etc.

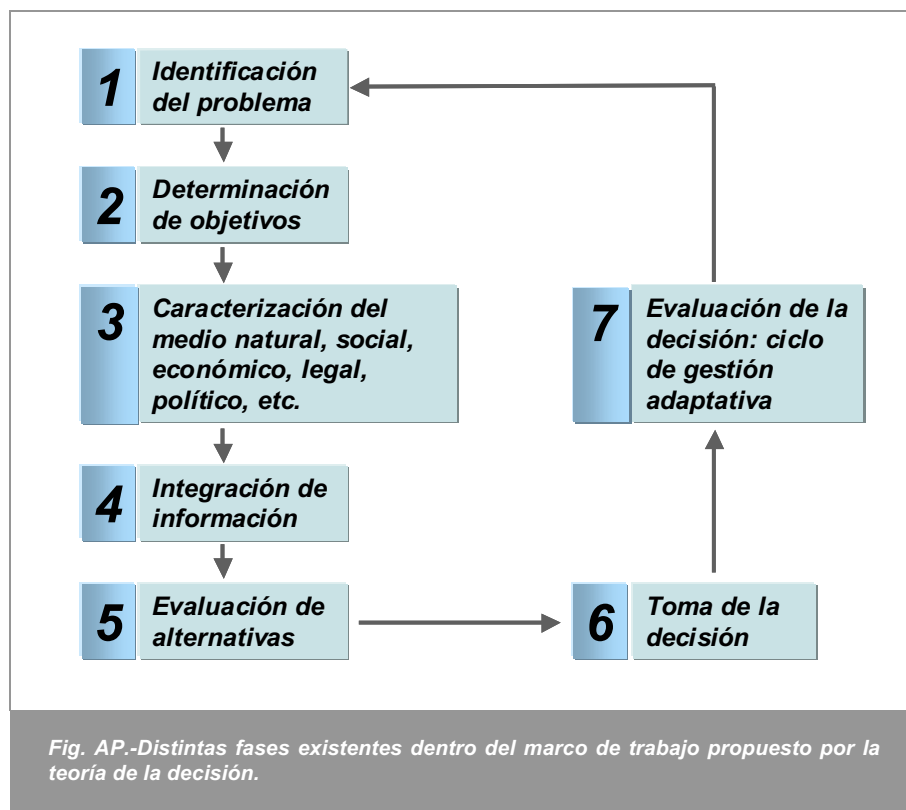
4. Integración de la información. Aquí se da el verdadero proceso de integración de información y es donde se lleva a cabo el análisis decisional propiamente dicho. La integración de toda la información requiere la existencia de una serie de métodos matemáticos relativamente complejos, en los que la lógica difusa (Ekel 1999, Klir 2000, Tyson 2001, Ameskamp & Lamp 2002) juega un papel importante (en los siguientes apartados abundaremos más sobre esto). Estos métodos de integración son enormemente variados, pero casi siempre comparten una característica: tratan de emular la forma de pensar de los seres humanos mediante la elaboración de un modelo conceptual o lógico. Mediante estos modelos se relacionan los factores determinados anteriormente. Así, siguiendo con el ejemplo, el modelo conceptual podría implicar el postulado de una serie de afirmaciones (axiomas): Las zonas aptas para albergar colmenas son aquellas ocupadas por matorrales melíferos **y** están en la solana **y** tienen cerca un punto de agua **o** están cerca de un carril transitable. Dentro de los métodos de integración también podemos encontrar otros más cuantitativos. La denominada ecuación universal para la evaluación de la pérdida de suelo (USLE), podría ser un ejemplo de este tipo (d'Angelo y col. 2000, Santón & García 2001). La aplicación de esta expresión no es más que la integración de una serie de factores (erosividad de la lluvia, erosionabilidad del suelo, etc.) mediante un modelo diseñado empíricamente.

En definitiva, es en este proceso donde se produce el verdadero análisis decisional del que hablábamos al principio. Aquí se plasman los criterios del centro decisor, así como su particular punto de vista sobre el problema considerado. El proceso de modelización resultante dará unos resultados u otros en función de cómo se hayan seleccionado e integrado los factores implicados en el problema a resolver. Esto dará lugar a varias soluciones alternativas, lo que nos lleva a la siguiente fase:

5. Evaluación de alternativas. Como consecuencia de la elaboración de un modelo que trate de resolver un problema ambiental dado, se obtiene una serie de alternativas en función de las distintas aproximaciones realizadas. Y es aquí donde suele acabar la asistencia a la decisión aportada por estas metodologías. Es el momento de que el centro decisor se decante por una u otra alternativa. Cada alternativa se caracterizará por un mayor o menor riesgo, o por diferente grado de incertidumbre, etc. El centro decisor elegirá entre la que más se adapte a sus criterios o necesidades. A pesar de que la última decisión siempre la toma una persona, el procedimiento realizado ofrece una predicción aproximada de las consecuencias de las distintas alternativas. Si continuamos con el ejemplo anterior, podremos obtener varios modelos de aptitud para albergar colmenas en el territorio, en función del tipo de parámetros considerados y de la metodología utilizada para integrarlos. Una alternativa puede considerar que sólo la vegetación es importante, mientras que otra considera que la aptitud final de una zona con vegetación poco adecuada se compensa con la bondad de otros parámetros. El centro decisor evaluará las distintas alternativas y optará por una determinada.

6. Evaluación de la decisión adoptada. Con esta última fase se cierra el ciclo de la asistencia a la toma de decisiones. Cuando el centro decisor opta por una de las alternativas se inicia un proceso de evaluación de las consecuencias de la misma. Esta evaluación es muy importante ya que nos permite aprender mucho sobre la dinámica del proceso modelizado. Se implanta de esta forma un ciclo de gestión adaptativa similar al propuesto por la aproximación ecosistémica para la gestión de los recursos (Montes y col. 1998, Mitchell 1999). En el ejemplo del mapa apícola se evaluará el grado de satisfacción de los apicultores a la hora de poner sus colmenas en las zonas propuestas por el modelo. Si éstos demandan lugares considerados como poco aptos por el modelo, habrá que evaluar con detalle esta situación. Los resultados de esta evaluación retroalimentarán al modelo, lo cual redundará en una mejora sustancial del mismo.

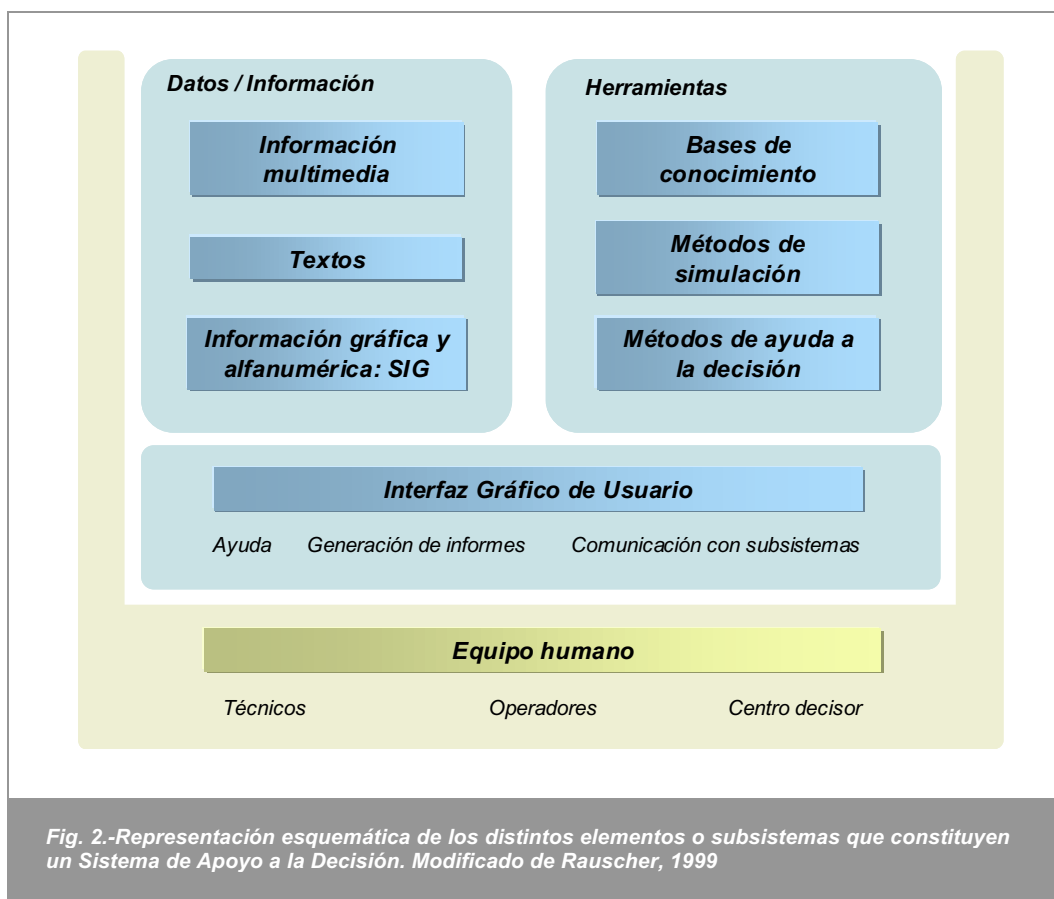
La figura 1 muestra de manera esquemática las distintas fases comentadas anteriormente.



Hasta ahora hemos descrito los aspectos más generales de los SAD, desde su definición hasta el marco de trabajo que proponen para afrontar problemas decisionales. Pero resulta fundamental conocer su estructura interna, así como los elementos necesarios para que estos sistemas sean realmente funcionales. En general podemos decir que un SAD cuenta con un número determinado de subsistemas (Rogers 1995, Rauscher 1999, Potter y col. 2000), los cuales se muestran en la figura 2.

El primero, y el más importante es el compuesto por las personas, es decir, el **centro decisor** propiamente dicho. Sin la "parte humana" no sería posible la existencia de un sistema de apoyo a la decisión porque no habría nadie a quién apoyar. Además de los decisores hay un grupo de técnicos que son los encargados de mantener la parte informática y metodológica del sistema. Éstos ayudan a los decisores a organizar sus ideas, formular relaciones entre los parámetros, a la vez que facilitan la comprensión del problema a abordar. Hay SAD en los que el equipo humano es, prácticamente, el único subsistema realmente estructurado. Se trata de sistemas basados en la negociación y en un pensamiento argumentativo. Esta

metodología se fundamenta en la sistematización del proceso de debate o discusión que puede haber en un grupo multidisciplinar sobre la forma de resolver un problema dado. Así, hay técnicas que evalúan las preferencias de cada miembro del grupo mediante una serie de preguntas concretas. El análisis de las distintas respuestas permite una mayor comprensión del problema y de sus posibles soluciones (Lal y col. 2001).



