

2. LA DISPONIBILIDAD, COBERTURA Y CALIDAD DE LOS DATOS

El diseño de los SIG para el análisis de riesgos en América Latina también enfrenta problemas críticos de disponibilidad, cobertura y calidad de los datos. Encontrar fuentes de datos de una calidad y cobertura adecuadas para el análisis de riesgos es un obstáculo al diseño e implementación de SIG en la región. En muchos casos, datos actualizados con una cobertura geográfica completa, sencillamente, no existen. En el Perú, por ejemplo, sólo 49% del territorio tiene cartografía a escalas de 1:100,000 o mayor. La cobertura cartográfica de gran parte del país es incompleta, debido a nubosidad o humo en las fotos aéreas o imágenes de satélite que se utilizan como fuentes, dado que la estación seca en gran parte del país coincide con la época de quema, cuando los agricultores preparan sus campos para el cultivo. Son pocas las áreas del país (generalmente donde se desarrollan actividades mineras significativas) que tienen mapas geológicos detallados. El mapa catastral agrícola más reciente ya tiene más de 20 años. En toda América Latina, los censos nacionales de población, optimistamente, se llevan a cabo cada 10 años.

Dados los problemas de disponibilidad y cobertura de datos cartográficos analógicos, los sensores remotos ofrecen a veces la única fuente de datos con una cobertura temporal y espacial completa. El uso de imágenes LANDSAT o SPOT y de fotos aéreas está documentado extensivamente en la literatura. Existen técnicas para la clasificación de imágenes que permiten la detección de laderas susceptibles a deslizamientos (British Geological Survey, 1995) y otras amenazas. Sin embargo, en muchas zonas, las imágenes de satélite sufren de los mismos problemas de nubosidad y humo que los mapas que se derivan de ellos, mientras que las imágenes de radar aún no están fácilmente disponibles. Un obstáculo para el uso de imágenes remotas en la región es su alto costo, signifi-

cando que son inaccesibles para muchas instituciones públicas y privadas. Mientras que en los países de la región hay una disponibilidad creciente de cartografía digital, muchas instituciones no tienen acceso a ella, salvo que estén dispuestas a pagar precios exorbitantes. Como resultado, sigue siendo más barato, en muchos casos, digitalizar los datos requeridos de la cartografía analógica.

Aparentemente, estos problemas son comunes a los países en desarrollo. En una aplicación documentada en Ghana (Minor, 1994), la incertidumbre referente a la disponibilidad, calidad y formato de datos frenó el progreso del desarrollo de bases de datos. Los diseñadores de SIG concluyeron que muchas demoras y problemas podrían haberse evitado, si es que se hubiera planificado desde el principio para el peor escenario posible de disponibilidad de datos. Problemas aún más serios ocurren en Beirut, Líbano (Azar, Ferreira, 1994), donde la mayoría de los datos disponibles tiene una antigüedad de por lo menos 30 años. Davis y Bickerman enumeran una larga lista de problemas que tienen que enfrentar para obtener datos destinados al desarrollo de un sistema nacional de gestión de desastres hipotético en un país en vías de desarrollo (Davis et al., 1993). Éstos incluyen: la ausencia de datos referenciales, de mapas topográficos, de datos históricos sobre ocurrencia de amenazas, de datos sociales y económicos sobre patrones de vulnerabilidad; problemas de calidad, formato y confiabilidad de los datos. Todos estos problemas se presentan en América Latina.

Los datos en América Latina también padecen problemas propios de la región. Como se ha mencionado en el punto 5, un número importante de desastres, de pequeña y mediana escala ocurren en América Latina, en regiones periféricas con bajas densidades poblacionales y poca importancia económica. No es sorprendente que muchas de estas regiones tengan datos cartográficos y estadísticos con una cobertura incompleta y de baja calidad. A la vez, debido a la existencia de ritmos dinámicos e inestables de cambio social, demográfico, territorial y económico, particularmente en regiones de este tipo, tanto los datos espaciales como sus atributos se desactualizan con mucho mayor rapidez que en países más estables del mundo. Por ejemplo, la aplicación del SIG en San Martín, Perú (Minaya, 1994), fue desarrollada en un contexto donde la población de la región se había duplicado en sólo 10 años. Dentro de la misma región, el crecimiento poblacional de algunas ciudades tenía una tasa anual de más del 28%. Movimientos demográficos, mediante la migración o el desplazamiento forzado, caracterizan a países como Colombia, Perú, Brasil y los países de América Central. Los cambios en la geomorfología y el paisaje son igualmente dinámicos. La deforestación afecta entre 0.5% y 1.5% de los bosques de la región ca-

