

CAPÍTULO 10

ALMACENAMIENTO, ACCESIBILIDAD Y DIFUSIÓN DE DATOS [HOMS G06]

10.1 INTRODUCCIÓN

La disponibilidad de datos de calidad suficiente es la base de la hidrología, desde la investigación hidrológica hasta la evaluación de recursos hídricos, pasando por todo tipo de aplicaciones operacionales. El término “suficiente” no tiene el mismo significado en todas las aplicaciones, como se explica en otra parte de esta Guía. En este capítulo se explicarán los conceptos de “buena calidad” y accesibilidad de los datos, así como la disponibilidad de datos para distintos tipos de usuarios.

10.1.1 Importancia de los datos

La compilación de datos hidrológicos, ya sean éstos mediciones de precipitación, registros de nivel de agua, mediciones de caudal, datos de observación de aguas subterráneas o de muestreo de la calidad del agua, culmina en una serie de datos de utilidad para la adopción de decisiones. Las decisiones se pueden adoptar directamente a partir de mediciones de datos en bruto, estadísticas derivadas de éstos o resultados de distintas etapas de modelización de datos elaborados, pero los datos obtenidos serán el fundamento de todas las decisiones.

Los datos en bruto –tanto en forma de formularios de campo como de gráficas o informes– deberán estar disponibles después del tratamiento. Algunos errores de transcripción y tratamiento podrían no ser evidentes hasta que los examinen los usuarios. También podría ser necesario comprobar las transcripciones de datos originales, o reevaluar la interpretación de una procedencia dudosa efectuada por el operario. Podría ser necesario tomar nuevas muestras con objeto de verificar los registros de un emplazamiento que haya experimentado modificaciones, o mejorar las pautas de trabajo en respuesta a cambios tecnológicos. En ambos casos, podría ser necesario procesar de nuevo los datos. Por todas estas razones, los datos originales deberán ser archivados en condiciones de seguridad. El lugar de almacenamiento estará apartado de la base de datos electrónica, y será un lugar físicamente seguro.

Los datos son muy valiosos, ya que su obtención requiere gran cantidad de tiempo y de dinero. Su gestión es, por consiguiente, una tarea importante que se ha de llevar a cabo con eficacia para aprovechar al máximo los resultados de la inversión. Un

archivo hidrológico sólido y bien gestionado será el compendio de los esfuerzos dedicados a la obtención de los datos, y constituirá una fuente de datos fiables y de alta calidad durante decenas o centenares de años. Un archivo de mala calidad, bien por falta de previsión, bien por mala gestión, puede dar lugar a años de trabajo adicional en la obtención de datos o en la elaboración de modelos, que se traducirán en decisiones equivocadas. Un archivo puede llegar a ser redundante en muy poco tiempo. Además, la mala calidad de los datos y de las bases de datos perjudicará la adopción de decisiones de planificación y el diseño de estructuras de ingeniería.

La integridad de los datos es también un aspecto importante. Frecuentemente, el éxito o las limitaciones de las actividades hidrológicas solo serán comprensibles analizando la calidad de los datos en que aquéllas están basadas.

La magnitud de las tareas de gestión de datos depende, por supuesto, de la magnitud de las operaciones: un proyecto hidrológico detallado y de gran escala requerirá unas técnicas de gestión más complejas y un mayor volumen de almacenamiento informático que un proyecto de menores dimensiones destinado a medir un número limitado de variables durante un breve período de tiempo. La magnitud de la gestión de datos depende también de otros factores; además del volumen de datos, suele haber limitaciones presupuestarias que impiden dedicar suficiente tiempo al archivado y disponer de sistemas de gestión de datos en gran escala. La capacidad de gestión de datos del personal puede ser una limitación, y el nivel de eficacia en la gestión de los datos dependerá de la experiencia y de las aptitudes del personal.

A pesar de las posibles diferencias de escala, todos los sistemas de gestión de datos hidrológicos tienen varios aspectos esenciales en común. En el presente capítulo se describen detalladamente todos estos aspectos de la gestión de datos desde un punto de vista general y, ocasionalmente, se señalan algunos casos extremos.

10.1.2 Procesos de gestión de datos

Los datos hidrológicos han de seguir una serie de etapas desde el lugar de observación, pasando por su introducción en el sistema y su validación hasta

su difusión y utilización para la adopción de decisiones. Estas etapas son esencialmente las mismas con independencia del tipo de operación y del nivel de la tecnología utilizada para la gestión de los datos, como puede observarse en el diagrama esquemático de la figura I.10.1. En la lista de la tabla I.10.1 se resumen algunos de los conjuntos de datos utilizados en diversas etapas del proceso de gestión de datos. La lista ofrece una perspectiva general de todo el proceso de gestión de datos. Algunos de los aspectos de este proceso se examinan detalladamente en el capítulo 9. En el presente capítulo se aborda básicamente el almacenamiento, la accesibilidad y la difusión de los datos, y se indican las etapas correspondientes en el proceso de gestión.

La *Guía de prácticas climatológicas* (OMM-N° 100) contiene una descripción completa de los procedimientos recomendados de almacenamiento y catalogación de datos climatológicos. Los datos hidrológicos requieren un tratamiento ligeramente diferente por razones de eficacia en su almacenamiento, pero muchas de las consideraciones climatológicas son también válidas en el ámbito hidrológico.

Las ingentes cantidades de datos climatológicos e hidrológicos que están reuniendo muchos países podrían hacer imposible almacenar todos los datos originales. Sin embargo, es posible efectuar copias (por ejemplo, mediante escaneo electrónico) en un soporte físico que ocupe mucho menos espacio que los documentos originales, lo cual permitirá desear el material original. Los datos almacenados estarán protegidos frente a excesos o variaciones de temperatura, humedad, polvo, insectos u otros parásitos, radiaciones e incendios.

Siempre que sea posible, se conservarán dos ejemplares de los registros: uno, en el centro principal de recogida, y el otro, en un centro regional o en la oficina del observador.

En el capítulo 9 se describen los diversos tipos de datos y los sistemas de proceso y control de calidad a ellos aplicados. Los datos pueden proceder de registros manuscritos de observadores, registradores gráficos, registradores cronológicos automáticos, hojas manuscritas o archivos digitales que recojan mediciones instantáneas del caudal, y en los que se especifiquen la sección fluvial, la profundidad y los

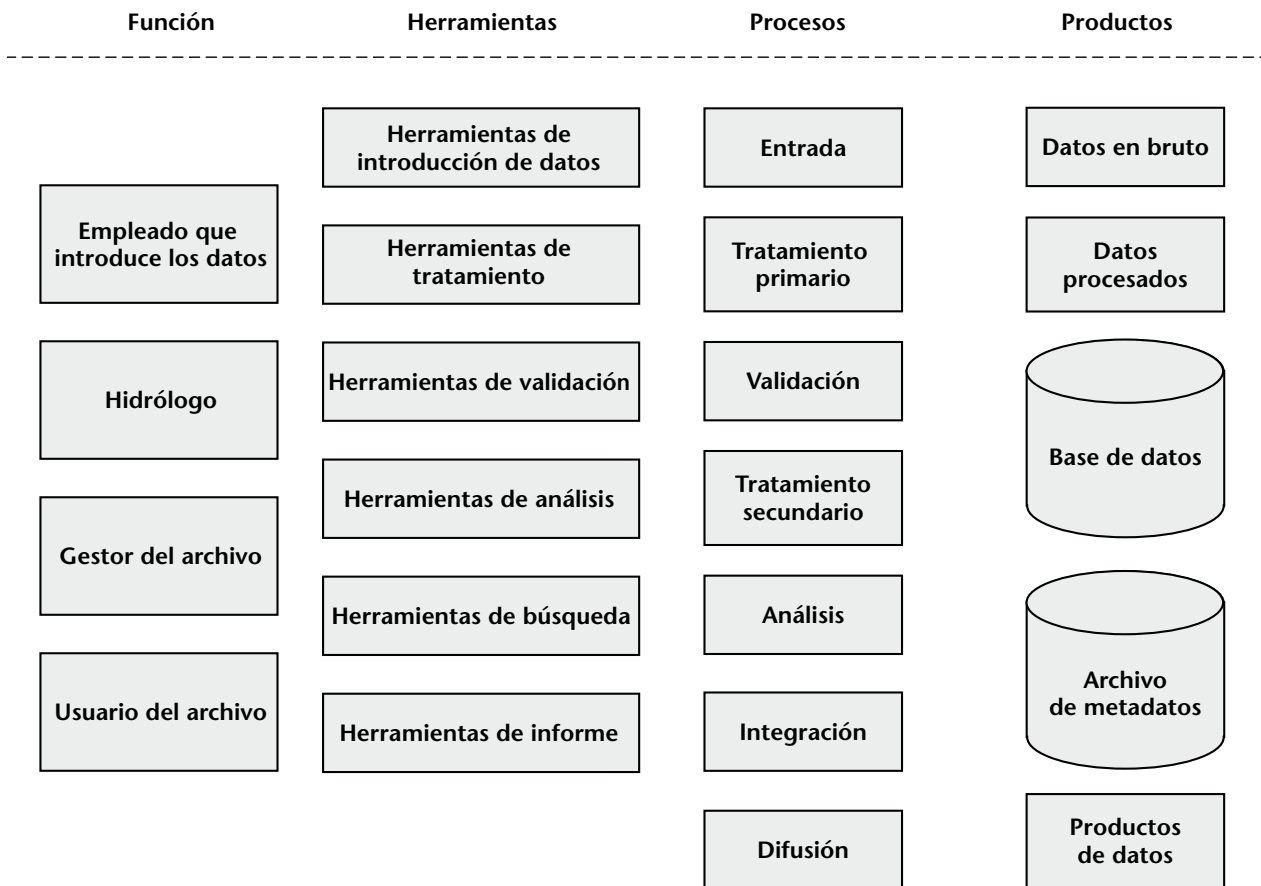


Figura I.10.1. Plan de gestión de datos

perfiles de velocidad, frecuentemente acompañados de las correspondientes descripciones.