El siguiente gran subsistema es denominado **subsistema de información**, donde se gestiona la información gráfica y/o alfanumérica. En su interior se organiza la información disponible sobre los procesos ecológicos a los que se refiere el SAD. Exceptuando los casos donde la suerte (o el azar) hace su aparición, la calidad de la decisión que se toma no es nunca mejor que la calidad del conocimiento que hay detrás de la misma. Es decir, una información escasa o incorrecta, dará lugar casi seguro a una decisión equivocada. La información es, por tanto, uno de los factores limitantes en el correcto funcionamiento de los SAD. Esto cobra especial

importancia cuando tratamos de poner en marcha SAD relacionados con la gestión del medio natural, donde como ya comentamos, la escasez de información de calidad es muy importante. Este subsistema de información sería equivalente a los sistemas o redes de información ambientales que proliferan en las distintas administraciones. En ellos se almacena información relativa a todos los aspectos del medio ambiente: suelos, atmósfera, vegetación, hidrología, relieve, vertidos, aprovechamientos, y un largo etc. Asimismo cuentan con bases cartográficas de referencia como ortofotografías aéreas o imágenes de satélite. Pero no basta con tener gran cantidad de información para que un sistema de estas características funcione correctamente. Es básico que la información esté organizada coherentemente. Para ello se requieren potentes herramientas informáticas y físicas que permitan estructurar la información generalmente en sistemas de gestión de bases de datos relacionales. Estas bases de datos pueden contener tanto información gráfica (fundamentalmente a través de los Sistemas de Información Geográfica) como alfanumérica. La adaptación de las metodologías de captura de información en campo a una serie de procedimientos tendentes a su normalización, es también fundamental a la hora de que la información sea realmente utilizable. Es decir, es importante que la información obtenida en campo (un mapa de vegetación, por ejemplo), se ajuste a una serie de criterios que permitan su inclusión en un sistema de información. Este proceso se denomina normalización y permite homogeneizar todas las fuentes de datos que hay en un sistema de información. Todavía no está muy extendido el concepto de que la información que generamos puede ser utilizada por otros para otros fines muy distintos, y por ello se suele considerar que no es importante la normalización.

Siguiendo dentro del subsistema de información, consideraremos otro aspecto de interés. Se trata del grado de precisión de la información y el coste asociado a su consecución. Solemos creer que cuanto más precisa es la información con la que contamos, mejores serán las soluciones que tomemos con dicha información. En principio esto es cierto, aunque hay que hacer una consideración importante a la hora de determinar el coste de adquisición de dicha información. Normalmente el aumento en la precisión lleva acompañado un aumento en el coste de consecución, bien sea en términos económicos o de cualquier otro tipo. Esta relación directa se verifica en la mayoría de los casos. Pero llega un momento en que por mucho que

aumentemos el coste de adquisición de información, el aumento en la calidad de las decisiones adoptadas con tal información, no se realiza con tanta intensidad (Burrough y col. 1996). Es decir, hay un nivel de coste que no merece la pena superar porque no se obtienen mejoras sustanciales en los resultados. Esto se debe a la incertidumbre: al aumentar el grado de conocimiento sobre algo, se reducen los errores debidos al desconocimiento que teníamos del proceso, pero esto no afecta a la incertidumbre, que siempre está presente. La figura 3, muestra esta situación. Observamos que el beneficio de invertir en una información precisa aumenta al hacerlo el coste. Sin embargo, llega un momento en que este crecimiento se hace semiasintótico. Es en este punto donde la incertidumbre se manifiesta, ya que el error debido a la carencia de información se ha reducido notablemente. El punto 1 representa la situación más frecuente en la realidad: el coste en nueva información es bajo y por tanto los errores producidos al generar modelos sobre la dinámica de los ecosistemas son evidentes. El punto 2 representaría una situación ideal en la que hay una información de calidad que permite tomar decisiones con cierta solvencia. En el punto 3 tendríamos una situación en la que los errores del modelo se deberían casi exclusivamente a la incertidumbre inherente al sistema.

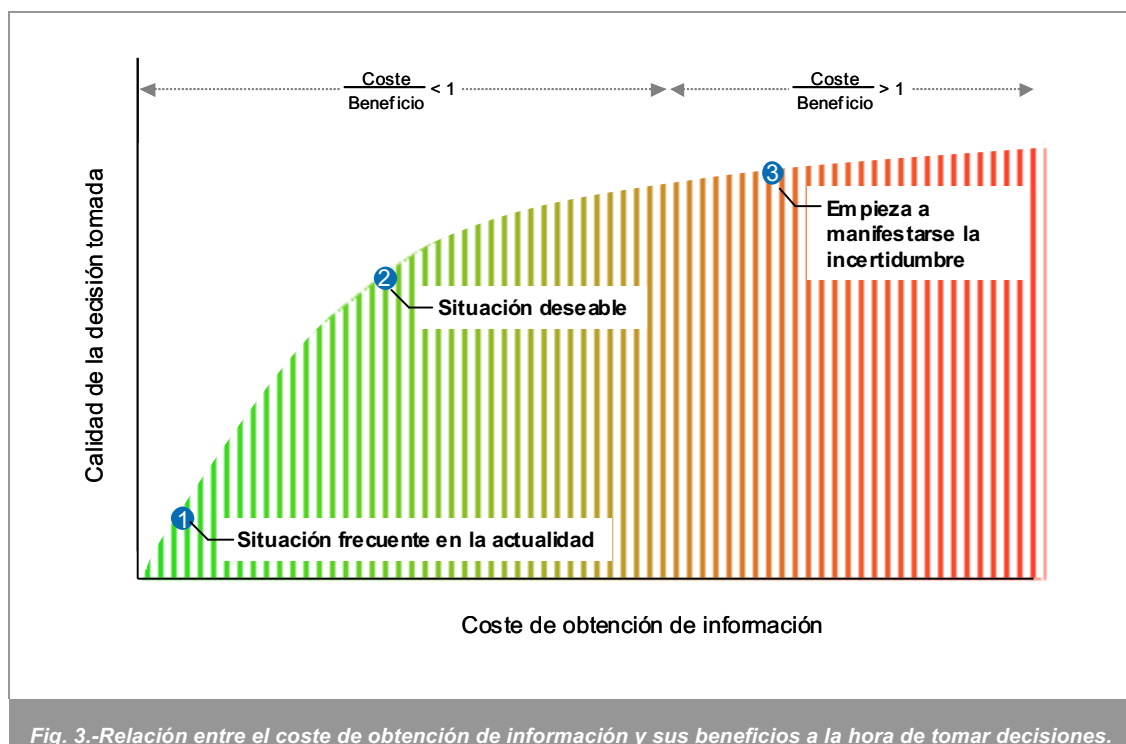



Fig. 3.-Relación entre el coste de obtención de información y sus beneficios a la hora de tomar decisiones.



Pero en un SAD no sólo hay información gráfica o alfanumérica perfectamente estructurada en un SIG o similar. En muchos casos hay gran cantidad de información en forma de fotografías, textos, videos, o bien la experiencia de los individuos constituyentes del SAD (Rauscher 1999). Esta información no reglada o normalizada se suele tratar en cuatro subsistemas similares: el que gestiona las **bases de conocimiento**, el de los **métodos de simulación**, el manejo de la **información de tipo texto** y la **visualización de la información**. La información no basada en el lenguaje se representa en forma de vídeos, fotografías, etc. Es registrada por el subsistema de visualización de la información. Este subsistema es equivalente a un archivo multimedia donde se pueden gestionar, analizar y consultar toda la información de este tipo: fotografías, videos, sonidos, etc. También se incluyen en este subsistema todas las aplicaciones informáticas que permiten mostrar al usuario los resultados del proceso de modelización. En este sentido tiene especial interés la visualización realista en tres dimensiones. El resto de la información, la asimilable a texto escrito, puede gestionarse a través de tres subsistemas. El más importante de ellos es el que gestiona las bases de conocimiento. Una base de conocimiento no es más que el resultado de plasmar en un documento, la información que tenemos sobre un proceso determinado (ecológico en nuestro caso). Es como si plasmáramos en un documento (texto, gráfico, etc.) la experiencia personal que tenemos sobre la dinámica de un proceso (Reynolds 1998b, Reynolds 2001). Hay múltiples formas de implementar el conocimiento que se tiene sobre algo en una base de conocimiento. Se puede realizar mediante modelos numéricos, mediante axiomas, redes neuronales (Deadman & Gimblett 2002), etc. Por último tenemos el subsistema de metodologías de simulación, diseñado para suministrar un marco consistente dentro del cual se organizan los distintos modelos sobre diversos procesos ecológicos. Aquí cabrían todos los modelos matemáticos que nos permiten conocer más cómo funcionan los ecosistemas. Digamos que aquí se guardan los modelos que utilizan la información existente en el subsistema anterior. En los últimos 20 años se ha producido un gran auge en el proceso de modelización. Algunos autores (Schuster y col. 1993) han identificado más de 250 aplicaciones informáticas útiles para simulación en el campo de la planificación forestal. Estos modelos van desde la

simulación del ciclo de nutrientes o del agua en un bosque, hasta la evaluación de las alteraciones producidas por el cambio climático. A pesar de este gran esfuerzo por cuantificar los procesos ecológicos y describirlos en modelos matemáticos, la mayoría de lo que sabemos todavía debe de ser expresado cualitativamente y de forma imprecisa. Frecuentemente este conocimiento cualitativo ha sido recopilado durante décadas de experiencia por una persona o grupo. Los avances teóricos y prácticos en el campo de la inteligencia artificial nos han permitido capturar sólo una pequeña parte de este conocimiento cualitativo, en forma de modelos informáticos complejos que se registran en este subsistema y que toman información del anterior.

El último subsistema que trataremos aquí es el que gestiona los **métodos de análisis de la decisión**. Este subsistema proporciona las herramientas necesarias para asistir en la elección de alternativas, análisis de sensibilidad del medio, identificación de variables implicadas en el proceso, etc. En definitiva, en este subsistema se analizan los resultados del proceso de modelización realizado en los anteriores, para facilitar la elección de una de las alternativas.

Para concluir este análisis sobre las características principales de los SAD, comentaremos brevemente cómo ha sido su desarrollo histórico. En general podemos decir que la evolución histórica de los SAD ha estado íntimamente ligada al desarrollo de los ordenadores personales y de los sistemas de gestión de información. De hecho, el concepto de sistema de asistencia a la toma de decisiones surge sólo cuando la tecnología informática y los procedimientos de gestión de información son lo suficientemente maduros como para plantearse este tipo de aproximaciones. Esto ocurre aproximadamente durante la década de 1980. El desarrollo de ordenadores cada vez más potentes y baratos, junto con la existencia de dispositivos para el almacenamiento masivo de datos, es una tendencia general y constante desde la invención del primer dispositivo electrónico. Sin embargo, el desarrollo de la teoría de la información ha sufrido una serie de hitos importantes que describimos a continuación.