da año, y está relacionada con la creciente ocurrencia y magnitud de inundaciones, sequías, deslizamientos y otras amenazas hidrometeorológicas. Los rellenos y drenajes en zonas costeras o pantanosas pueden aumentar dramáticamente la amenaza sísmica. Ciudades en expansión rápida causan cambios en los ecosistemas que las rodean. Regiones que hace 25 años estuvieron completamente deshabitadas y donde la manifestación de amenazas de gran escala no hubiera causado ni pérdidas ni desastres; a menudo, tienen ahora una población grande y creciente con redes complejas de ciudades, actividades económicas e infraestructura en riesgo. En regiones como éstas, es muy posible que modelos espaciales-temporales que se basan en datos con una antigüedad de sólo 10 años no representan los niveles actuales de riesgo.

La calidad de los datos también está relacionada con problemas de incompatibilidad entre series de datos de períodos diferentes. En América Latina, los tipos de datos recolectados y los sistemas utilizados para clasificarlos a menudo cambian de censo a censo, haciendo difícil producir series temporales de un atributo determinado. Adicionalmente, hay una falta de confianza en la calidad de datos censales oficiales en algunos países, como México, aun cuando estos datos son actualizados. En algunas zonas, datos detallados derivados de estudios particulares pueden utilizarse para llenar el vacío; sin embargo, entre ellos hay normalmente problemas de incompatibilidad y de falta de cobertura que impiden su integración en un SIG, salvo en el caso de aplicaciones locales.

El cambio social, económico y territorial, a la vez, agrava estos problemas de compatibilidad. El SIG de San Martín (Minaya, 1994) utilizó datos de los censos nacionales de 1981 y 1993. En ese período, el crecimiento poblacional en la región fue tan rápido que se crearon dos nuevas provincias y una gran cantidad de nuevos distritos, mediante la subdivisión y la amalgamación de entidades político-administrativas existentes, significando que las topologías político-administrativas de 1981 y 1993 fueron completamente diferentes. Adicionalmente, muchos distritos no tenían límites claramente definidos. Los mapas producidos por el Instituto Nacional de Estadística e Informática (INEI), que implementa los censos nacionales y los mapas producidos por el Instituto Geográfico Nacional (IGN), responsable de la cartografía nacional, presentaron áreas y topologías diferentes. De por sí, la cartografía producida por el IGN, a una escala de 1:100,000, se proyecta de escalas menores (1:1'000,000), llevando a problemas de imprecisión en la ubicación, área y forma de centros urbanos y otras entidades espaciales. Como resultado de este tipo de problemas, la construcción de la topología en un SIG puede volverse un proceso arbitrario e incierto.

La calidad de los datos también tiene relación con las metodologías utilizadas para recolectarlos. En el caso de los datos sobre pérdidas y desastres ocurridos, los que a menudo son importantes para el análisis de riesgos; la calidad especulativa y no uniforme de estas metodologías impide su incorporación en análisis comparativos. En la mayoría de los países, no existen bases de datos oficiales sobre la ocurrencia de desastres y pérdidas, sino múltiples informes sobre desastres específicos elaborados con metodologías diferentes. Frente a esta realidad, a menudo la única fuente de datos sobre los desastres ocurridos son los reportes periodísticos; los cuales, sin embargo, presentan problemas propios, tal como fue visto en la aplicación DesInventar (LA RED, 1996). Es posible que exista una relación inversa entre la cobertura de los datos sobre un desastre y la distancia en el espacio y el tiempo del desastres de la fuente periodística consultada. Es muy probable que muchos desastres, ocurridos en regiones aisladas, antes de la era de las telecomunicaciones modernas, nunca fueron registrados. Esto implica que los datos periodísticos pueden tener una distorsión a favor de desastres recientes en áreas accesibles, y que su uso daría origen a series temporales y espaciales distorsionadas.