En la tabla I.10.1 se resumen los procesos de gestión de datos, desde la introducción de datos de mediciones en bruto hasta la difusión de

Tabla I.10.1. Procesos que intervienen en la gestión de datos

<i>Proceso</i>	<i>Descripción</i>	<i>Ejemplos de tipos de datos intervinientes en el proceso</i>
Introducción	Se obtienen los datos de la fuente de información o del instrumento mediante registro manual, registro cronológico de datos, digitalización u otros métodos, y se convierten al formato apropiado para su almacenamiento	Cuaderno/hoja de lecturas de medición (datos de altura del agua leídos manualmente con periodicidad diaria o inferior en un limnómetro, con las notas adicionales necesarias para describir el contexto) Archivo de datos en bruto obtenido de un registrador cronológico automatizado vinculado a una sonda de medición, frecuentemente en formato ordinario/personalizado Gráficas obtenidas de un limnógrafo Hojas de trabajo/archivos digitales con información sobre las mediciones instantáneas de caudal (aforo), incluidos los perfiles de la sección fluvial, de la profundidad y de la velocidad del agua y otras informaciones descriptivas
Tratamiento de datos en bruto	Se almacenan los datos por medios manuales o digitales, frecuentemente por ambos	Archivos de datos digitalizados a partir de gráficas Series de datos contenidas en bases de datos o en archivos digitales de registros de medición introducidos manualmente y de información sobre la medición Archivos digitales de series de datos convertidas a partir de archivos de registro cronológico personalizados, por lo general en forma textual
Validación	Se comprueban los datos en bruto para identificar los valores erróneos, y se someten a corrección para producir una serie de datos en bruto corregida	Series de datos contenidas en bases de datos para cada etapa de corrección/validación Texto descriptivo de los cambios introducidos, y en particular de los métodos utilizados y de las razones a que respondan las correcciones
Tratamiento secundario	Esta etapa abarca el subsanamiento, en su caso, de los datos faltantes, la conversión en intervalos secundarios (por ejemplo, el cálculo de las series de valores medios y totales), la construcción de curvas de gasto a partir de las nuevas mediciones del caudal, y la conversión de los datos de altura de agua en bruto en valores de flujo o de almacenamiento embalsado	Conjuntos de datos correspondientes a cada serie nueva creada en el proceso de conversión Ecuaciones de gasto obtenidas a partir de las mediciones del caudal, y acompañadas de un texto descriptivo del proceso de decisión empleado para formular la ecuación. Por lo general, las ecuaciones de gasto evolucionarán a lo largo del tiempo, por lo que deberá mantenerse un historial satisfactorio
Seguridad y archivo	Los datos deberán ser archivados de manera que sean accesibles, aunque en condiciones seguras, además de bien documentados e indexados	Metadatos que permitan un acceso rápido y sencillo a los conjuntos de datos, así como un índice completo de la información disponible
Integración con otros datos	Operaciones que permiten visualizar los datos conjuntamente con otras fuentes de datos, como las series de datos SIG	Nuevos conjuntos de datos, como la cobertura espacial de la precipitación de lluvia obtenida de datos de pluviometría puntuales o de mapas de superficie de aguas subterráneas confeccionados en base a datos de perforaciones
Difusión de datos	Distribución de los datos en la forma apropiada para uso de los encargados de la modelización y de la adopción de decisiones, organismos públicos, etc.	Resumen de datos, anuarios, etc. Cada vez más frecuentemente, la difusión de los datos se efectúa mediante sitios web

los datos procesados, junto con los datos correspondientes.

Es evidente que cada etapa del proceso de gestión puede abarcar un gran número de conjuntos de datos, lo cual obligará a adoptar decisiones sobre los datos que serán almacenados y sobre la manera de hacerlo para poder conseguir un archivo hidrológico eficaz. Se describen a continuación diversos métodos de almacenamiento y análisis de datos, y para la elaboración, accesibilidad y difusión de la información a todos los usuarios.

Una vez completadas las fases del proceso de datos y control de calidad descritas en el capítulo 9, se archivarán los datos en un formato que permita, mediante distintos tipos de herramientas, analizarlos y transformarlos en productos, así como facilitar su accesibilidad y difusión.

10.2 **ALMACENAMIENTO Y RECUPERACIÓN DE DATOS**

10.2.1 **Almacenamiento de datos**

En la gestión de datos hidrológicos, lo más importante es determinar cuáles de los numerosos conjuntos de datos producidos han de ser almacenados. El proceso de gestión consta de numerosas etapas, desde la de registro hasta la de divulgación, y cada una de ellas puede representar uno o varios conjuntos de datos diferentes. Si se almacenasen todas las permutaciones posibles de los datos, se obtendría un archivo confuso e inmanejable. En el otro extremo, si un archivo hidrológico contiene solo conjuntos estáticos de datos procesados y validados, será imposible saber por qué medio han sido obtenidos o medidos los datos, y no será posible conocer las limitaciones del conjunto final de datos. Por ejemplo, un conjunto de datos de flujo procesados no aporta información sobre el método de medición ni sobre el proceso de obtención de aquéllos a partir de datos de nivel de agua, y no permite saber si los datos han sido corregidos ni cómo. Es, pues, necesario determinar un mecanismo de almacenamiento viable que esté situado en un término medio.

Al determinar el nivel de detalle de los datos almacenados, la consideración básica será su reproducibilidad. En cualquier proyecto hidrológico, por importante que sea, es necesario conocer y, en su caso, saber reproducir las etapas que conducen desde los datos en bruto hasta los datos procesados finales. Los usuarios de los datos procesados deberían poder conocer de un vistazo el

proceso a que han sido sometidos los datos, así como sus posibles limitaciones. Esto no quiere decir que haya que conservar todos los cambios para la posteridad, sino que habrá que conservar los conjuntos de datos en bruto, y documentar y almacenar los cambios efectuados y los supuestos en que estén basadas las etapas de validación y tratamiento. Será también importante que los usuarios de los datos puedan distinguir entre los datos originales, los añadidos para cubrir lagunas de datos, y los corregidos.

También en este caso, el volumen del almacenamiento de datos dependerá de diversos factores, como la capacidad de almacenamiento, los fondos disponibles para su almacenamiento y documentación, o el personal disponible. Inevitablemente, habrá que encontrar un punto de equilibrio entre la completitud del archivo y los recursos utilizados. El caso más complejo podría ser un gran proyecto hidrológico cuyo sistema de almacenamiento de datos permita verificar automáticamente todas las modificaciones introducidas en los conjuntos de datos del sistema, la fecha y hora en que éstas se han efectuado y el usuario que las ha efectuado, y que permita revertir secuencialmente las modificaciones para recrear cualquiera de las versiones anteriores del conjunto de datos. Un sistema más sencillo podría consistir únicamente en el conjunto de datos en bruto más el conjunto de datos finales, junto con un fichero de notas que documenten las decisiones y modificaciones acometidas. En ambos casos, el proceso es esencialmente el mismo:

- a) se conservan los ficheros de datos en bruto, ya sea en formato impreso (cuadernos de lectura de mediciones, registros en forma gráfica) o digital (ficheros de registradores de datos en bruto o de datos obtenidos por telemedición);
- b) se asocia a todos los conjuntos de datos procesados un registro de metadatos descriptivos de su procedencia y que los vincule al conjunto de datos del que ha sido obtenido;
- c) se almacenan los datos correspondientes a las etapas importantes del tratamiento, aun cuando los datos procesados constituyan solo una etapa intermedia entre los datos en bruto y los difundidos. Su importancia se determinará en función de la magnitud del sistema de gestión de datos. Por ejemplo, si fuera necesario convertir únicamente una serie de niveles de agua en bruto en una serie de flujos medios mensuales, convendría almacenar, por lo menos, el conjunto de datos validados junto con los datos de flujo diario correspondientes, además de los datos en bruto y de los valores finales del flujo medio mensual;

- d) se documentan los cambios parciales masivos introducidos en cada serie tomando como referencia el conjunto de datos, anotando, por ejemplo, la aplicación de un valor de referencia a un período de un registro de valores de altura, o la conversión de éste en valores flujo mediante una curva de gasto, que constituirá un conjunto de datos por derecho propio;
- e) se documentan los cambios introducidos en los datos, por interpolación de datos faltantes o por corrección de valores individualizados, tomando como referencia cada uno de los datos modificado, adjuntando al registro una nota que indique al usuario que la serie ha sido alterada; y
- f) llegado a ese punto, se obtiene un catálogo completo de las correcciones introducidas y de sus causas, y el usuario de los datos podrá conocer las razones por las que se han modificado los valores en bruto y los métodos utilizados, y podrá reproducir el conjunto de datos a partir de los datos en bruto.
- b) la facilidad de mantenimiento;
- c) los costos, incluidos los desembolsos iniciales y los costos recurrentes, incluidas las licencias informáticas necesarias, el mantenimiento y el almacenamiento;
- d) la facilidad de consulta;
- e) la capacidad de las herramientas de búsqueda disponibles;
- f) la facilidad para el desarrollo de herramientas de búsqueda adicionales;
- g) la capacidad para incorporar/enlazar otras fuentes de datos o programas de visualización de datos, por ejemplo el SIG;
- h) la adaptación a la infraestructura/requisitos existentes de tecnología de la información y a las capacidades de personal;
- i) un sistema de metadatos que proporcione información apropiada sobre los datos contenidos en la base; y
- j) la capacidad de acceso en red o a distancia (enlaces con servidores de red y de web).