En 1948, Claude E. Shannon publicó un artículo denominado "*A Mathematical Theory of Communication*". Nació así la teoría de la información como una respuesta al desarrollo tecnológico del momento. Comenzaba a generalizarse el uso de

dispositivos para almacenar y (sobre todo) transmitir información. Esto requería un fundamento teórico que permitiera el análisis y comprensión de este fenómeno. La propuesta de Shannon permitía modelizar el proceso de transmisión de información entre un emisor y un receptor basándose en el concepto termodinámico de entropía y asemejándolo a la "cantidad de información". Sin embargo, su teoría no contemplaba el aspecto del almacenamiento de la información. En esta época la información "digital" se almacenaba en tarjetas perforadas o bien en forma de conexiones físicas entre cables. Esto hace que no podamos hablar de información real, sino más bien de datos. La diferencia entre ambos conceptos estriba en que los datos se convierten en información cuando sufren un proceso de análisis y de estructuración determinado (Nebert 2000). La precariedad de los dispositivos de almacenamiento limitaba este proceso, por lo que sólo se guardaban datos. Sin embargo, fueron desarrollándose aspectos teóricos propios de los sistemas de información, como las bases conceptuales del procesamiento electrónico de los datos.

Paralelamente, en la década de los 50 surgió el concepto de inteligencia artificial, íntimamente ligado a los procesos de transmisión y almacenamiento de información. Este nuevo concepto fue acuñado en 1956 por John McCarthy, y tenía como objeto describir la posibilidad de crear dispositivos electrónicos capaces de desarrollar alguno de los procesos mentales humanos. Lo interesante de este concepto (al menos lo que nos interesa aquí), es que para crear inteligencia artificial es necesario comprender los procesos mentales de los humanos. Esto último está muy relacionado con los procesos de toma de decisiones y con los aspectos teóricos que tratan de simularlos. Por tanto, en la década de los 50 ya había un incipiente desarrollo conceptual que pretendía comprender y simular el proceso del pensamiento humano para aplicarlo a la resolución de problemas concretos.

Conforme aumentaba la capacidad de almacenar e integrar los datos digitales, fueron avanzando los aspectos prácticos de las teorías anteriores. Así, en la década de 1970 ya había dispositivos de almacenamiento masivo de información. Esto requería el desarrollo de procedimientos de gestión de dicha información para que mantuviera su coherencia. Así surgieron los sistemas de gestión de información,

cuyo desarrollo está relacionado con el concepto de base de datos. Los sistemas de gestión de bases de datos supusieron un gran avance en el desarrollo de todo lo relacionado con la teoría de la información. Mediante estos sistemas se puede mantener, actualizar, consultar y modelizar la información almacenada en un ordenador o en otro dispositivo. La potencia de estas herramientas hizo que se pudiera hablar de información procesada y no sólo de datos inconexos. También en la década de 1970 se desarrollaron los primeros sistemas expertos. Se trata de programas informáticos que tratan de simular el pensamiento humano mediante la representación simbólica del mismo. Este tipo de sistemas es el que se ha extendido hoy en los SAD y otros sistemas de información.

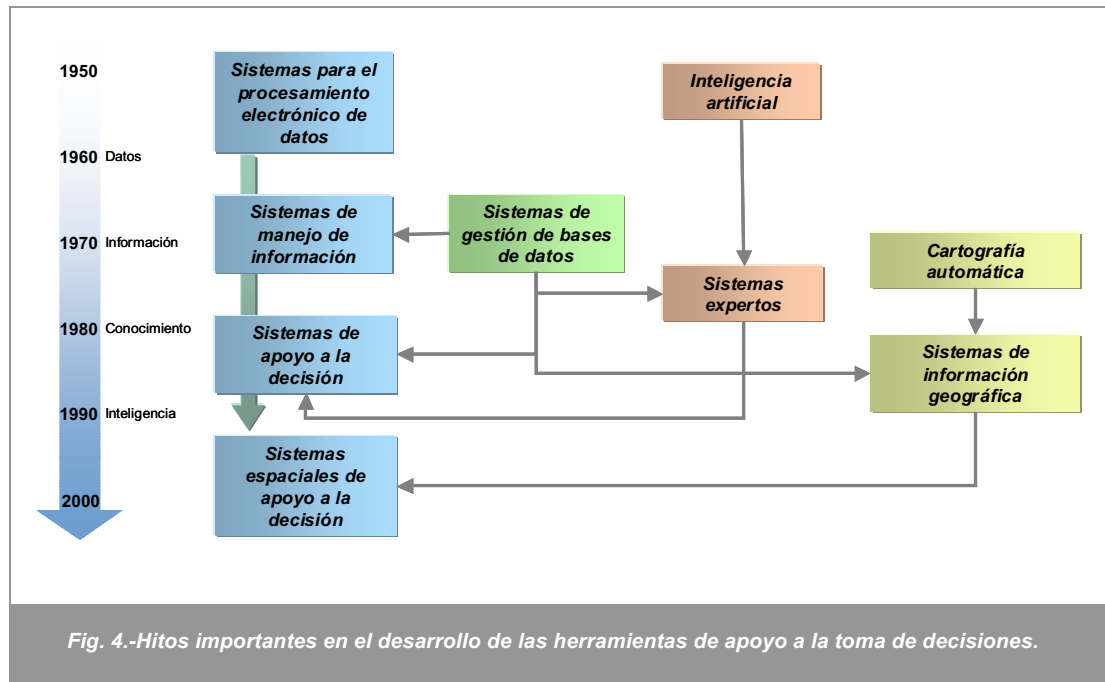
De forma paralela a lo anterior, fueron desarrollándose sistemas semiautomáticos para la producción de cartografía. Estos sistemas se perfeccionaron y evolucionaron de manera aislada a todo lo relacionado con la teoría de la información. Sólo en algunos casos muy puntuales se crearon verdaderos sistemas de información geográfica, donde se combinaban conocimientos procedentes de ambas disciplinas (teoría de la información y representación automática de cartografía).

En los años 80 se dio una confluencia de situaciones que permitieron el desarrollo de los primeros sistemas de apoyo a la toma de decisiones. En primer lugar, los procedimientos de almacenamiento y análisis de la información permitieron la aparición de un grado importante de sinergia. Surgía así el concepto de conocimiento almacenado en un sistema de información. El conocimiento se puede definir como la suma de la información en sí más la información extra debida a las relaciones que se establecen en un sistema de información. Pasamos por tanto de guardar información a almacenar conocimiento. Por otro lado, aparecieron los sistemas de bases de datos relacionales, que revolucionaron los procedimientos de gestión de información. Los antiguos sistemas automáticos de cartografía se aprovecharon del desarrollo de estos últimos sistemas, surgiendo así el concepto de Sistema de Información Geográfica. En los últimos años hemos asistido a un aumento casi exponencial en las funcionalidades de estos sistemas. De esta forma, podemos decir que hoy por hoy, los SIG son una herramienta casi insustituible a la hora de abordar cualquier estudio que implique el uso de cartografía básica o temática.

A principios de los 90 se pone de manifiesto la unión entre los conceptos procedentes de los sistemas expertos, los sistemas de gestión de bases de datos y los antiguos sistemas de manejo de información. Esto da lugar a los primeros sistemas de apoyo a la toma de decisiones, tal y como los conocemos en la actualidad.

También en la década de los 90 se produjo el último hito de interés en este resumen histórico. Después de muchos años de desarrollo independiente, los Sistemas de Información Geográfica comenzaron a incorporar funcionalidades propias de los Sistemas de Apoyo a la Decisión. La incorporación de algunas herramientas típicas de un SAD en ciertos SIG comerciales, ha permitido la "especialización" de los métodos de asistencia a la toma de decisiones (Heywood y col. 1994). De esta forma han surgido los denominados Sistemas espaciales de apoyo a la decisión (o SAD espaciales) (Gold 1993, Gold & André 1998, Riedl y col. 2000, Chen & Gold 2000). Mediante estos nuevos sistemas se pueden generar modelos que integran distintos aspectos del medio natural, dando respuesta a las necesidades científicas que tienen los gestores a la hora de tomar decisiones (Clare & Ray 2000). El uso de los SAD espaciales permite a técnicos o científicos procedentes de diversas disciplinas la participación en la elaboración de modelos cuyo resultado final es un "mapa de gestión" que permite al gestor tomar decisiones informadas sobre un aspecto determinado de sus competencias. Podemos definir aquí el "mapa de gestión" (Bonet García y col. 2001b) como el resultado de la integración de varias coberturas cartográficas temáticas referentes a distintos aspectos del medio (naturales, abióticos, sociales, económicos, políticos, etc.), mediante el uso de metodologías procedentes de la teoría de la decisión, con el objetivo final de obtener un documento que muestre distintas unidades homogéneas en lo que a su gestión se refiere. Asimismo, los SAD espaciales permiten realizar un análisis del tipo "¿qué pasaría si?" (del inglés "What if"). Esta función predictiva es una de las más interesantes de estos sistemas, ya que nos permite analizar los resultados modificando los posibles escenarios del problema. Como se comentará en el siguiente apartado, uno de los objetivos de este trabajo es mostrar las aplicaciones de estos sistemas espaciales de apoyo a la decisión, tomando como ejemplo la gestión forestal de un territorio concreto.

El siguiente esquema muestra el desarrollo histórico de los SAD comentado anteriormente.



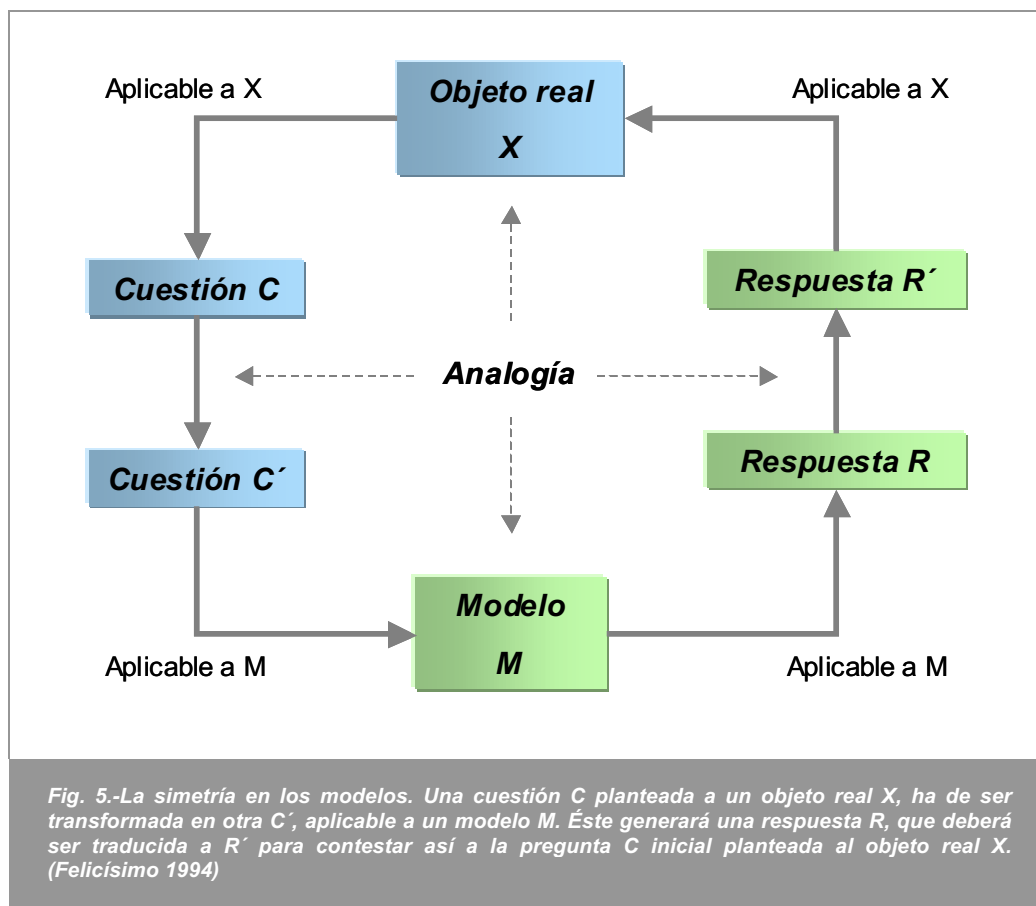
I.1.3 La modelización ambiental.