La calidad de los datos también se ve afectada por las variaciones en la terminología que se utiliza para describir amenazas, pérdidas, poblaciones y desastres. En la aplicación de DesInventar, por ejemplo, las entidades espaciales fueron georreferenciadas utilizando los códigos censales y los nombres oficiales de las unidades político-administrativas respectivas en cada país. En las fuentes de datos sobre desastres ocurridos, sin embargo, se utilizan con frecuencia nombres coloquiales para la georreferenciación. La topología del espacio en uso popular raramente coincide con la topología oficial; sin embargo, influye en la definición de muchas bases de datos, incluso las oficiales. El uso repetido de nombres comunes; por ejemplo, San Juan, Santa Rosa, complica el problema. Problemas similares ocurren con la terminología utilizada para describir las amenazas y las pérdidas con múltiples variaciones locales en los términos utilizados. Ciertos atributos, como el número de muertos en un desastre, son robustos y pueden fácilmente medirse y compararse en diferentes países. En cambio, en el caso de amenazas como huracanes o deslizamientos, hay muchos términos locales que se utilizan para describir diferencias sutiles, pero importantes, en el tipo de amenaza. Las definiciones utilizadas en los datos sobre el impacto social y económico de un evento son más variables aún y el significado de uno u otro término varía según quién lo utiliza y quién lo interpreta.

Los datos sobre los desastres y los riesgos a menudo se manipulan para servir a intereses particulares, lo cual también afecta la calidad de los mismos. Poblaciones locales o gobiernos municipales a veces exageran los datos sobre las pérdidas sufridas o riesgos enfrentados, para negociar más recursos o asistencia de un gobierno nacional. Éstos, a su vez, hacen lo mismo para obtener mayor ayuda internacional. En cambio, a veces se subestiman las pérdidas y el riesgo, para dar una imagen de previsión y competencia. Los datos sobre desastres también cambian con el tiempo, empezando a veces con informes exagerados y sensacionalistas e imprecisos, hasta que aparecen informes menos sensacionalistas y mejor informados sobre las pérdidas y daños ocurridos.

En resumen, hay numerosos problemas para conseguir datos con una cobertura espacial, temporal y calidad adecuada para alimentar un modelo espacial-temporal que sirva al análisis de riesgos en América Latina. La ausencia de datos actualizados y con una cobertura espacial adecuada, particularmente en regiones aisladas que experimentan procesos rápidos de cambio; el alto costo de datos remotos y digitales; la incompatibilidad entre diferentes series y fuentes de datos; la falta de precisión y existencia de error en la cartografía; y la ausencia de métodos homogéneos para la producción de datos sobre los desastres y pérdidas ocasionadas son sólo algunos de los problemas que tienen que enfrentarse comúnmente en el diseño de un SIG para el análisis de riesgos en la región. Tomando en cuenta estos problemas, es preciso explorar estrategias alternativas para poder satisfacer los requerimientos de datos.

Una primera estrategia consiste en explorar la disponibilidad de datos en paralelo con el diseño del modelo espacial-temporal. A menudo, se gastan esfuerzos significativos en el diseño de modelos espaciales, para luego descubrir que no existen los datos necesarios para implementar el modelo. El levantamiento de inventarios o metabases de datos, entonces, debería acompañar al proceso de diseño de modelos espaciales, mediante un proceso iterativo de selección de datos y desarrollo de modelos, hasta que se llegue a un modelo funcional a las necesidades de la aplicación, que puede implementarse con los datos disponibles. En la medida que el modelo escogido sea lo más sencillo posible, tal como se recomienda en el punto 1, esto puede reducir a su vez los requerimientos de datos. Esta estrategia puede significar un ahorro significativo en tiempo y recursos en el desarrollo de una aplicación. A la vez, antes de iniciarse el levantamiento de inventarios de datos, es preciso hacer explícita la terminología convencional que se utilizaría en una aplicación. En la aplicación DesIn-

ventar, por ejemplo, se desarrolló un diccionario de términos por utilizarse a nivel regional, antes de dar inicio a la recopilación de los datos por incorporarse.