10.2.2 Métodos de almacenamiento

Cuando se archivan datos digitalmente, una de las consideraciones más importantes es la base de datos utilizada. El término base de datos se utiliza a menudo de manera equivocada, y en el ámbito hidrológico y en otros contextos se utiliza a menudo para referirse al sistema de base de datos propiamente dicho y al programa informático utilizado para la búsqueda, visualización y análisis de los datos.

Ambos son aspectos importantes de cualquier archivo; por esa razón, se abordarán por separado en el presente capítulo.

Una base de datos es sencillamente un sistema de archivado de datos electrónicos. Cualquier conjunto organizado de datos digitales es, de hecho, una base de datos. Varios aspectos importantes de estos conjuntos permiten determinar la base de datos más apropiada para cada caso, en función de los objetivos principales perseguidos por los administradores de los datos, indicados en la sección 10.2.1.

10.2.2.1 Criterios importantes respecto a los sistemas de almacenamiento de datos

Para desarrollar un sistema de almacenamiento de datos deberán tenerse en cuenta varios criterios importantes, entre otros:

- a) la seguridad, que abarca la gestión del acceso y los derechos administrativos de los diversos usuarios;

Por supuesto, cada archivo hidrométrico otorgará diferentes grados de importancia a cada uno de estos aspectos, y también en este caso se describen los casos extremos. Los modernos requisitos de una gran red nacional, como la telecarga automatizada de datos en tiempo real, los enlaces con herramientas de análisis sofisticadas y el acceso multiusuario desde un gran número de organizaciones distribuidas, conllevan a su vez gastos considerables en apoyo técnico, cursos de capacitación para los usuarios y, a menudo, desarrollo de herramientas a medida. La base de datos debe funcionar de manera fiable en una computadora de altas prestaciones, y ha de ser posible hacer automáticamente copias de seguridad en cintas magnéticas conservadas en un lugar seguro a prueba de fuego. En proyectos pequeños, puede ser necesario que la base de datos sea manejada por un solo hidrólogo. En tales casos, las operaciones de carga, corrección y análisis de los datos deberán ser simples y rápidamente asimilables. Podría ser necesario utilizar una base de datos tan pequeña como para ser enviada por correo electrónico a otros usuarios. El archivo hidrométrico de un país pequeño puede contener datos esenciales para el futuro social, medioambiental y financiero del país, aunque a veces podría ser necesario mantenerlo con un presupuesto muy limitado. La obtención de datos es onerosa, y el dinero invertido en sistemas informáticos excesivamente complejos podría impedir que el archivo cumpla su cometido, es decir, obtener y publicar datos hidrométricos de buena calidad. Una base de datos deberá ser también sostenible, es decir, segura, fácil de gestionar con la infraestructura disponible, y provista de las herramientas necesarias.

Las distintas bases de datos (en este caso, los sistemas electrónicos de gestión de datos) se pueden clasificar con arreglo a las categorías siguientes.

10.2.2.2 **Ficheros ASCII simples**

El tipo más simple de base de datos puede consistir en un conjunto de archivos ASCII que contengan datos, indexados en una PC o unidad de red. Se puede utilizar un fichero aparte para almacenar datos de una determinada serie temporal, posiblemente con un directorio aparte para almacenar los datos de cada estación. Este tipo de sistema tiene, entre otras ventajas, un costo no superior al de la computadora que almacena los datos, es muy fácil de instalar con conocimientos elementales o incluso nulos de informática, y los archivos pueden ser fácilmente localizados en un formato de texto que permita leer los datos inmediatamente y almacenar datos de cualquier tipo para consultarlos posteriormente. Sus inconvenientes son: la evidente falta de seguridad del sistema, las limitaciones inherentes de un sistema de almacenamiento manejado por un solo usuario, la ausencia de herramientas de análisis y representación gráfica de los datos, y la dificultad de desarrollar herramientas para trabajar con los datos. Sin embargo, muchas organizaciones mantienen todavía sistemas de este tipo, que pueden considerarse apropiados si se trata de pequeñas empresas que almacenan copias de sus datos archivados cuando no es necesario analizarlos, y la seguridad no plantea problemas, aunque un nivel mínimo de mantenimiento es muy importante.

10.2.2.3 **Formatos personalizados para bases de datos**

Muchos sistemas, en particular los desarrollados antes del auge de la informática a finales de los años noventa, utilizan su propio formato de almacenamiento de datos. Son a menudo formatos muy comprimidos, que permiten almacenar grandes volúmenes de datos en el reducido espacio de los discos utilizados por aquel entonces. Además, los protocolos especiales creados para acceder a los datos en determinados formatos permiten recuperar y guardar datos muy rápidamente, ya que evitan los gastos generales que conlleva un acceso genérico a los datos. Además de estas ventajas, una entidad que compila su propia base de datos adquiere también un acervo considerable de conocimientos y puede atender eficazmente a sus propias necesidades de almacenamiento y visualización de datos, así como de instrumentos de análisis. Los inconvenientes son la imposibilidad de comunicar con otras tecnologías disponibles, o de incorporarlas (una característica de los sistemas más genéricos),

y el costo de mantenimiento de las herramientas desarrolladas a medida que evolucionan los sistemas operativos y plataformas utilizados. Además, para una organización es arriesgado depender solamente de conocimientos internos, ya que pueden surgir dificultades si no dispone de un sistema de transferencia de conocimientos o si, por alguna causa, se perdieran éstos.

10.2.2.4 **Sistemas de gestión de bases de datos relacionales**

Como su propio nombre indica, los sistemas de gestión de bases de datos relacionales no son simples bases de datos. Suelen tener un formato de archivo específico para el almacenamiento de datos (la base de datos propiamente dicha), además de unos protocolos de gestión y de herramientas de acceso informáticas. Las más complejas pueden tener herramientas de búsqueda integral, redacción de informes, trazado de gráficos y publicación. Pueden adquirirse en el mercado varios sistemas de gestión de bases de datos relacionales muy conocidos. Se utilizan mucho en todo el mundo y, por lo tanto, han sido probadas por los usuarios y ofrecen un buen servicio de atención al cliente. Es fácil encontrar profesionales capaces de desarrollar herramientas adicionales. La seguridad, el nivel de apoyo técnico, la disponibilidad de herramientas de búsqueda, el precio, etc., variarán según el tipo de sistema.

10.2.2.5 **Sistemas especializados de bases de datos hidrométricos**

Los sistemas de bases de datos precedentemente descritos son lo que su propio nombre indica: herramientas genéricas para almacenar datos. Han de ser adaptados por el usuario para atender a sus necesidades específicas. Lo más probable es que un sistema especializado de base de datos hidrométricos responda a las necesidades de casi todos los hidrólogos. Se trata esencialmente de programas informáticos comerciales (aunque en algunos casos su instalación puede representar un trabajo considerable), que es posible adquirir u obtener por otros medios. Constan en su mayor parte de un sistema de base de datos, de uno de los tipos anteriormente descritos, adaptado para manejar específicamente tipos habituales de datos hidrológicos. Por ejemplo, tablas de bases de datos y rutinas de acceso que manejan específicamente datos hidrológicos y que almacenan la información descriptiva apropiada y los metadatos creados.

El sistema suele disponer de un programa informático que permite gestionar, corregir y representar gráficamente los datos, y es mucho más fácil de

utilizar que la propia base de datos. Además, muchas de las herramientas de gestión de datos han sido dotadas de capacidad de análisis, por ejemplo, para producir curvas de duración de flujo a partir de datos de flujo, y de métodos estadísticos para ajustar las distribuciones a los picos de crecida. Estos sistemas informáticos de bases de datos son, entre otros, HYDATA, HYMOS, TIDEDA, HYDSYS y WISKI.

Las ventajas y los inconvenientes de estos sistemas se enmarcan en las categorías mencionadas en la sección 10.2.1. Los sistemas generalmente disponibles son de distinta envergadura. Los más pequeños son más baratos y fáciles de instalar, utilizar y mantener. Los más grandes tienden a ser más caros, pero tienen funcionalidades más avanzadas y gestionan bases de datos más grandes con mayor seguridad, aunque a menudo el programa informático requiere el pago de una licencia. La selección de un sistema de base de datos dependerá, pues, de las necesidades del usuario y de su capacidad para adquirirlo y mantenerlo.

10.2.2.6 Competencias para la gestión de bases de datos

Las bases de datos pueden ser gestionadas por una sola persona o por equipos de muchas personas, pero sus procesos suelen requerir conocimientos particulares, que determinan las funciones de la persona que actúa en el proceso. En la tabla I.10.2 se enumeran algunos de estos conocimientos.

10.2.2.7 Resumen

En resumen, existen numerosos tipos de sistemas de almacenamiento de datos digitales. Aunque cualquiera de los sistemas especializados de bases de datos hidrométricos cubriría las necesidades de la mayoría de los hidrólogos, algunos usuarios avanzados podrían necesitar un sistema de base de datos especialmente adaptado. Esencialmente, un

sistema de base de datos es un simple medio para el almacenamiento de datos digitales. Una buena gestión de los archivos solo se conseguirá mediante una gestión adecuada de los datos.