Como vimos anteriormente, la resolución de un problema mediante herramientas de apoyo a la decisión, suele requerir la elaboración de un modelo más o menos complejo para simular el proceso en cuestión. Por ello la teoría de modelos constituye un aspecto básico a la hora de comprender las herramientas de apoyo a la decisión y de aplicarlas en un caso concreto. En este apartado describiremos en primer lugar el concepto de modelo, para luego analizar alguno de sus tipos, utilidades y marco de trabajo general.

Al igual que ocurre con otras palabras de uso general, resulta difícil dar una definición de modelo. Si buscamos en un diccionario, veremos que una de las definiciones podría ser: "representación en pequeño de alguna cosa", o "representación simplificada de la realidad en la que aparecen algunas de sus propiedades". Sin embargo, resulta complejo dar una definición más formal. Lo

cierto es que estamos más o menos familiarizados con el concepto de modelo como simplificación de la realidad realizada para explicar algo más complejo (Voinov 1999). Los modelos pueden ser maquetas físicas, esquemas mentales, procesos físicos, etc. Pueden ser sencillos o tremendamente complejos. También pueden tener carácter explicativo, o bien predictivo. Pueden ser cualitativos o cuantitativos, basados en la teoría o en hechos empíricos. Se pueden representar en forma de pensamiento, gráfico, tabla, texto, ecuaciones matemáticas o programas informáticos (Rogers & Johnson 1998).

A pesar de la gran diversidad de modelos existentes sobre casi cualquier aspecto, todos ellos surgen de una realidad que ya hemos comentado: Resulta imposible conocer la totalidad de los aspectos que constituyen los sistemas complejos (Margalef 1991). Esto nos pone ante un dilema interesante: un modelo funciona tanto mejor cuanto más se acerca a la realidad. Pero esto lo hace cada vez menos útil y manejable. Se trata por tanto de alcanzar un equilibrio entre complejidad y operatividad. En general los modelos que describen los ecosistemas o sus componentes han sido aceptados como simplificaciones de la realidad necesarios para predecir y/o comprender su funcionamiento (Rasinmäki 2002). El uso de modelos implica asumir que tienen ciertas limitaciones y que sus resultados deben de ser analizados según los objetivos del modelo y según la información utilizada para su puesta en marcha (Rogers & Johnson 1998). Para que los modelos puedan ser realmente útiles, es necesario que se elaboren estableciendo una relación con la realidad que debe de ser simétrica. Es decir, la relación de correspondencia entre el objeto real y el modelo debe ser al menos parcialmente reversible y debe permitir la traducción de algunas propiedades del modelo a la realidad (Felicísimo 1994). La siguiente figura muestra con más detalle esta característica de los modelos.



Podemos plantearnos si realmente merece la pena hacer un esfuerzo de síntesis para generar un modelo que simule un proceso determinado. Aunque esta pregunta no tiene respuesta fácil y depende de las distintas situaciones particulares, podemos esbozar las ventajas e inconvenientes de la modelización a la hora de resolver problemas concretos (Varcoe 1990).

- **Ventajas:** podemos citar las siguientes.
 - Permiten la integración de múltiples variables complejas en un marco definido y comparable con otros. La complejidad del modelo depende del grado de comprensión que se tenga del fenómeno modelizado.
 - El almacenamiento de la información está estandarizado, lo que facilita el análisis y consulta.
 - La creación de modelos requiere la formación de grupos interdisciplinarios que trabajen coordinadamente. En este sentido los modelos son un buen “foro” para trabajar en equipo.

- La simulación es un proceso rápido y barato de investigación. Esto es especialmente cierto cuando los requerimientos económicos no permiten una fase intensa de captura de información para resolver un problema determinado.
- Gracias a los modelos se puede simular el transcurso de grandes periodos de tiempo en cuestión de segundos o minutos.
- Los modelos pueden ser permanentemente actualizados y mejorados para reflejar tanto las novedades técnicas como los avances en el conocimiento del proceso modelizado.
- La creación de un modelo casi siempre acaba mejorando la comprensión del proceso modelado por parte de los autores. No sólo se trata de resolver un problema, sino de aprender del proceso subyacente al anterior.
- La simulación permite la extrapolación de resultados experimentales a lugares y a momentos temporales diferentes de los utilizados para generar el modelo.
- Los modelos permiten un análisis del tipo "¿qué pasaría sí?", lo que nos permite modificar ciertos parámetros del sistema modelizado para evaluar los resultados en los distintos casos. Así surge el concepto de escenario de un modelo como el conjunto de parámetros que se tienen en cuenta en su elaboración y que pueden ser modificados por los autores.
- **Desventajas:** entre otras tendríamos las siguientes.
 - Uno de los principales problemas es la interpretación de los resultados. Éstos han de ser analizados críticamente, teniendo en cuenta el proceso de modelización y la información tomada para ejecutarlo. El hecho de que los resultados salgan de un ordenador, no indican que sean inherentemente correctos.
 - Todos los modelos requieren una validación, con objeto de aplicar correctamente sus resultados. El problema es que no hay una metodología estándar para realizar esta comprobación. Además, hay casos en los que no es posible llevarla a cabo por la propia definición del modelo.
 - En muchas ocasiones, los modelos requieren gran cantidad de datos difíciles de conseguir.
 - También hay casos en los que el desmesurado interés por realizar modelos resta importancia a la fase de captura de datos en campo. Su versatilidad

puede llevar a esta situación porque es relativamente fácil hacer un modelo sobre cualquier aspecto del medio. Otra cosa es que se ajuste a la realidad con ciertas garantías. Algunos autores suelen decir en este sentido (de modo irónico), que los modelos se usan frecuentemente para confirmar lo que ya se sospechaba (Burrough y col. 1996).

- Otra desventaja consiste en exceder los límites de aplicación del modelo. A veces se extrapolan sus resultados más allá de lo que fue pretendido inicialmente. Esta extrapolación excesiva se aplica tanto en el aspecto espacial (aplicar un modelo diseñado en ambientes tropicales para gestionar la vegetación aquí, por ejemplo), como en el temporal.
- Por último, los generadores de modelos suelen caer en el error de que su modelo es el mejor. Cuando uno hace un modelo casi siempre se ajusta perfectamente a la realidad según el autor. Lo que ocurre en realidad es que el modelo se ajusta perfectamente a la percepción de la realidad que tiene su autor, y no necesariamente al mundo real en sí. Por esto se produce esta falsa (a veces) sensación de éxito.

El frecuente uso de modelos en casi todas las disciplinas científicas ha ocasionado que haya múltiples formas de clasificarlos. Esbozaremos alguna de ellas.

En primer lugar podemos distinguir varios tipos en función del objetivo que persiguen (Riviera 1999):

- **Modelos descriptivos:** Tienen como objeto la observación de procesos de un sistema, así como su descripción verbal o mediante otros medios. Sólo importa la descripción del proceso, no su funcionamiento interno.
- **Modelos explicativos:** En este caso están orientados a identificar los elementos y procesos de un sistema, haciendo especial hincapié en la determinación del esquema causal que subyace al proceso. Es decir, pretenden desentrañar las reglas que rigen el funcionamiento de un sistema determinado.
- **Modelos de decisión:** Se trata de modelos cuyo objetivo es asistir a la toma de decisiones en condiciones de incertidumbre. Este tipo de modelo es el que más nos interesa en este caso.

Desde el punto de vista de los métodos utilizados para la construcción de los modelos, podemos distinguir también varios tipos (Dale & English 1999):

- **Modelos heurísticos:** son modelos basados en conceptos mentales y son relativamente sencillos. Muchos autores los consideran poco rigurosos, ya que se suele recurrir a la experiencia personal para formularlos. Pueden ser expresados en forma de gráficos, esquemas o tablas. Su interés radica en que suelen ser sencillos de explicar y de entender. En algunas ocasiones se usan modelos heurísticos como base conceptual para generar otros más complejos de los tipos siguientes. A pesar de su sencillez, estos modelos están empezando a jugar un papel importante en la disciplina de la modelización. Esto se debe a que cada vez hay más procedimientos que permiten plasmar en un documento la experiencia de una persona o personas. La experiencia personal suele tener mucho que ver con la heurística.
- **Modelos físicos:** Se suele decir que estos modelos son la realidad a una escala reducida. Generalmente los modelos físicos son representaciones a escala del proceso o fenómeno analizado. Entre ellos podemos citar los túneles de viento, u otros más complejos utilizados para simular las mareas o los procesos de inundación. La ventaja de estos modelos radica en que no es necesario conocer mucho sobre el proceso a simular. Basta representarlo lo más fielmente posible a una escala asequible. Por ejemplo, para saber el grado de erosionabilidad de una serie de suelos, podríamos extraer sus perfiles y simular la lluvia en el laboratorio. Sin embargo, no siempre es tan sencillo. Simular el efecto del cambio climático en un bosque requeriría reproducirlo en un ambiente controlado. Y esto requiere a su vez un profundo conocimiento de la dinámica del ecosistema. En general estos modelos no suelen aplicarse en los ámbitos en los que se desarrolla este trabajo, por lo que no abundaremos en ellos.
- **Modelos matemáticos:** la característica fundamental de estos modelos es que en ellos se describen las relaciones entre los elementos mediante ecuaciones matemáticas. Los modelos demográficos derivados de las leyes de Lotka y Volterra son un buen ejemplo de este tipo. El hecho de que aparezcan ecuaciones matemáticas les da cierto cariz de fiabilidad y corrección. Sin embargo, no siempre es cierto. Resulta difícil definir las relaciones entre los elementos de un ecosistema mediante ecuaciones matemáticas. Normalmente la naturaleza no es tan rígida como las ecuaciones, por lo que no siempre se ajusta a las mismas. De todas maneras los últimos desarrollos de las ciencias matemáticas están ayudando a flexibilizar este tipo de ecuaciones. Nos

referimos a la lógica difusa (Ekel 1999,Klir 2000,Tyson 2001) y a otras aproximaciones como las matemáticas de los sistemas complejos (Pimm 1991,Solé 1998,Manson 2001). En definitiva, los modelos matemáticos se basan en el concepto de algoritmo. Un algoritmo es un método de resolución de problemas complejos mediante la repetición de operaciones matemáticas sencillas. En los algoritmos se “comprime” la realidad que se trata de representar en el modelo. En función de los distintos tipos de algoritmos utilizados, podemos definir varios tipos de modelos matemáticos (Burrough y col. 1996).