En América Latina, metabases de datos en general aún no existen, salvo en el caso de datos remotos como imágenes LANDSAT, que están disponibles en el Internet. Por lo tanto, es preciso adoptar una actitud creativa y proactiva para identificar, acceder y combinar fuentes de datos inconsistentes, incompatibles e incompletos. Aunque muchos datos, supuestamente, son de dominio público, se requiere en muchos casos aplicar técnicas de negociación de diversa índole para poder obtener acceso a ellos. Tejer y remachar los retazos de datos conseguidos, para lograr la cobertura espacial, temporal y de atributos requeridos, también exige una voluntad y capacidad para tomar decisiones, aun cuando los datos y sus atributos son aparentemente incompatibles y contradictorios. Al igual que en el diseño de modelos espaciales-temporales de riesgo, la mejor estrategia para conseguir, completar e integrar los datos necesarios para alimentar el modelo es la aplicación de *inteligencia* acerca de los escenarios de riesgo bajo análisis. Para el diseñador, la *inteligencia* es crucial, a fin de poder tomar decisiones y eliminar errores en la integración de los datos.

En San Martín, Perú, por ejemplo, la cartografía de la región a escala 1:1'000,000 muestra tanto las carreteras existentes como las proyectadas, a pesar de que algunas de éstas últimas nunca fueron construidas. En casos como éste, la aplicación de *inteligencia* sobre la región es la única manera de tomar decisiones acerca de qué carreteras proyectadas existen o no, y para evitar la producción de errores en la información producida. La *inteligencia*, entonces, puede ayudar al diseñador a: eliminar errores en los datos cartográficos y de sus atributos, tomar decisiones cuando existen contradicciones, saber cuándo y cómo aplicar técnicas de interpolación o de corrección de contornos y escoger atributos que reflejen los supuestos implícitos en el modelo espacial. En otras palabras, la imaginación geográfica del diseñador es clave para poder identificar errores y navegar con éxito en un mar de datos absurdos.

Una segunda estrategia consiste en la producción de información primaria para satisfacer las necesidades de una aplicación. En zonas donde no existe una cobertura de datos actualizados, a una escala adecuada, es posible generar datos primarios mediante métodos como los autocensos o el análisis de riesgos participativos (Medina, 1996). Aplicaciones en el Perú (Maskrey et al., 1992) demuestran que los autocensos, donde comunidades locales levantan sus propios datos sociales y económicos, pueden organizarse para cubrir áreas relativamente grandes, produciendo series de datos actualizados similares a los encontrados en los

censos nacionales. El análisis de riesgos con participación local, en el cual la población produce datos primarios sobre la ocurrencia histórica de amenazas y desastres en cada localidad, fue aplicada en la región San Martín, cubriendo una población de más de 600,000 personas en una área de más de 50,000 km². En otras palabras, si bien la producción de información primaria utilizando técnicas como PRA (Hall, 1996) puede ser adecuada para aplicaciones detalladas a nivel local, es posible también generar datos primarios para alimentar aplicaciones a nivel regional. En algunos contextos, en comparación con las dificultades y costos implícitos en la adquisición e integración de datos secundarios, esta estrategia puede resultar un ahorro significativo de tiempo y recursos en la implementación de una aplicación. Asimismo, permite mayor control sobre la calidad de los datos producidos y su relevancia para las necesidades de la aplicación.

Cuando las fuentes existentes de datos cartográficos son incompletas, desactualizadas, demasiado caras o tienen problemas de nubosidad, el uso de la videografía puede representar otra estrategia de costo relativamente bajo para generar datos cartográficos directos (Doyle et al., 1994). Imágenes de video tomadas desde avionetas o helicópteros, utilizando equipos no profesionales, ofrecen una tecnología a bajo costo para capturar los atributos espaciales y temporales de escenarios de riesgo y de desastres ocurridos. Si es que se combina con el uso de GPS y técnicas sencillas de análisis de imágenes, la videografía puede ser una estrategia eficiente y de bajo costo, particularmente en áreas remotas e inaccesibles. Otra alternativa que puede tener mayor importancia en el futuro es el uso de orthophotos digitales. Éstas son representaciones digitales de fotos aéreas con distorsiones debido a topografía, condiciones atmosféricas y sistemas de cámara corregidos. Particularmente, en el caso de aplicaciones locales, las orthophotos digitales permiten una resolución más alta que las imágenes de satélite (Tessele, 1994). Es posible, además, que la disponibilidad creciente de imágenes de satélite declasificadas de los países de la ex Unión Soviética genere una reducción en los precios de datos remotos en América Latina, aumentando su accesibilidad y uso en aplicaciones para el análisis de riesgos.