10.2.3 Tipos de datos y de información almacenable

En esta sección se detalla la información almacenable en un archivo hidrológico. Quizá la mejor manera de acometer la organización de un archivo sería imaginarse un usuario sin ningún conocimiento previo de condiciones meteorológicas, dimensiones de ríos, características de cuenca, red de estaciones de aforo, usos del agua en las cuencas, o volúmenes de datos. Tendrá que ser posible conocer rápidamente todo el contenido del archivo y recuperar fácilmente los datos necesarios. Los usuarios del archivo deberán ser capaces de valorar todos los cambios introducidos en los datos y de obtener rápidamente información sobre los datos disponibles, así como estadísticas resumidas y conjuntos de datos completos. De ese modo podrá, sin demasiado esfuerzo, comenzar a trabajar con un conjunto de datos en cualquier etapa de su proceso de gestión. Además, el sistema de archivado simplificará y facilitará el trabajo de documentación. El administrador del archivo deberá poder producir de manera simple y fácil conjuntos de datos adicionales con el fin de introducirlos en modelos o en procesos de datos adicionales, distribuirlos a otros usuarios de los datos o crear publicaciones, por ejemplo anuarios.

10.2.3.1 Metadatos de archivos

Cuando el usuario estudia un archivo hidrológico, el primer nivel de datos que debe ver es el de los datos que describen el propio archivo. Éstos son, en realidad, metadatos, es decir, información sobre el propio archivo, que debería ser publicada por el administrador de los datos con el fin de ofrecer

Tabla I.10.2. Competencias necesarias para la gestión de una base de datos

<i>Función</i>	<i>Descripción</i>
Introducción de datos	No son necesarios grandes conocimientos de hidrología o de tecnología de información, aunque los datos pueden tener que descargarse de un registrador cronológico o extraerse de formatos de archivos diferentes
Hidrólogo	Para las tareas de validación es necesario un conocimiento experto de hidrología y de los regímenes hidrológicos locales. Para las tareas de análisis son necesarios conocimientos expertos en hidrología
Gestor de archivos	Para la gestión y difusión general de los archivos son necesarios conocimientos de hidrología. Para la integración de los datos de archivo en otros procesos son necesarios conocimientos prácticos de hidrología y tecnología de información

información sobre el archivo. Este tipo de datos podría adoptar la forma descrita en la tabla I.10.3.

Los metadatos archivados pueden ser obtenidos de un sistema informatizado complejo, posiblemente provisto de una interfaz SIG para permitir el acceso a los datos y a resúmenes automáticamente actualizados de la disponibilidad de datos que el usuario podría consultar, o pueden ser tan simples como una carpeta de documentos, en cuyo caso la responsabilidad corresponderá al administrador del archivo. En este último caso, la carpeta se actualizará regularmente cuando se añadan nuevas estaciones o nuevos datos.

10.2.3.2 Metadatos de estaciones

Cuando los usuarios del archivo conozcan el acervo de datos, necesitarán información adicional. Los datos descriptivos de las estaciones son importantes para indicar el contexto en el que opera la estación. Pueden constituir también un recurso común que ayude a sus usuarios a conocer, por ejemplo, en qué medida influyen en los datos los aparatos de medición utilizados, o la configuración de la estación, y que ayude al personal administrativo de la estación a almacenar información sobre la ubicación de la estación, los accesos, y los datos y direcciones postales de los operarios locales. La mayoría de los datos indicados en la tabla I.10.4 tienen relación con estaciones meteorológicas, estaciones de aforo u otros emplazamientos de medición, aunque algunos campos son específicos de las estaciones fluviales de medición hidrológica.

En el sitio web <http://www.wmo.int/pages/prog/www/WDM/Metadata/documents.html> se encontrará información en inglés sobre el desarrollo de la norma de metadatos básicos de la OMM.

En el Informe 31 del Centro mundial de datos de escorrentía se encontrará un resumen de la situación actual en relación con los sistemas de metadatos hidrológicos (Maurer, 2004).

En los sitios web siguientes se encontrará información adicional sobre las normas aplicables a las bases de metadatos:

- USGS – Federal Geographic Data Committee, “Content Standard for Digital Geospatial Metadata”: <http://www.fgdc.gov/metadata/>;
- Dublin Core Metadata Element Set, Version 1.1: <http://dublincore.org/documents/dces/>;
- ISO 8459-5 *Information and documentation – Bibliographic data element directory – Part 5: Data elements for the exchange of cataloguing and metadata*: <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=27176&ICS1=35 &ICS2=240&ICS3=30>.

En las direcciones siguientes se encontrarán ejemplos de metadatos para sistemas hidrológicos:

- a nivel mundial: <http://www.wsag.unh.edu/metadata/>;
- a nivel nacional: http://www.epa.gov/Region8/gis/data/r8_hyl.html;
- a nivel de circunscripción estatal: <http://www.isgs.uiuc.edu/nsdihome/webdocs/st-hydro.html>; <http://www.wy.blmgov/gis/hydrologygis-meta.html>.

10.2.3.3 Series cronológicas de datos

La mayoría de los datos utilizados en hidrología están agrupados en series cronológicas, es decir, son mediciones de una variable durante cierto tiempo en un determinado lugar: precipitación, flujo fluvial, nivel de agua, capacidad de embalse, nivel de agua en pozos, humedad del suelo o pH.

Tabla I.10.3. Descripción de los datos archivados

<i>Tipo de datos</i>	<i>Descripción</i>	<i>Ejemplos</i>
Descripción del archivo	Texto sucinto descriptivo de los antecedentes y finalidades del proyecto de control de datos	Nombre y descripción del proyecto, fecha de comienzo del proyecto/archivo, finalidades del proyecto, resumen de las vías de difusión
Mapas geográficos	Mapas indicativos del contexto físico de los datos archivados	Límites de la cuenca, ubicación de la estación de aforo/meteorológica, y otras ubicaciones de medición de datos, red fluvial, lagos y otras particularidades importantes
Resúmenes de datos	Lista de conjuntos de datos y de disponibilidades	Resumen, por tipos de datos, de los datos almacenados en la base de datos, con referencia a la ubicación de las mediciones, más otros datos adicionales (por ejemplo, datos espaciales o de un SIG derivados de otras fuentes), además de un resumen, para cada conjunto de datos, de su disponibilidad en términos de fechas y a una escala apropiada

En una estación (o ubicación geográfica) se miden a menudo numerosas series cronológicas de datos, cada una de las cuales puede tener características diferentes. Cada estación almacenará un resumen de las series cronológicas medidas en el emplazamiento (historial del conjunto de datos, en la tabla I.10.4), y anotará los atributos correspondientes a cada una de las series cronológicas (tabla I.10.5).

10.2.3.4 Datos en tiempo real

Los datos obtenidos mediante telemetría que sean necesarios para aplicaciones en tiempo real, por ejemplo, para predicción de crecidas, utilización de embalses o control de bajos niveles de flujo por razones ecológicas, podrían ser archivados y consultados en un sistema diferente al de los datos obtenidos para la observación periódica

Tabla I.10.4. Ejemplos de metadatos de estación

<i>Metadato de referencia</i>	<i>Descripción</i>	<i>Ejemplos</i>
Identificación	Información identificativa actual de la estación y resumen de los fines a los que se destina	Nombre(s) de la estación, número(s) de estación, nombre de la cuenca, nombre de la masa de agua, nombre del área hidrométrica, elevación, superficie de la cuenca, finalidad primaria, finalidad secundaria, método de medición primario (por ejemplo, tipo de presa), método de medición del flujo elevado, descripción general de la estación
Ubicación	Información sobre la posición geográfica de la estación	Latitud/longitud (o posición en un sistema de coordenadas local), núcleo urbano/rasgo físico más cercano, ubicación y altura de referencia, información sobre el propietario, carreteras, accesibilidad, tiempo de acceso apropiado, información sobre los accesos en períodos de crecida, etc.
Operador	Información sobre la organización que opera la estación, si ésta es operada por otra organización, por ejemplo de ámbito regional	Nombre del operador, datos de contacto, responsabilidades, etc.
Observador	Información sobre el personal que efectúa las mediciones en la estación	Nombre del observador, datos de contacto, responsabilidades, fecha de comienzo, frecuencia de las visitas, método y periodicidad de la presentación de informes
Historial de la estación	Descripción del historial de la estación, indicando los cambios que pudieran afectar a los datos medidos	Fecha de apertura, fecha de cierre (para las estaciones que han cerrado), historial de la ubicación, historial del operador, historial del equipo, historial de los elementos de referencia
Equipo/telemetría	Información descriptiva de los registradores cronológicos de datos o sistemas de telemetría automatizados que se utilicen en la estación	Nombre de los sistemas, fabricante, finalidad, referencia que remita a las publicaciones apropiadas, fecha de instalación, alturas de antena, etc., intervalo y frecuencia de presentación de informes, parámetros notificados, información descriptiva adicional
Estadísticas	Estadísticas resumidas de los datos en la estación	Estadísticas, valores, período de validez de las estadísticas, cálculo de fechas, etc.
Gráficas	Fotografías de la estación y del área circundante	Fotografía, descripción, fecha, referencias a archivos de imagen digitales, etc.
Historial del conjunto de datos	Información descriptiva de los conjuntos de datos producidos para la estación	Parámetros medidos, series derivadas, trayectoria del flujo para los datos medidos en la estación, resumen de la disponibilidad de datos
Historial de mediciones y de curvas de gasto	Información descriptiva de las mediciones instantáneas del caudal y de la obtención de la ecuación de curvas de gasto; los datos de las mediciones deberían almacenarse en una base de datos	Descripción de la sección o secciones fluviales utilizadas para las mediciones, historial de las modificaciones en la sección por efecto de desplazamiento de ésta o de erosión, etc., dibujos de la sección, descripción de las cuestiones a considerar