- **Basados en reglas** (modelos lógicos): este tipo de modelos está muy extendido en la actualidad. Se basan en la formulación de una serie de reglas axiomáticas consideradas como válidas antes de ejecutar el modelo (Loh y col. 1998). Los algoritmos matemáticos utilizados se centran en la representación de cada una de estas reglas y en las relaciones entre las mismas. Este tipo de modelos están muy relacionados con los modelos heurísticos, ya que permiten expresar mediante ecuaciones matemáticas aspectos relacionados con la experiencia personal de una persona o grupo. En este trabajo utilizaremos modelos semejantes a este tipo, que serán descritos con más detalle en el apartado III.2.1.2.
- **Modelos empíricos**: Se trata probablemente de los modelos más complejos, ya que se usan el empirismo como base teórica. Esto implica que el proceso de modelización está íntimamente relacionado con los resultados obtenidos tras un análisis experimental del problema. Estos modelos están próximos al mecanicismo, ya que asumen que la comprensión de cada una de las partes del sistema implica una comprensión de la globalidad. Esta comprensión se lleva a cabo mediante la experiencia propia de la filosofía newtoniana. En estos casos, la estructura del modelo no es importante. Lo interesante es que las ecuaciones matemáticas que transforman las entradas en salidas, funcionen correctamente. Se les suele denominar “modelos tipo caja negra” (Voinov 1999), ya que operan como dispositivos cerrados en lo que se refiere al flujo de la información a su través. Los problemas derivados de la gestión del medio natural son especialmente difíciles de abordar bajo esta perspectiva, ya que se requiere un amplio y detallado conocimiento del medio. Como consecuencia de este tipo de modelos se suele obtener una

ecuación que permite relacionar una serie de variables conocidas para obtener el valor de otra desconocida de interés. Se han construido múltiples modelos de este tipo que tratan de simular procesos como la fisiología de una planta, o el clima planetario, etc.

- **Modelos físicos deterministas:** Si los anteriores modelos se basaban en la obtención de datos mediante experimentos, éstos asumen que se conoce el funcionamiento exacto de todo el sistema. Son modelos mecanicistas puros en los que el concepto de sistema no adquiere demasiada importancia. Hay pocos campos en los que pueden aplicarse en la realidad. Estos modelos resultan difícilmente implementables en un SIG. Esto se debe a que estos modelos se centran en simular la dinámica de un individuo aislado (una planta, o animal o bosque), por lo que no tiene demasiado interés una espacialización de la información.
- **Modelos físicos estocásticos:** Aunque confían en las leyes físicas y se basan en la experiencia, estos modelos incorporan un factor probabilístico que los hace más versátiles. La incorporación de la estocástica (azar o probabilidad) implica que en estos modelos hay factores desconocidos de los que sólo sabemos que hay cierta probabilidad de ocurrencia. Por ello son más interesantes que los anteriores en lo que a aplicación a aspectos ambientales se refiere. Al igual que los anteriores, resulta difícil implementar en un SIG estos tipos de modelos.

Por último distinguiremos dos tipos de modelos en función de su relación con los sistemas de información geográfica:

- **Modelos individuales:** Se trata de modelos que simulan el funcionamiento de una entidad individual no representable espacialmente. Se considera que el objeto a modelizar está más o menos aislado y por ello no se tienen en cuenta los parámetros espaciales (vecindad, topología, etc.). Como ejemplo de estos modelos podríamos citar aquellos que simulan la fisiología de plantas aisladas (Rodà y col. 1999), o los que simulan la dinámica poblacional de una especie mediante gráficas o tablas de abundancia.
- **Modelos espaciales:** Estos modelos son los que más nos interesan en este trabajo. Se caracterizan porque los procesos que simulan no afectan sólo a un individuo, sino a un conjunto de ellos que se encuentran relacionados

espacialmente. Los modelos espaciales suelen utilizar el modelo de datos **raster** para su representación gráfica. En este modelo de datos se asemeja una celdilla (casi siempre cuadrada), a cada uno de los elementos individuales simulados por los modelos anteriores. Pero no se trata sólo de unir individuos en una capa de información. También hay que tener en cuenta las relaciones de vecindad que pueden establecerse entre los anteriores (Slothower y col. 2001). En los últimos años estamos asistiendo a un progresivo proceso de espacialización de los distintos modelos individuales (Kemp 1997, Roy & Tomar 2000, Slothower y col. 2001). De hecho, algunas disciplinas relativamente recientes como la ecología del paisaje (Turner 1998), basan buena parte de sus trabajos en la simulación mediante modelos espaciales. Esto se debe a la extensión en el uso de los SIG y a su gran potencia de análisis (Fall & Fall 1999, Hall 2000, Logsdon 2001).

Al igual que planteamos con los SAD, hay un marco operativo que facilita el proceso de modelización. Se trata de identificar los pasos necesarios para generar un modelo. Dentro del esquema conceptual de los SAD (Figura 1), el proceso de modelización sería similar al mostrado en dicha figura, pero empezaría en los puntos 3 y 4 (Caracterización del medio e integración de la información). Los distintos pasos que comentamos a continuación son muy similares a los descritos para los SAD.

Elaborar un modelo es un proceso iterativo (Voinov 1999), lo que indica que requiere la consecución de una serie de pasos que se repiten un número determinado de veces. Comenzamos planteando el objetivo general del modelo (no del problema decisional en general, según la figura 1). A continuación identificaremos aquellas propiedades del sistema que han de ser consideradas para conseguir el objetivo anterior. Esto puede requerir la recopilación de información ya elaborada, generación *de novo*, o consulta a expertos. Una vez hecho esto, y con el objetivo en mente, comenzamos a caracterizar el sistema a modelizar según tres dimensiones básicas (Voinov 2002): **espacial, temporal y estructural**.

Desde el punto de vista **espacial**, es básico identificar el tamaño del objeto que vamos a analizar. Así sabremos la escala aproximada de nuestro modelo, su dinamismo, relaciones con otros objetos, etc. La escala de los modelos espaciales

es un aspecto clave, ya que nos permitirá saber el ámbito espacial en el cual se pueden aplicar los resultados obtenidos (Lovell y *col.* 2002, Storch 2002). No tiene sentido, por ejemplo aplicar un modelo generado a escala de un país, a un parque natural determinado. El **tiempo** es otro factor clave, ya que su incorporación en un modelo espacial complica considerablemente el proceso (Rasinmäki 2002). Añadir el factor tiempo a un modelo espacial no consiste sólo en volver a ejecutar el modelo en un momento temporal diferente. El factor tiempo es una dimensión más que hay que incorporar al modelo, al igual que lo es la altura o la latitud (Yuan 2001). Este aspecto es uno de los más complejos dentro de la modelización con SIG, y todavía no está resuelto en su totalidad (Yuan 2001). Asumir el tiempo en los modelos implica responder a preguntas como: ¿Todos los procesos discurren a la misma velocidad?, ¿necesitamos saber cómo evoluciona el proceso en el tiempo, como en un video, o nos basta con una foto fija de la realidad?, etc. La **estructura** del modelo debe de ser claramente descrita antes de su elaboración. Necesitamos saber qué elementos constituyen el sistema, cómo se relacionan y en qué grado de detalle se van a representar. Un aspecto fundamental, muchas veces minusvalorado, es el conocimiento de los factores limitantes. En la naturaleza hay pocas reglas fijas (Margalef 1991), pero una de ellas es que los factores limitantes son los que condicionan casi siempre el funcionamiento de los sistemas. Por tanto, resulta básico identificar aquellos factores que puedan limitar el desarrollo del proceso que estamos modelizando.

Una vez identificados los aspectos espaciales, temporales y estructurales, procedemos a elaborar un **modelo conceptual** del sistema considerado. Un modelo conceptual es el esbozo inicial del modelo. De su correcta realización depende buena parte del éxito. Puede representarse de forma mental, en un esquema o dibujo, gráfico, diagrama de flujo, etc. Para su expresión no se recurre a un lenguaje concreto y definido. El modelo conceptual debe de identificar claramente los siguientes aspectos del sistema:

Los **límites** que permiten definir el sistema que estamos modelizando. Son importantes para decidir qué material e información sale y entra del sistema, qué procesos son externos y cuáles internos. Las **variables** son los factores que pueden cambiar y que caracterizan el modelo. Sin embargo, los **parámetros** son estáticos, aunque también permiten caracterizar el modelo.

La elaboración de un modelo conceptual es a veces un ejercicio muy subjetivo, en el que el autor o autores deben de ser capaces de plasmar toda su experiencia sobre el aspecto a modelizar. A veces se considera algo más parecido al arte que a la ciencia, ya que no hay guías claras para describir cómo se hace. Puede haber recomendaciones y sugerencias, pero cualquiera puede hacerlo a su manera. Una forma de evaluar de alguna manera el modelo conceptual es tratar de adaptarlo a los términos matemáticos utilizados para describir los modelos. La formalización consiste en sustituir los conceptos, imágenes o palabras del modelo conceptual, por ecuaciones matemáticas que permiten la implementación real del modelo. Una vez pasado este proceso, tenemos un modelo cuyo comportamiento puede ser comparado con el objeto real simulado. Al ejecutarlo obtenemos una serie de resultados que satisfarán en mayor o menor medida los objetivos planteados al principio. Llega el momento de evaluar estos resultados. Como ya hemos comentado resulta muy difícil evaluar los resultados de un modelo determinado. En ocasiones no es posible comparar los resultados del modelo con la realidad, bien por la dificultad técnica que esto implica, o bien porque el modelo incorpora aspectos espaciales o temporales no constatables. Por ello no hay una metodología clara a la hora de evaluar los resultados de un modelo.

Una vez que el modelo ha superado la fase de evaluación, puede ponerse en uso. Esto no implica que el modelo ya no volverá a evaluarse. Los resultados reales se obtienen una vez que el modelo está en explotación. Esta experiencia es básica para aprender los posibles errores cometidos en las distintas fases del proceso de modelización. Además la propia construcción de un modelo permite aumentar el conocimiento del proceso modelizado.

El siguiente esquema muestra de forma resumida las distintas etapas en la construcción de un modelo.