o la evaluación amplia de los recursos hídricos. Por lo general, este tipo de datos han de ser sometidos a algún proceso de validación relativamente simple antes de ser archivados para su introducción en modelos utilizables en tiempo real. La validación puede consistir simplemente en comprobar que cada valor de datos recibido no rebasa los límites predeterminados para la estación, y que la diferencia respecto de valores precedentes no es excesiva. Así, los datos de precipitación por períodos de 15 minutos deberán ser siempre un número positivo, aunque inferior al valor máximo registrado en la región considerada, posiblemente con un margen del 10 por ciento. Los datos de nivel fluvial deberán ser también superiores al nivel de lecho o al nivel de coronación del vertedero de medición, y suele ser posible fijar un valor máximo apropiado. Además, a partir del análisis de las fuertes crecidas anteriores puede establecerse una tasa máxima de elevación apropiada para cualquiera de los períodos de 15 minutos. Cuando los datos rebasen esos límites, seguirá siendo necesario almacenarlos en el archivo de datos en bruto, marcándolos como

datos sospechosos y asegurándose de que los operadores del modelo visualizarán un mensaje de aviso.

Cuando se identifiquen datos dudosos, todos los modelos de predicción o de ayuda a las decisiones en tiempo real dispondrán de varias opciones:

- aceptar los datos dudosos, y utilizar el modelo normalmente, aunque ésta no suele ser una opción razonable;
- utilizar el modelo considerando los datos dudosos como datos faltantes, es decir, suponer que no ha habido nuevas precipitaciones durante el período en cuestión o que no se dispone de datos de nivel fluvial y de flujo observados que permitan comprobar las predicciones de flujo; o
- sustituir los datos faltantes por algún tipo de datos de reserva. Así, los niveles fluviales faltantes podrán ser extrapolados a partir de valores anteriores, anotando los datos de precipitación en referencia a otros medidores en funcionamiento o presuponiendo unos valores estacionales medios.

Tabla I.10.5. Características de las series cronológicas de datos

<i>Campo de datos</i>	<i>Descripción</i>
Nombre	Idealmente, el nombre asignado a una serie cronológica debería permitir reconocer al instante el contenido almacenado; por ejemplo, flujo medio diario o flujo total mensual en lugar de serie de flujo 1 o serie de flujo 2
Tipo de serie cronológica	Datos objeto de medición; por ejemplo, precipitación de lluvia, flujo o nivel de agua
Estadísticas de medición	Indica el proceso de derivación de los datos, o las estadísticas que se almacenan: valores medios, instantáneos, totales, máximos, etc.
Unidad	Indica la unidad en que se almacenan los datos
Intervalo	Frecuencia con que se efectúan las mediciones, o período a lo largo del cual se calculan las estadísticas, por ejemplo, diario, mensual, cada 15 minutos. Se toman también en cuenta los datos registrados de manera irregular, frecuentemente denotados como series cronológicas instantáneas
Período de registro	Fecha de comienzo (y, en ocasiones, de finalización) de la serie de datos
Límites estadísticos	Se suele recomendar una estimación inicial de los valores de datos máximos y mínimos de una serie antes de comenzar la medición, con el fin de validar los datos. Este proceder es particularmente útil cuando los métodos de validación automatizada permiten seleccionar valores que rebasan el intervalo recomendado. Tras la medición de los valores extremos, podrán reajustarse con mayor exactitud los valores límite. Podría ser también útil establecer un nivel de aumento o disminución máximo en el transcurso de la serie cuando los métodos utilizados para la validación de los datos pueden beneficiarse de tales estadísticas, al igual que la determinación de unos límites para la obtención de estadísticas más complejas. Éstas deberían ser únicamente orientativas, ya que los datos externos que sobrepasen tales límites podrían ser válidos y no deberían ser excluidos
Nivel adicional de series cronológicas de datos	En este nivel se pueden almacenar también otros tipos de información aplicable a una serie cronológica, como los días de las mediciones hidrológicas (indicando los períodos en que habría que calcular los promedios y otros valores), o los valores de referencia, si ello fuera apropiado respecto de la totalidad del período de datos

El tratamiento de los datos faltantes dependerá de la situación y de los requisitos de modelización. El tema de la modelización se aborda en el capítulo 6 del volumen II.

10.2.3.5 Datos espaciales

Casi todos los datos examinados anteriormente son metadatos descriptivos o series cronológicas de datos de variables. Hay otro tipo de datos que se examinará en la presente sección. Los datos espaciales tienen un componente geográfico no desdeñable. Suelen ser mapas de emplazamientos de estaciones de aforo, modelos de elevación digitales o isoyetas de precipitación de lluvia. Los datos espaciales se pueden visualizar mediante un SIG, y suelen utilizarse para

integrar conjuntos de datos hidrológicos y espaciales.

En un SIG, las características geográficas pueden representarse de diversas formas (figura I.10.2):

- polígonos: los datos son demarcaciones de áreas, por ejemplo, países o cuencas;
- líneas: los datos aparecen en forma de líneas, con sus atributos correspondientes, por ejemplo ríos;
- puntos: los datos son puntos individuales, por ejemplo, estaciones de aforo fluvial o pluviómetros; y
- retículas: la región aparece dividida en cuadrículas y la variable (por ejemplo, precipitación) correspondiente a la cuadrícula se almacena junto con otras variables.

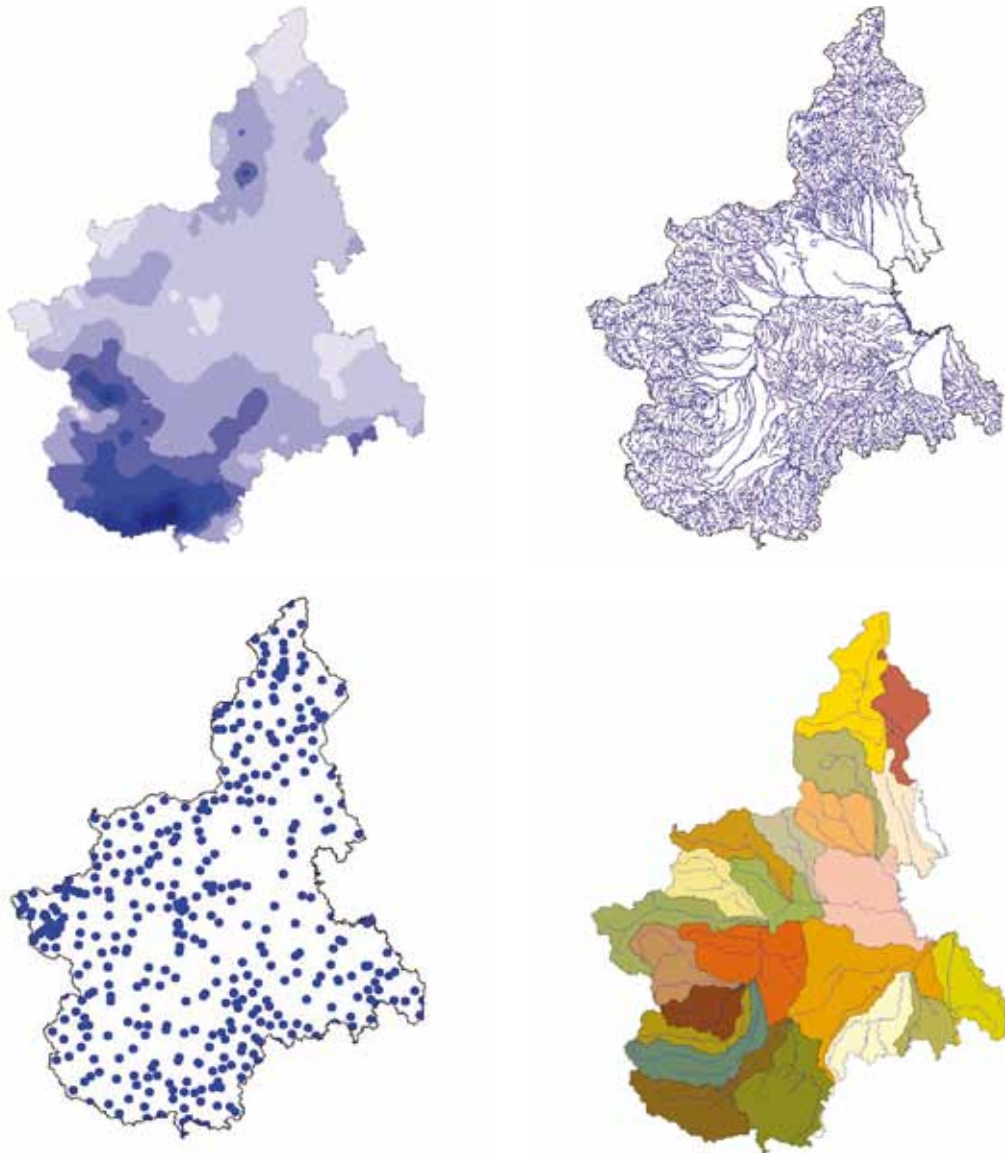


Figura I.10.2. Ejemplos (comenzando por la figura superior izquierda, en el sentido de las agujas del reloj) de datos reticulados (lluvia), lineales (ríos), puntuales (estaciones de aforo) y poligonales (cuencas)

Las características de estos accidentes geográficos se denominan atributos, por ejemplo, cada polígono de una extensión geológica puede contener atributos tales como su litología o el tipo de acuífero.