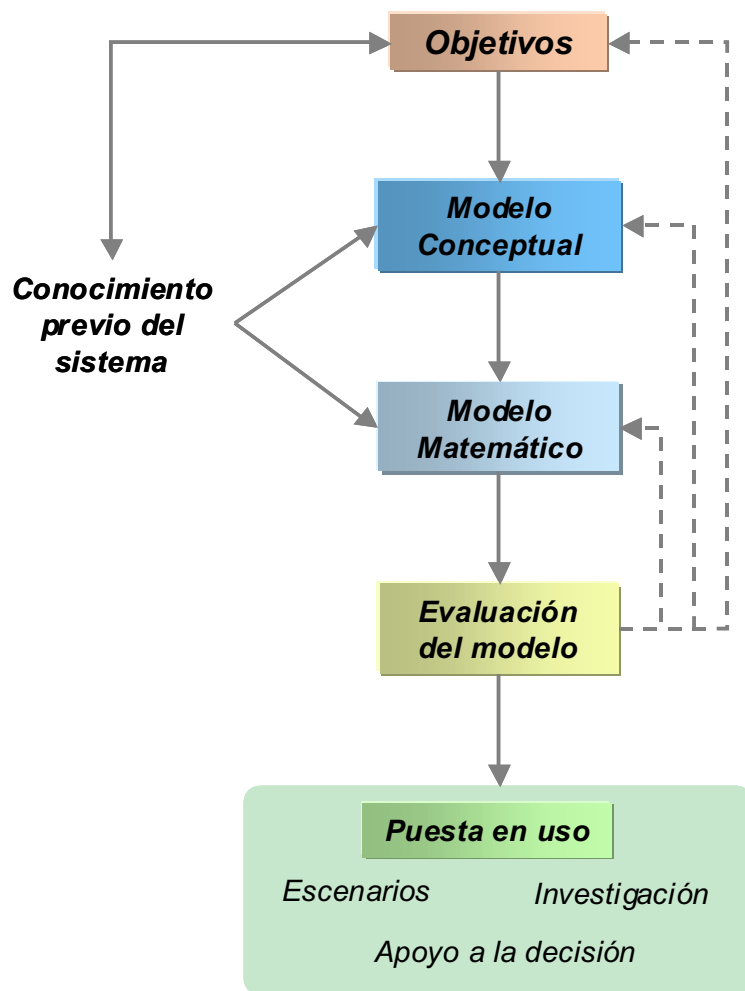


Fig. 6.-Fases principales que facilitan la elaboración de modelos (Voinov 1999)

1.1.3.1.1 Métodos de integración de información: el interior de un SAD.

Hasta ahora hemos descrito las principales funciones de los SAD, así como su estructura interna y características básicas. En este apartado trataremos con detalle la metodología que utilizan estos sistemas para integrar la información y obtener lo que ya hemos denominado un “mapa de gestión” (Bonet García y col. 2001a).

El concepto de integración de información es relativamente ambiguo y hace referencia a un proceso mental típicamente humano. Por eso quizás, resulta difícil de explicar. Sin embargo, los SIG ayudan en este sentido, ya que en ellos la “integración de información” tiene una importante vertiente gráfica, lo cual facilita el proceso. Se trata, en definitiva, de generar una información cartográfica sintética que resuelva un problema decisional determinado. Este problema se encuentra implementado en un modelo lógico de un SAD y ahora procedemos a resolverlo. Por ejemplo, supongamos que queremos saber la aptitud de un territorio determinado para albergar asentamientos apícolas. Necesitamos saber dónde es más adecuado ubicar una serie de colmenas con objeto de maximizar la producción. Para ello procedemos tal y como ya describimos en la figura 1: formulamos el problema, planteamos unos objetivos y mediante un SAD elaboramos un modelo conceptual. Supongamos que es tan sencillo que sólo considera tres parámetros: vegetación, cercanía a vías de acceso y presencia de agua. De esta forma, una zona apta para la implantación de colmenas deberá reunir las siguientes características:

Será un sitio con vegetación apta para la producción de miel (romerales, tomillares, etc.). Además debe de ser un lugar al que se pueda acceder en vehículo a motor, con objeto de manejar la explotación y retirar la cosecha. Por último, los lugares con presencia constante de agua son más adecuados, ya que las colmenas necesitan beber con abundancia.

Pues bien, este modelo conceptual puede ser resuelto de múltiples maneras en función de las distintos métodos de integración de información disponibles en la actualidad. De esta forma, el proceso de integración debe de reflejar la aportación de cada una de las variables anteriores (independientes) a la variable dependiente.

Esta variable dependiente o resultado contiene un valor de **idoneidad** o **aptitud** de cada punto del territorio para el objetivo del análisis. Los procedimientos matemáticos que hacen posible esta relación entre las variables independientes y la dependiente, se denominan **reglas de decisión** (Suárez López 2001).

En los siguientes párrafos describiremos las principales metodologías disponibles para resolver los modelos conceptuales implementados en un SAD espacial. Para ello esbozaremos cómo se resolvería el ejemplo anterior con cada uno de los métodos descritos.

En principio podemos dividir los métodos de integración de mapas (o regla de decisión) en 6 grupos (ver tabla 3).

- **Reglas basadas en operadores de lógica discreta** (Heywood y col. 1994, Eastman 1997, Husdal 2001, Eastman 2001): se trata de la forma de combinación más sencilla. Los factores se representan mediante variables binarias, lo que se traduce en un resultado también binario (Bosque y col. 2000). Esto indica que la leyenda de cada uno de los mapas que representa a los factores implicados, tiene sólo dos valores: **0**, que indica incompatibilidad total, y **1**, que indica compatibilidad. Estos factores binarios se relacionan mediante una serie de operadores que se traducen en operaciones matemáticas entre las distintas capas de información. En este caso los operadores son: **Y** (AND) y **O** (OR). Al unir dos capas mediante el operador **Y**, estamos diciendo que las zonas aptas en el resultado son aquellas que también lo son en los dos factores considerados. El operador **O** se comporta de manera opuesta: sólo es necesario que una de las capas de entrada sea apta para que el resultado final también lo sea. Si aplicamos al ejemplo anterior este tipo de operadores, tendríamos lo siguiente: “una zona buena para albergar colmenas debe de tener unidades de vegetación melífera **Y** disponibilidad de agua **O** estar cerca de un carril transitable”.

Como se puede observar este tipo de operadores da un resultado demasiado rígido, ya que simplifica la realidad a zonas totalmente aptas y absolutamente inhábiles. En realidad este método de combinación sólo se usa cuando se trabaja

con factores realmente limitantes, como incompatibilidades debidas a usos del suelo prohibidos, o situaciones similares.

- **Reglas basadas en operadores de lógica difusa** (Zadeh 1965, Reynolds 1998b, Ekel 1999, Klir 2000, Aranda y col. 2000, Sainz de la Maza 2000): esta forma de combinación se basa también en operadores similares a los anteriores. Sin embargo, hay algunas diferencias. En primer lugar los factores no se representan de manera binaria, sino que pueden tomar múltiples valores continuos. Además los operadores que combinan los factores no son tan rígidos como los anteriores. No abundaremos más aquí en este tipo de reglas decisionales ya que serán descritas con más detalle en el siguiente apartado.
- **Reglas basadas en sumas ponderadas** (Heywood y col. 1994, Eastman 1997, Husdal 2001, Eastman 2001, Ramos & Mendes 2001): esta aproximación se basa en la existencia de una serie de pesos específicos para cada uno de los factores combinados. De esta manera se puede matizar la combinación en función de la información que tengamos sobre el funcionamiento del sistema. Se distinguen dos tipos de sumas ponderadas:
 - Factores binarios (Eastman y col. 1993): En este caso los factores a combinar son binarios y cada uno de ellos está afectado por un peso específico. Se suele convenir que la suma de los pesos asignados a todos los factores debe de ser uno (para evitar posteriores reclasificaciones). Al trabajar con factores binarios, obtendríamos un resultado también de este tipo. En el ejemplo del mapa apícola podríamos tener algo así: "el factor vegetación es 2 veces más importante que el resto de los factores". Aplicando los factores binarios tendríamos que la aptitud global para un punto determinado, sería:

$$S(F_1, F_2, F_3) = (F_1 \cdot W_1) + (F_2 \cdot W_2) + (F_3 \cdot W_3) = (1 \cdot 0.5) + (0 \cdot 0.25) + (0 \cdot 0.25) = 0.5,$$

Siendo W_i los pesos de los factores (el primero dos veces mayor que los demás) y F_i , el valor que adquieren los distintos factores en cada punto (valores binarios).

- Factores no binarios (Eastman y col. 1993): Ahora los factores no son binarios (pueden tomar cualquier valor dentro de una escala determinada), lo cual da gran versatilidad al método. Es decir, no sólo hay un peso para cada factor, sino que se define un peso específico para cada una de las clases definidas en cada factor. La ecuación sería similar a la anterior, variando únicamente el rango que toman los valores dentro de cada factor. Esta metodología ha sido ampliamente desarrollada por el SIG raster **Idrisi** (Eastman y col. 1993, Heywood y col. 1994, Eastman 1997, Eastman 2001), en el cual se han implementado procesos más complejos que el presente ejemplo. En general las sumas ponderadas tienen dos desventajas. Por un lado, la existencia de pesos aporta cierta carga de subjetividad al proceso. Aunque hay métodos que intentan sistematizar en la medida de lo posible esta asignación (Saaty 1977), éste suele ser uno de los aspectos más criticados del método. Por otro lado, la naturaleza aditiva de la combinación produce efectos de compensación de factores. Es decir, puede que en un lugar donde el factor más importante tiene poca aptitud, la gran aptitud de los otros factores compense esta situación.
- **Reglas basadas en probabilidad bayesiana** (Stassopoulou y col. 1996, Kangas y col. 2000): esta metodología se basa en el concepto de probabilidad de ocurrencia del suceso que estamos modelizando. Más concretamente se trata de evaluar la probabilidad de ocurrencia del suceso problema, condicionado a la ocurrencia de otros sucesos conocidos (que actúan como variables independientes). Aplicando esta metodología al ejemplo utilizado, diríamos: la aptitud del territorio para albergar colmenas se puede expresar en términos de probabilidad. En ese caso, la probabilidad de que un lugar sea apto para poner colmenas, dependerá de la probabilidad de que esto ocurra sabiendo el tipo de vegetación en cada caso, y la probabilidad de disponibilidad de agua y de estar cerca de un carril de acceso. Conocemos para cada punto las probabilidades de ocurrencia de los sucesos representados por los factores anteriores. Gracias a esto podemos calcular la probabilidad de ocurrencia del suceso problema, mediante las probabilidades condicionadas.

En general estos métodos de integración de información se utilizan en la generación de modelos de distribución potencial. Sabemos el patrón de distribución de un fenómeno (especie biológica, yacimiento minero, etc.) y queremos extrapolar este conocimiento al resto del territorio. Estos métodos no son demasiado adecuados para generar modelos de aptitud como el del ejemplo. En principio podemos distinguir dos tipos de aproximaciones: perfiles corregidos y perfiles ponderados. Aunque no los describiremos, sí merece la pena apuntar que los primeros se usan en la elaboración de modelos de distribución potencial de especies (Austin 2002), mientras que los segundos tienen una clara utilidad en la identificación de potenciales yacimientos mineros o arqueológicos (Raines y col. 2000, Sahoo y col. 2000).

- **Regresión logística** (Felicísimo y col. 2000, Palomares Fernández 1999, Mladenoff & Baker 1999, García Santiago y col. 2001, Hirzel 2001, Fernández Requena 2003) al igual que el método anterior, mediante la regresión logística se trata de establecer la probabilidad de ocurrencia de un suceso determinado. Sin embargo, en este caso la relación entre las variables independientes y la dependiente se hace a través de una ecuación logística, del tipo:

$$P(i) = \frac{1}{1 + e^{a \cdot F_1 + b \cdot F_2 + c \cdot F_3}}$$

, donde $P(i)$ es la probabilidad de ocurrencia del fenómeno problema (en nuestro caso sería probabilidad de que el punto en cuestión sea apto para la producción de miel). Las constantes a , b , c , etc. son factores asociados a los distintos parámetros considerados (F_1 , F_2 y F_3). Estas constantes se obtienen mediante una fase de campo en la que se determina el patrón de distribución del fenómeno problema. En nuestro caso, se trataría de identificar qué factores determinan la aptitud del territorio para albergar colmenas. Mediante la regresión logística se integran todos los datos de campo y se obtienen las constantes comentadas anteriormente.