A los efectos de la presente guía, los datos espaciales hidrológicos pueden clasificarse con arreglo a dos categorías simples que se describen a continuación.

Mapas físicos

Los mapas físicos son valiosísimos en los estudios hidrológicos, y siguen siendo la fuente principal de datos espaciales en muchos países. Pueden ser mapas especializados, que indiquen la extensión de suelo, la geología o la precipitación, o mapas nacionales que muestren múltiples aspectos, tales como ciudades, carreteras, contornos o ríos. Los mapas físicos serán un elemento central del archivo hidrológico, y constituirán un primer punto de referencia muy útil que proporcionará valiosos datos meteorológicos y acerca de la estación, tanto contextuales como directos. Serán almacenados en consonancia, preferiblemente en cofres, estanterías o monturas apropiadas. El archivo cartográfico estará adecuadamente documentado, en particular con los datos siguientes:

- a) números de referencia y origen/fuente del mapa;
- b) título y descripción del mapa;
- c) escala;
- d) proyección;
- e) número y nombre del mapa en su serie; y
- f) referencias.

La comunicación de datos cartográficos físicos puede ser difícil, ya que frecuentemente están sujetos a derechos de autor. No obstante, si fuera necesario dar a conocer fotocopias completas o parciales de un mapa, se incorporará la información precedente de modo que permita determinar el origen de los mapas y leyendas y facilitar su interpretación.

Si se creasen mapas físicos en un archivo, por ejemplo, de escorrentía obtenidos de datos de flujo en estaciones de medición, se indicará de manera apropiada la existencia de esos mapas, junto con los detalles correspondientes. Por ejemplo, en libros de datos hidrológicos, o por conducto de organismos cartográficos nacionales o regionales.

Datos digitales

En los 10 últimos años se ha generalizado la utilización de mapas digitales en sustitución de los mapas físicos. Los avances tecnológicos han permitido digitalizar los mapas y utilizarlos en un SIG. En estos sistemas es mucho más fácil manipular e

integrar mapas. También es más fácil extraer información y difundir los cambios.

Muchos mapas digitales son simples versiones digitalizadas de mapas físicos. Por ejemplo, los contornos de un mapa de uso general normal se pueden digitalizar en términos de demarcaciones lineales, o un mapa de tipos de suelo, en términos de demarcaciones poligonales. Como en todos los procesos de gestión de datos, se documentará cuidadosamente la procedencia de los datos, así como las correcciones formales introducidas, con el fin de que los usuarios de los datos resultantes conozcan su origen.

Es posible también crear mapas digitales. Por ejemplo, se puede crear una extensión cuadrículada de datos de precipitación a partir de datos pluviométricos de fuente puntual, utilizando para ello diversos procesos. Las líneas de contorno digitalizadas con un alto grado de exactitud pueden ser extrapoladas, con objeto de crear una retícula de alturas topográficas para un modelo digital de elevaciones. Mediante el uso de contornos o de modelos de elevación digital, pueden añadirse manualmente límites de una cuenca a una nueva extensión lineal. Cuando los mapas obtenidos hayan sido almacenados en un archivo hidrológico, se adoptarán las precauciones indicadas en la sección 10.2.1 por lo que respecta a su reproductibilidad. Los mapas obtenidos se archivarán conjuntamente con los metadatos que describan el proceso utilizado para crearlos. Los conjuntos de datos intermedios creados se archivarán como corresponda, siempre que sean importantes y útiles.

10.2.3.6 Consideraciones de gestión

En la gestión de datos e información hidrológicos será importante incluir los elementos siguientes:

- a) banderines de validación o de control de calidad (secciones 9.8 y 9.9);
- b) comentarios textuales de usuarios/procesadores de datos (secciones 9.7 y 9.8); e
- c) historial de auditoría, es decir, información sobre la introducción de los datos en la base de datos y sobre las modificaciones o ajustes subsiguientemente introducidos (capítulo 9).

Estos tipos de datos deberán también conservarse y mantenerse fácilmente accesibles.

10.2.3.7 Control del flujo de datos

Ya se ha mencionado la importancia de un control apropiado de los conjuntos de datos recibidos en las

operaciones de introducción de datos. Será también esencial conocer el estado de todos los conjuntos de datos en las diversas etapas de evaluación y actualización. Esto será particularmente importante cuando se hayan obtenido datos dudosos y el personal hidrológico encargado del control de calidad espere algún tipo de respuesta.

Inicialmente, todo el proceso de control puede ser manual, pero en último término será posible automatizar algunas funciones, en el marco de las actividades generales de tratamiento informatizado de datos.

La automatización permite controlar en régimen habitual el estado de los lotes de datos, así como los resúmenes de validación y la disposición física de los datos en el sistema, por ejemplo, mediante sus números de volumen en cinta o disco y mediante los nombres de los conjuntos de datos. Este tipo de control será esencial cuando se manejen grandes cantidades de datos.

El personal de control de los datos asumirá las responsabilidades siguientes:

- a) registro cronológico de los lotes de datos entrantes y envío de éstos al correspondiente sistema de introducción de datos;
- b) control y registro cronológico del estado de introducción de los datos y envío de datos para su validación inicial y tratamiento;
- c) envío de informes de validación al personal hidrológico apropiado y recepción de datos corregidos;
- d) repetición de los pasos a) a c) hasta que todos los lotes de datos hayan sido aceptados para su actualización; y
- e) envío de estadísticas resumidas mensuales y anuales a los organismos y al personal correspondientes.

La naturaleza exacta de las tareas dependerá del grado de acceso de los usuarios a efectos de corrección formal de los datos. En los sistemas en línea cuyos usuarios adopten sus propias medidas de control de la calidad, las responsabilidades de la central serán menores. Los usuarios, sin embargo, deberán disponer de algún medio para indicar que han realizado el control de calidad y que los conjuntos de datos están preparados para su ulterior tratamiento.

10.2.3.8 Procedimientos de actualización

La actualización de las bases de datos hidrológicos con fines de archivado consta, como mínimo, de dos etapas. El proceso puede verse en la figura I.9.2. La primera etapa es el ciclo de actualizaciones

mensuales correspondiente a un período de informe típico. El número de sesiones de computadora en que se subdividan las cuatro actividades de la primera etapa dependerá del usuario y de los recursos físicos del sistema. Cuando la mayoría de los ficheros estén archivados en cinta, será probablemente imposible realizar el tratamiento mensual completo mediante un solo programa, ya que se necesitarán demasiados lectores de cinta. Puede adoptarse también como norma no computar valores derivados, por ejemplo, flujos o evapotranspiración potencial, hasta que todos los valores básicos hayan sido comprobados manualmente.

Para el usuario, los principales resultados de la primera fase de actualización serán los informes resumidos mensuales. A efectos de gestión de la base de datos, los resultados más importantes serán los ficheros de trabajo anuales actualizados. Si el sistema de esta primera fase maneja únicamente datos en bloques mensuales, podría ser necesario mantener archivos de datos incompletos. Esto se debe a la utilización de registradores compatibles con computadora, cuyo soporte de registro cambia por lo general a intervalos irregulares. Así, mientras se procesa el mes 1, el soporte de registro puede contener varios días del mes 2. En tal caso, los datos del mes 2 se guardarán en un archivo temporal hasta que los datos complementarios estén disponibles durante el mes 3. Seguidamente se repite el ciclo, y se genera un archivo completo del mes 2, más un nuevo archivo incompleto del mes 3. Este problema suele aparecer rara vez en las estaciones manuales informantes o telemétricas. Cuando el soporte compatible con computadora requiera un tratamiento previo, existe la posibilidad de realizar la subdivisión y subsiguiente agregación de meses en la (micro) computadora de tratamiento antes de someter los datos a la máquina de tratamiento principal.

Después de las comprobaciones de validación (y, en su caso, del tratamiento primario) (capítulo 9), los lotes de datos mensuales serán agregados al archivo de datos anual en curso. Los datos que no pasen satisfactoriamente las comprobaciones de validación serán examinados manualmente y, si se observasen errores, se adoptarán las medidas pertinentes, como puede verse en la figura I.9.2.

Para conseguir una productividad de datos adecuada, suele ser necesario comenzar el tratamiento de cada lote de datos mensuales a partir del décimo o el decimoquinto día del mes siguiente. Si el tratamiento no comenzase en ese momento, podrían sobrecargarse los procesos de tratamiento, introducción y proceso de datos para la actualización anual de los archivos.

El ciclo de actualización anual tiene por objeto añadir el archivo de trabajo anual a la base de datos históricos. Esta transferencia entraña un cambio del estado de los datos, que dejan de ser un conjunto de datos de trabajo para convertirse en un conjunto de referencias hidrológicas de calidad controlada. Así, convendrá asegurarse de que se han resuelto todas las dudas posibles en relación con los datos antes de emprender la actualización anual. Los resultados de la etapa de tratamiento anual pueden utilizarse en anuarios hidrológicos.