La regresión logística se aplica a la obtención de modelos de distribución potencial de especies, análisis de riesgos de inundaciones (Conselleria d'Obres Públiques 2001, Jones & Thornton 2002), aludes, etc.

- **Redes neuronales** (de la Rosa y col. 1999, Deadman & Gimblett 2002, Thurston 2002). Esta última clase de metodologías es la más interesante desde el punto de vista teórico, ya que pretende simular el modo en el que nuestro cerebro afronta un problema determinado para resolverlo. Las redes neuronales consisten en una estructura en red: un conjunto de nodos unidos entre sí por enlaces. Los nodos se encuentran jerarquizados en una serie de capas. A cada nodo le llega información procedente de todos los nodos de la capa anterior. La procesa y la envía a todos los nodos de la capa siguiente. El paso de una capa a otra lleva asociado la asignación de un peso determinado. Además cada nodo aplica una función determinada a la información que pasa a su través. Se trata de una estructura un poco "desordenada" que simula de alguna manera la forma en la que se relacionan las neuronas de un cerebro. Cuantos más nodos hay en una red, más compleja resulta y mejor resuelve los problemas. Actualmente hay múltiples redes neuronales que resuelven problemas complejos y que asisten a la toma de decisiones, sobre todo en aspectos médicos e ingenieriles. Todavía hay pocas aplicaciones al campo del medio ambiente (Benito Garzón 2002). El principal problema de su aplicación es que estos modelos carecen de significado intrínseco. Es decir, pueden replicar los datos de una muestra y obtener un resultado coherente, pero no se pueden interpretar y no ayudan a comprender el funcionamiento del proceso. Se trata, en definitiva, de modelos de tipo "caja negra".

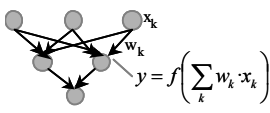
| Metodología | | Expresión | Características |
|------------------------|----------------------|---|--|
| Lógica discreta | AND | $AND(F_1, F_2, F_3) = \begin{cases} 1 & \text{si } F_1, F_2 \text{ y } F_3 = 1 \\ 0 & \text{si } F_1, F_2 \text{ ó } F_3 = 0 \end{cases}$ | <ul style="list-style-type: none"> Los factores tienen carácter binario, adquiriendo valores de 0 ó 1 únicamente. Se usan para distinguir zonas compatibles e incompatibles. Todos los factores tienen la misma importancia. Fáciles y rápidos de aplicar. Son muy rígidos (resultados binarios). |
| | OR | $OR(F_1, F_2, F_3) = \begin{cases} 1 & \text{si } F_1, F_2 \text{ ó } F_3 = 1 \\ 0 & \text{si } F_1, F_2 \text{ y } F_3 = 0 \end{cases}$ | |
| Lógica difusa | FUZZY AND | $fuzAND(F_1, F_2, F_3) = \min(F_1, F_2, F_3)$ | <ul style="list-style-type: none"> Son tan intuitivos como los operadores discretos, pero más potentes. El producto y la suma algebraica producen efectos antagonista y sinérgico respectivamente al "cruzar" la información. Permiten trabajar con bases de conocimiento. Se pueden relacionar con programas de modelización lineal (Stella, VENSIM, etc.). Aprovechan las ventajas de la lógica difusa. No hay control sobre la asunción de riesgo y compensación. No se establecen pesos entre los factores. |
| | FUZZY OR | $fuzOR(F_1, F_2, F_3) = \max(F_1, F_2, F_3)$ | |
| | Producto algebraico | $fuzPRO(F_1, F_2, F_3) = (F_1 \cdot F_2 \cdot F_3)$ | |
| | Suma algebraica | $fuzSUM(F_1, F_2, F_3) = 1 - [(1-F_1) \cdot (1-F_2) \cdot (1-F_3)]$ | |
| | GAMMA | $fuzGAMMA(F_1, F_2, F_3) = [fuzSUM(F_1, F_2, F_3)]^\gamma \cdot [fuzPRO(F_1, F_2, F_3)]^{1-\gamma}$ | |
| Sumas ponderadas | Factores binarios | $S(F_1, F_2, F_3) = (F_1 \cdot W_1) + (F_2 \cdot W_2) + (F_3 \cdot W_3)$ <p>, para $F_i = 0 \text{ ó } 1$, y $W_i \in [0, 1]$</p> | <ul style="list-style-type: none"> Modelos de idoneidad más precisos que los basados en operadores discretos. Fáciles de aplicar. Se usan pesos para ponderar los factores. Naturaleza aditiva de la combinación. Los factores son binarios. |
| | Factores no binarios | $S(F_1, F_2, F_3) = (F_1 \cdot W_1) + (F_2 \cdot W_2) + (F_3 \cdot W_3)$ <p>, para $F_i \in N$, y $W_i \in [0, 1]$</p> | |
| Probabilidad bayesiana | Perfiles corregidos | $P(i) = P(F_1) + P(F_2) + P(F_3) - [P(F_1) \cap P(F_2)] - [P(F_1) \cap P(F_3)] - [P(F_2) \cap P(F_3)] + [P(F_1) \cap P(F_2) \cap P(F_3)]$ | <ul style="list-style-type: none"> Permiten establecer pesos entre parámetros de forma objetiva. Útiles en modelos de distribución potencial. Basados en la teoría de la probabilidad. Es fácil introducir datos de campo. Requieren una cierta representatividad muestral de los datos utilizados. |
| | Perfiles ponderados | | |
| Regresión logística | | $P(i) = \frac{1}{1 + e^{a \cdot F_1 + b \cdot F_2 + c \cdot F_3}}$ | <ul style="list-style-type: none"> Evaluación de riesgo (avenidas, inundaciones, etc.) Evaluación de hábitats o distribución potencial de especies. Están implementados en muchas aplicaciones. Basados en la teoría de la probabilidad. Requieren cierta representatividad muestral de los datos utilizados. |
| Redes neuronales | |  | <ul style="list-style-type: none"> Conjunto de nodos que reciben datos y generan otros en función de una serie de reglas. Se parecen conceptualmente a las sumas ponderadas. Son compatibles con muchos sistemas. Simulan bien la forma de razonar del cerebro. Actúan como "cajas negras", ya que no ayudan a comprender los procesos modelizados. |

Tabla 3.-Resumen de algunos métodos de integración de información en un SAD.

1.1.3.1.2 Lógica proposicional

Según hemos visto, una base de conocimiento puede ser casi cualquier "cosa" donde se almacene información sobre cómo hacer "algo". Pues bien, la forma en la que se almacena este conocimiento condiciona considerablemente la utilidad de la base de conocimiento. En este caso nos hemos basado en la lógica proposicional como fundamento teórico para almacenar las reglas, restricciones y relaciones entre las distintas entidades de la base de conocimiento.

La lógica (del griego, logos, "palabra", "proposición", "razón"), es una disciplina de la filosofía que estudia los principios formales del conocimiento humano. Su principal análisis se centra en la validez de los razonamientos y argumentos, por lo que se esfuerza por determinar las condiciones que justifican que el individuo, a partir de proposiciones dadas, llamadas premisas, alcance una conclusión derivada de las primeras. La lógica se encarga de analizar la estructura y el valor de verdad (o grado de cumplimiento) de las proposiciones.

El concepto clave en la lógica es el de **proposición**. Se puede definir como un enunciado en el que se afirma algo, que puede ser verdadero o falso. Suele ser la expresión de un juicio y, por lo tanto, todo lo que se considera en un juicio tiene su reflejo en la proposición. Según la definición clásica, una proposición es un discurso enunciativo que expresa un juicio y posee un significado que es verdadero o falso (McElvain 2001).

Pero ¿qué ventajas tiene generar una base de conocimiento basándose en la lógica proposicional?:

- En primer lugar se trata de una metodología que trabaja con conceptos abstractos y relativamente ambiguos, similares a los que se usan en la gestión de ecosistemas. Estos conceptos o entidades son difícilmente expresables con funciones matemáticas clásicas.
- Por otro lado, aunque la solución matemática es posible en principio, casi nunca tenemos el conocimiento necesario para formular de esta manera la información que tenemos sobre el medio. Entendemos más la naturaleza en forma de conceptos ambiguos y de afirmaciones más o menos abstractas. Hay muy pocos

casos en los que un proceso ecológico se pueda resumir en una serie de ecuaciones matemáticas.

La forma en la que se generan bases de conocimiento en EMDS se fundamenta en la lógica proposicional. Es decir, el conocimiento se almacena mediante la formulación de proposiciones que se relacionan entre sí mediante una serie de operadores que luego describiremos. Estas proposiciones tienen un carácter axiomático, por lo que han de ser consideradas válidas para que el modelo funcione. Dentro de estos axiomas se almacena realmente el conocimiento que tenemos sobre la forma de resolver el problema en cuestión. Por ejemplo, si queremos elaborar una base de conocimiento que nos ayude a identificar lugares aptos para la implantación de colmenas, tendríamos que incluir los siguientes axiomas: "los lugares donde abunden las plantas melíferas serán aptos para el establecimiento de colmenas", "es mejor un lugar que está cerca de un carril transitable", "es preferible un lugar con poco riesgo de heladas". Para que el sistema resuelva nuestro problema hemos de asumir que las proposiciones anteriores son válidas (es decir, son axiomas). Este paso supone uno de los más importantes de todo el proceso, ya que implica una simplificación de la realidad en algunas ocasiones. Obtener una serie de axiomas lo suficientemente válidos como para que el modelo funcione, pero lo suficientemente sencillos como para poder trabajar con ellos, implica un importante trabajo de síntesis y recopilación de información.

Cada uno de los axiomas puede ser representado mediante un mapa que indicaría la aptitud del territorio desde el punto de vista del factor descrito en la proposición. Así, un mapa de vegetación donde las distintas clases lleven asignado un valor de aptitud desde el punto de vista de su potencialidad melífera, nos permitiría plasmar en el territorio el conocimiento que subyace al primer axioma del ejemplo.

Pero todavía quedan dos aspectos importantes a tener en cuenta:

- Cómo se implementan en una base de conocimiento conceptos como "lugares donde abunden las plantas melíferas", o "... poco riesgo de heladas", etc.
- Cómo se relacionan entre sí los distintos axiomas de una base de conocimiento.