10.2.3.9 Compresión y exactitud

Al actualizar una base de datos será importante comprimir los datos para aprovechar de manera óptima el espacio de almacenamiento. La técnica de compresión aparece descrita en la publicación *Guidelines for Computerized Data Processing in Operational Hydrology and Land and Water Management* (WMO-No. 634). Las técnicas de compresión dependen, sin embargo, del tipo de aparato utilizado, y en otros sistemas de bases de datos hidrológicos se utilizan las técnicas que se indican a continuación:

- a) se almacenan los datos en números enteros y a una escala apropiada, en función de los resultados deseados. Así, por ejemplo, una precipitación diaria medida con una precisión de 0,1 mm puede almacenarse en décimas de milímetro (expresadas en forma entera), dividiendo el valor por 10 al producir el resultado. El espacio de almacenamiento se reduce así a la mitad. Un entero normal ocupa 2 bytes de espacio de almacenamiento, frente a los 4 bytes ocupados por un número real (decimal);
- b) se utilizan ficheros de datos (binarios) no formateados en lugar de ficheros ASCII normales. Además de ocupar menos espacio, los datos binarios son almacenados y recuperados con mayor rapidez;
- c) se utiliza un contador para los valores constantes repetidos. Así, un período de 10 días sin precipitación de lluvia no se almacena como un conjunto de 10 ceros, sino como un factor de repetición de 10 seguido del valor cero;
- d) se eliminan completamente los datos redundantes. Éstos proceden de registros sobreabundantes de fenómenos hidrológicos efectuados por ciertos tipos de instrumentos, en particular los registradores de intervalo fijo. Por ejemplo, en la secuencia 40, 50, 60, es evidente que el valor central se puede obtener por interpolación de los valores adyacentes. Así pues, es posible desarrollar un programa que analice los datos y suprima los valores interpolables linealmente con arreglo a un margen de tolerancia determinado. Esta técnica reduce notablemente el espacio de

almacenamiento y no entraña pérdidas significativas de la información contenida en los datos. En Nueva Zelandia, la utilización del sistema TIDEDA (componente G06.2.01 del HOMS) ha permitido reducir entre 2 y 12 veces el espacio de almacenamiento. Este método es una versión más perfeccionada del descrito en el apartado c); y

- e) se utilizan valores de datos relativos en lugar de valores absolutos. Así, por ejemplo, el nivel de agua en un pozo de sondeo puede ser expresado en valores absolutos de elevación o, para ahorrar espacio, remitiéndose a un nivel de referencia local o a un nivel de agua promedio. Solo será necesario almacenar la diferencia respecto al valor de datos anterior. Estos métodos permiten obtener cifras más pequeñas, que pueden ser almacenadas en espacios también más pequeños. Será necesario encontrar un cierto equilibrio entre los niveles de compresión aplicados. La eficacia del almacenamiento aumenta a expensas de las rutinas de compresión y expansión que se ejecutan en cada operación de almacenamiento o recuperación de datos. En cada instalación, el nivel correcto de compresión de datos reflejará las limitaciones relativas del espacio de almacenamiento y de la capacidad de cálculo, y de la eficacia en el desarrollo del programa informático. Por lo que respecta a la exactitud de los datos almacenados, será excepcional obtener una exactitud superior a 1 por 1 000 en la observación de datos hidrológicos. Por ello, en muchos casos las bases de datos hidrológicos almacenan los datos con una precisión de tres o cuatro cifras significativas. Así, un flujo expresado en la forma $234,56 \text{ m}^3 \text{ s}^{-1}$ puede ser almacenado como 235. Este método se utiliza también para ahorrar espacio de almacenamiento de datos.

10.2.3.10 Organización de los archivos físicos

La organización secuencial de los archivos es simple, sirve para todo tipo de soporte de almacenamiento, y está adaptada a las series cronológicas de datos, que suelen ser introducidas y consultadas secuencialmente. Los ficheros secuenciales indexados son muy convenientes para el almacenamiento de la mayoría de datos hidrológicos, ya que conservan el carácter secuencial inherente de los datos en el medio de almacenamiento y permiten seguir accediendo directamente a los distintos registros o grupos de registros.

El acceso aleatorio, como la indexación secuencial, sirve únicamente para ficheros almacenados en disco o disquete, pero conlleva una carga más elevada en términos de volúmenes de almacenamiento.

Para los registros individuales, el acceso aleatorio es más directo y rápido. Cuando se utilicen referencias cruzadas (punteros), los datos contenidos en archivos de acceso aleatorio pueden ser relacionados en términos complejos y eficaces.

Cuando se desarrolle una base de datos hidrológicos para que admita la manipulación de datos en línea (interactiva), los ficheros estarán disponibles en disco, y será posible utilizar archivos secuenciales indexados o de acceso aleatorio. De hecho, su utilización es probablemente esencial para obtener tiempos de respuesta aceptables cuando se manejan grandes volúmenes de datos.

Cuando el acceso en línea a los datos no sea prioritario, podría merecer la pena conservar series cronológicas de datos de una sola variable, por ejemplo, de niveles de agua o de valores de precipitación, en ficheros secuenciales, ya que su búsqueda suele tener por objeto abstraer secuencias temporales de datos. En archivos de series cronológicas con más de una variable, la organización secuencial indexada o de acceso aleatorio tiene varias ventajas.

Si una variable se mide solamente en un pequeño número de estaciones, para localizar los valores en un archivo secuencial habrá que consultar los datos de todas las estaciones. En algunos tipos de archivos de acceso aleatorio será posible almacenar, junto con cada valor de una variable, un puntero que señale la ubicación del registro de estación siguiente que contenga un valor de esa misma variable. Será posible de ese modo acceder directamente a esa ubicación. Esta técnica es interesante con datos de calidad del agua, ya que las variables observadas varían considerablemente de una estación a otra y, en una misma estación, según la hora.

Los datos almacenados en cinta magnética, que es el formato más común de los grandes archivos de bases de datos, se organizarán secuencialmente. Sin embargo, al transferir los archivos de cinta a disco será posible utilizar cualquiera de los métodos de acceso anteriormente descritos. Sea cual sea el método que se utilice, se recomienda que los grandes archivos de bases de datos no estén formateados (es decir, estén en formato binario).

Algunos sistemas de bases de datos utilizan una combinación de técnicas para maximizar la eficacia de almacenamiento y de recuperación, y almacenan grandes grupos de datos secuenciales en registros separados de acceso aleatorio o en archivos secuenciales indexados. Con este método, los datos diarios, o incluso horarios, tomados en una

estación durante un año se pueden almacenar como un solo registro físico en un archivo secuencial indexado o de acceso aleatorio. Para recuperar los datos de un mes determinado, será posible acceder directamente en el disco al registro anual de la estación correspondiente. Este registro se transfiere seguidamente a una memoria tampón interna, en la cual se podrán leer rápidamente los datos correspondientes al mes buscado. Quedan por mencionar los sistemas de gestión de bases de datos (SGBD), que utilizan sistemáticamente archivos de acceso aleatorio. Se recomienda utilizarlos con precaución, a menos que los formatos exactos de introducción y recuperación sean conocidos (y relativamente fijos) y se disponga de suficiente apoyo informático. Se recomienda adoptar para los SGBD un criterio evolutivo.

Muchos organismos están evaluando actualmente la utilización de sistemas de gestión de bases de datos relacionales para el almacenamiento conjunto de datos y otros tipos de información. Convendrá vigilar de cerca los avances en este terreno.

10.2.3.11 Organización lógica de archivos

La organización lógica de los datos abarca dos vertientes: los grupos principales, que determinan el número de archivos, y los conjuntos de valores de variables incluidos en los registros de cada archivo.

Una base de datos hidrológicos completa contiene los grupos de archivos siguientes:

- a) archivos de referencia del sistema, y en particular listas de claves (archivo de diccionario) utilizadas para comprobar los datos introducidos, para cifrar los datos a almacenar y para descifrar los datos antes de producirlos. Si se utilizara algún tipo de cifrado de datos espacial, se necesitarán también archivos de referencia hidrológicos y/o geográficos;
- b) archivos de descripción de estación, que pueden ser desde archivos simples que relacionan números de estación con nombre, tipo, ubicación e instrumentación de una estación, hasta archivos detallados, por ejemplo, datos completos de archivos cronológicos de pozos o perforaciones;
- c) archivos de calibración, que contienen la información general necesaria para calcular variables derivadas, normalmente estación por estación. Pertenecen a este tipo las curvas de gasto de estaciones de medición del flujo fluvial o los coeficientes de calibración de sensores climatológicos y de calidad del agua. Algunos datos son independientes de la estación, por ejemplo,

- los coeficientes de calibración de molinetes o las tablas de referencia de radiación incidente teórica y horas de insolación; y
- d) archivos de series cronológicas, que contienen series de observaciones efectuadas en estaciones hidrológicas. Las series pueden ser de variable única o de múltiples variables, y pueden haber sido observadas a intervalos regulares o irregulares.

La relación entre esos grupos de archivos se indica en la figura I.10.3.

Desde un punto de vista organizativo, es posible combinar toda la información de los apartados b) y c) en archivos comunes, o separar cada tipo en archivos corrientes e históricos. La ventaja radica en que será posible dar un formato y un tamaño estándar a los archivos corrientes. La decisión dependerá principalmente de la cantidad de datos descriptivos que se conserven en los archivos informáticos, en comparación con los conservados en archivos manuales.

Será útil considerar las diversas alternativas disponibles para el almacenamiento de distintos tipos de series cronológicas de datos en un mismo archivo físico.

En el nivel más simple, se asignará a cada estación su propio archivo, con los datos ordenados cronológicamente. Esta técnica será apropiada para

pequeños conjuntos de datos, o para la conservación de datos archivados en cinta. Sin embargo, las redes hidrológicas pueden contener varios millares de estaciones diferentes, en cuyo caso el sistema se hace extremadamente difícil de gestionar y de mantener para grandes cantidades de archivos.