Ambos aspectos se resuelven mediante la lógica difusa:

1.1.3.1.3 Lógica difusa

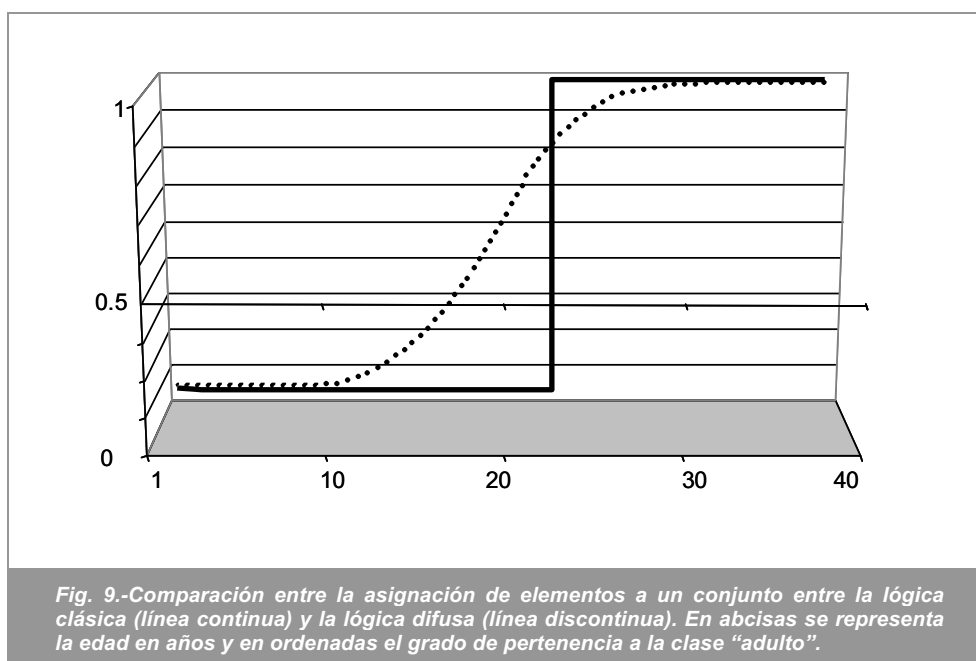
La lógica difusa (también denominada lógica borrosa, del inglés *Fuzzy logic*), es una generalización de la teoría clásica de conjuntos (Martínez 2000). Los conceptos básicos de esta teoría fueron postulados por Zadeh, en varios artículos publicados en los años 1960 y 70 (Zadeh 1965, Klir 2000). Hoy día se trata de uno de los campos más dinámicos de las matemáticas, habiéndose convertido en un pilar básico para la inteligencia artificial, redes neuronales, sistemas expertos, robótica, etc. Este gran éxito se debe a que aporta una nueva forma de tratar dos aspectos primordiales en casi todas las disciplinas: la cuantificación del grado de pertenencia de un elemento a un conjunto determinado, y la forma en la que se pueden relacionar conceptos mediante operadores difusos (Reynolds & Peets 2000). En general la lógica difusa es especialmente útil para procesar información cuyos atributos tienen zonas de transición gradual (Sainz de la Maza 2000) (transiciones entre tipos de suelo, formaciones vegetales, rangos climáticos, etc.). Asimismo, la lógica difusa nos ofrece las herramientas teóricas para tratar conceptos expresados en lenguaje natural.

La aplicación de la lógica difusa a la gestión de recursos naturales es todavía reciente pero se encuentra en pleno auge (Lexer y col. 2000). Podemos encontrar algunas aplicaciones en teledetección (Fithian y col. 1999, Hall 2000), evaluación de impacto ambiental (Holland 1994), e incluso en fitosociología (Moraczewski 1993a, Moraczewski 1993b), donde la lógica difusa puede jugar un papel importante en la caracterización de los sintaxones.

En la teoría clásica de conjuntos, la pertenencia de un elemento a un conjunto es una propiedad binaria (cierto o falso). Sin embargo, en la lógica difusa la pertenencia a un conjunto es una función cuyo valor puede estar en el rango 0-1. El elemento tendrá un valor nulo cuando haya una certeza sobre la no pertenencia al conjunto en cuestión. Será 1 cuando la pertenencia es absolutamente segura. Para el resto de situaciones, el elemento tendrá un grado de pertenencia entre 0 y 1.

Por ejemplo, todos tenemos un concepto más o menos claro de lo que significa ser adulto. Desde el punto de vista legal, alguien es adulto cuando cumple los 18 años de edad. Si tienes 17 años, 11 meses y 30 días, no eres adulto. Pero con unas

horas más, ya sí se puede considerar adulto. Esta distinción rígida, basada en la teoría de conjuntos clásica, es útil en algunos casos. Sin embargo, el concepto biológico de adultez es más difícil de acotar. Representar el grado de madurez que implica ser adulto mediante la cifra de 18 años, es poco fiable. Veamos cómo resuelve esto la lógica difusa. Es evidente que alguien con 5 años no es adulto. También está más o menos claro que una persona ha alcanzado su madurez a los 25 años. Pero, ¿y a los 13, se es adulto?. No totalmente, pero sí hay ciertos rasgos que denotan madurez propia de los adultos. De esta manera, podemos construir una curva "difusa" que traduzca la edad biológica en un grado de pertenencia al conjunto de los adultos. La función matemática que relaciona la variable independiente (edad biológica, en nuestro ejemplo) con el grado de pertenencia al conjunto en cuestión (variable dependiente), se denomina **función de pertenencia**. Estas funciones son uno de los dos aspectos claves de la lógica difusa comentados anteriormente. La figura 9 muestra gráficamente este ejemplo.



Mediante las funciones de pertenencia se pueden "fuzzydificar" (Sainz de la Maza 2000) todos los factores contemplados en la base de conocimiento que resuelve el problema que nos ocupa. En el caso de EMDS, la asignación de funciones de pertenencia a cada uno de los factores se realiza en la aplicación Netweaver. En

este programa, los rangos de aptitud difusa oscilan entre -1 (totalmente fuera del conjunto) y 1 (totalmente dentro del conjunto).

En general hay dos formas de convertir un rango discreto de valores (independientes) en un rango difuso: es posible aplicar una función matemática conocida, o bien podemos ir asignando valores concretos a puntos determinados y luego interpolar el resto del rango. Por ejemplo, retomemos uno de los axiomas comentados anteriormente "la aptitud del territorio para albergar asentamientos apícolas es mayor cuanto menor es el riesgo de heladas (expresado en días con heladas)". Es posible que sepamos la relación matemática que hay entre el riesgo de heladas y la aptitud del terreno para la supervivencia de las colmenas. En este caso bastaría con aplicar esta expresión (que puede ser lineal, logarítmica, exponencial, etc.) a todo el rango de días con heladas que hemos identificado. Esta función asignará un valor de pertenencia determinado a cada uno de los valores de riesgo de helada. Sin embargo, lo más normal es que no sepamos la función exacta, sino que conozcamos algunos valores umbrales o hitos. Podemos saber, por ejemplo que si hay más de 20 días de heladas al año, es casi seguro que la colmena muera congelada (la función de pertenencia asigna un valor de -1 en este caso). Si hay menos de 4 días con heladas, nos aseguramos la supervivencia de la colonia (ahora asigna un valor de 1 , total pertenencia al conjunto "lugares aptos para albergar colmenas" y total cumplimiento del axioma). El resto de valores nos es desconocido. Pero podemos suponer que hay una relación lineal entre el número de días de heladas y la aptitud entre el rango definido previamente. Pues bien, Netweaver permite establecer este tipo de interpolaciones lineales o de cualquier otro tipo entre valores conocidos de un rango de aptitud (Reynolds 1998b). Este proceso de "fuzzydificación" se aplica a cada uno de los factores identificados en la base de conocimiento construida para resolver el problema decisional propuesto en esta tesis: identificación de zonas aptas para el tratamiento selvícola de pinares de repoblación con objeto de maximizar la regeneración de la vegetación natural.

La lógica difusa incide notablemente en otro aspecto de interés dentro de nuestros objetivos: aporta un método de combinación de capas mucho más versátil que la lógica binaria clásica. Se trata de los **operadores difusos**, que describimos a continuación.

Como vimos en el apartado de III.2.1.1.3, la combinación de capas mediante operadores consiste en la conversión de un condicionante semántico (o, y, etc.) en una operación matemática entre dos capas de información. En este caso es equivalente hablar de capas o de elementos (axiomas) de una base de conocimiento, porque las últimas siempre se corresponden con una capa de información temática. Los operadores más importantes son:

- **Operador Fuzzy OR:** Si combinamos dos axiomas como por ejemplo: “una zona es apta para implantar colmenas si hay disponibilidad de agua **O** si hay carriles de acceso”, mediante el operador **OR** binario, observamos que en la capa resultado, las zonas aptas son aquellas que cumplen una o las dos condiciones. Generalizando este concepto obtenemos la definición del operador **fuzzy OR**: la capa resultante de combinar una serie de capas mediante este operador, adquirirá en cada punto el máximo de los valores contenidos en cada una de las capas iniciales. Se trata de un operador altamente conservativo que apenas asume ningún riesgo. Las zonas aptas serán aquellas que sean adecuadas para los dos factores o para cada uno de ellos individualmente.
- **Operador Fuzzy AND:** Al igual que antes, si combinamos las capas anteriores mediante el operador **AND** binario, obtendremos que el resultado es apto sólo si lo son los valores de los axiomas combinados. Sin embargo, el operador **fuzzy AND** implementado en Netweaver, asigna a la capa resultante el siguiente valor:

$$fuzAND (F_1, F_2) = \min (F_1, F_2) + [media (F_1, F_2) - \min (F_1, F_2)] \cdot [\min (F_1, F_2) + 1] / 2$$

Esta ecuación está pensada para producir un efecto conservativo (es decir, sesgado hacia valores cercanos a 1) en la combinación cuando hay evidencias parciales de incumplimiento de alguno de los axiomas combinados (es decir, cuando algún factor adquiere valores cercanos o poco inferiores a 0). Por ejemplo, si dos axiomas A y B tienen valores de cumplimiento de 1 (totalmente cierto) y 0 (indeterminado) respectivamente, su combinación mediante este operador dará el valor de 0.25. Si en este mismo caso usáramos el operador **AND** binario, obtendríamos un valor de 0. Por eso se dice que **Fuzzy AND** es relativamente conservativo.

- **Otros operadores:** Netweaver incorpora otros operadores difusos para combinar información (**NOT**, **XOR** y **SOR**). Sin embargo, no los describiremos aquí ya que no han sido utilizados en nuestra base de conocimiento y tienen una utilidad reducida en general. Por otro lado la lógica difusa aporta otros operadores que no se han implementado en Netweaver pero que consideramos interesantes. Se trata de los denominados **producto algebraico**, **suma algebraica** y **operador gamma**. La expresión para su cálculo puede observarse en la tabla 3. El producto algebraico aporta un concepto interesante en la combinación de capas. Produce un efecto antagonista entre las capas combinadas. Es decir, el resultado es siempre menor que el menor de los valores existentes en las capas combinadas. Lo contrario ocurre con la suma algebraica, que introduce un efecto de sinergia (el resultado siempre es mayor que cualquiera de los valores combinados). Ambos efectos son interesantes en la generación de modelos sobre aspectos ecológicos, ya que suelen darse con frecuencia en la naturaleza.