En el nivel siguiente, que se utiliza en la mayoría de los sistemas de bases de datos hidrológicos, los archivos contienen datos diferentes de muchas estaciones. Éstos pueden consistir en datos hidrológicos, por ejemplo valores de caudal diarios, o series cronológicas mixtas, por ejemplo varias variables medidas a intervalos fijos. En el primer caso, un archivo de caudales diarios contendría todos los datos de caudal diario de la red hidrológica completa. Si estuviera organizado secuencialmente, el archivo estaría ordenado por estaciones y, en cada estación, de manera cronológica. En el segundo caso, figurarían todos los datos diarios, independientemente del tipo hidrológico, y el archivo estaría ordenado por tipo y número de estación. Ambos casos están presentes en el sistema WATSTORE (Kilpatrick, 1981), que consta de cinco grandes archivos. Uno de ellos contiene los datos de encabezamiento (descripción) de la estación. De los otros cuatro, tres están agrupados por tipos hidrológicos (calidad del agua, caudal máximo, e inventario de los emplazamientos de medición de agua subterránea), y el cuarto, agrupado en una serie cronológica, es el archivo de valores diarios. Este último contiene datos observados en régimen diario o continuo, y ha sido

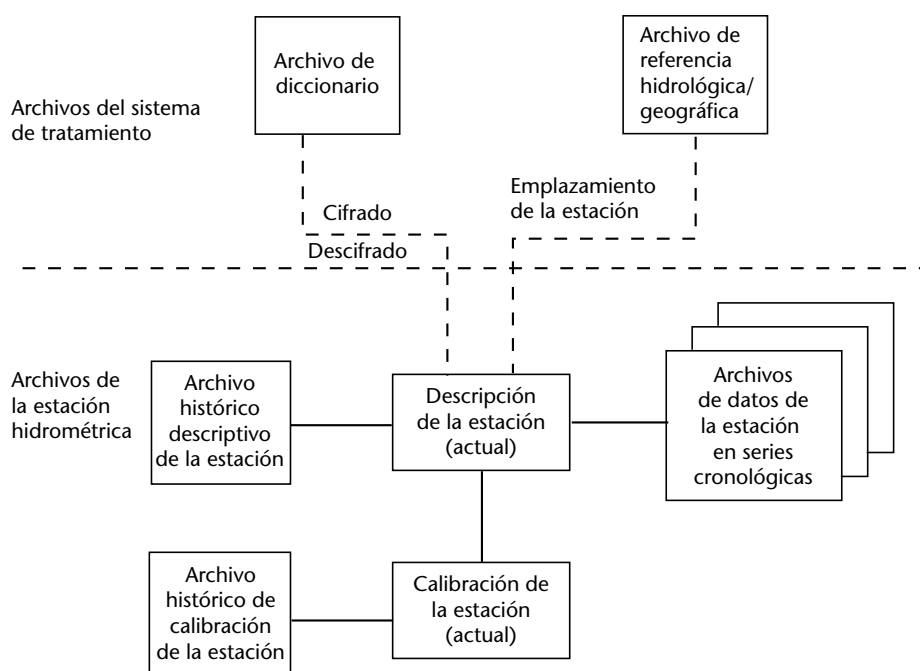


Figura I.10.3. Relaciones entre los archivos de datos de estación

reducido numéricamente a valores diarios. Pueden almacenarse también mediciones instantáneas a intervalos fijos, o valores y estadísticas diarios medios, por ejemplo en forma de máximos y mínimos diarios. En 1981, el archivo contenía 190 millones de valores diarios, consistentes en datos de caudal, nivel de agua, volumen de embalse, temperatura de agua, conductividad eléctrica, concentración de sedimentos, caudal de sedimentos y nivel de agua subterránea.

En el nivel de integración más elevado (que no esté basado en la utilización de un SGBD) se encuentran los sistemas que manejan todo tipo de series cronológicas en un solo formato de almacenamiento, y que almacenan todos los datos cronológicos en un solo archivo físico. Esta metodología, utilizada en el sistema TIDEDA de Nueva Zelanda, simplifica mucho el desarrollo de programas de gestión y recuperación de datos, ya que utiliza un formato de almacenamiento estándar. Otros sistemas similares de proceso y almacenamiento de datos que también son componentes del HOMS son HYDATA, en Reino Unido, e HYDSYS, en Australia. En *Guidelines for Computerized Data Processing in Operational Hydrology and Land and Water Management* (WMO-No. 634) se ofrecen detalles sobre los métodos utilizados para el manejo de datos en estos sistemas de proceso y almacenamiento.

10.3 RECUPERACIÓN DE DATOS

10.3.1 Herramientas de análisis de datos

Las herramientas de análisis de datos pueden consistir en un sistema integral que opere a partir de una misma base de datos, o en herramientas manuales e informatizadas que realicen las tareas necesarias para crear un archivo (tabla I.10.6).

Cuando se desarrollen herramientas de recuperación de datos, será necesario identificar las necesidades de los usuarios y asegurarse de que las herramientas responderán a tales necesidades. Para ello se tendrán en cuenta las necesidades para:

- una serie simple, por ejemplo, datos de flujo diarios o mensuales para un período determinado;
- una serie múltiple, por ejemplo, datos de flujo de un grupo de estaciones, o datos de precipitación y flujo fluvial coincidentes; y
- un solo valor de una serie (tales como series con fines de modelización o de visualización en un SIG), por ejemplo, el caudal máximo anual en un emplazamiento, o la precipitación anual media en varios emplazamientos.

Tabla I.10.6. Herramientas de análisis de datos

<i>Herramienta</i>	<i>Descripción</i>
Herramientas de introducción de datos	Teclado de uso manual, equipo y programas para descargar datos de los registradores cronológicos, programas de reformato, herramientas de hoja de cálculo ordinarias para el formateo y almacenamiento de los datos, herramientas de gestión automatizada de datos en tiempo real
Herramientas de tratamiento	Tratamiento primario: por ejemplo, equipo y programas para digitalizar los registros gráficos; tratamiento secundario: por ejemplo, programas para la conversión de los niveles de agua en flujos
Herramientas de validación	Programas para visualizar gráficas y corregir datos, programas para producir curvas de doble masa/hidrógrafos de máxima y mínima, etc.
Herramientas de análisis	Programas (incluidas las hojas de cálculo) para la elaboración de estadísticas, por ejemplo curvas de duración del flujo
Herramientas de búsqueda	Programas para recuperar determinados valores de datos o estadísticas de los archivos
Herramientas de informe	Programas para elaborar informes/datos destinados a la difusión y procedentes de los archivos

En función también de las necesidades de los usuarios, los datos y la información de la base de datos podrán ser recuperados en varios formatos, entre ellos:

- archivos descriptivos: información de distintas fuentes que describa los datos disponibles y sus características;
- archivos ASCII (sección 10.2.2.2);
- archivos CSV: un formato de datos delimitados, en el que las columnas están separadas por comas, y los registros o filas, por saltos de línea. Los campos que contengan un carácter especial (coma, salto de línea o comillas) figurarán entre comillas. Sin embargo, cuando una línea contenga una sola entrada consistente en una cadena de caracteres vacía, será posible incluirla entre comillas; y
- otros formatos definidos por el usuario.

10.3.2 **Recuperación de datos de una sola variable**

A veces el almacenamiento de datos en forma de series cronológicas múltiples no es eficaz. Ello se debe al gran número de variables que es posible observar en cada emplazamiento y al medio utilizado para recuperar los datos.

Considérense, por ejemplo, unos datos climatológicos que, después de ser utilizados inicialmente para el cálculo de la evapotranspiración potencial, pueden ser consultados tan solo para recuperar variables individuales. Esta última operación suele ser necesaria para la interpolación espacial y/o representación gráfica de los datos, por ejemplo, con datos de temperatura destinados al cálculo de la fusión de la nieve, o con datos de radiación destinados a evaluar potenciales de producción agraria. El proceso de recuperación no sería eficaz, ya que habría que buscar en todas las estaciones aun cuando la variable solo hubiera sido observada en algunas de ellas.

Ya se ha visto (sección 10.2.3.10) que estos problemas pueden solucionarse almacenando con cada valor punteros de datos que indiquen la ubicación del registro que contiene el valor siguiente de la variable. Sin embargo, si se utiliza esta técnica para un gran número de variables, se necesitará mucho espacio para almacenar los punteros. Este problema podrá solucionarse eliminando variables importantes –las que son frecuentemente leídas una a una– y almacenándolas como series cronológicas de una sola variable. Ésta es una práctica corrientemente utilizada para los datos de lluvia observados en estaciones climatológicas. El mejor momento para efectuar esta extracción de variables importantes es la actualización anual, en que los datos validados serán transferidos al archivo histórico.

Conviene señalar que la decisión de extraer variables individuales dependerá de la manera en que se prevea recuperar los datos. La recuperación frecuente de valores para una variable específica sugiere la extracción de esa variable del conjunto de variables múltiples. Cuanto menor sea el número de estaciones en que se ha observado la variable, menos eficaz será la búsqueda de variables múltiples y más justificada estará la utilización de un formato de una sola variable.

Cuando, como suele ocurrir con los datos de calidad del agua, se recuperen datos de varias variables relativas al mismo tiempo de observación, es posible que siga siendo más conveniente el formato original de variables múltiples.

10.3.3 **Sistema de recuperación de datos**

La recuperación de datos se examina de manera pormenorizada en la publicación *Guidelines to Computerized Data Processing in Operational Hydrology and Land and Water Management* (WMO-No. 634). La capacidad de recuperar rápidamente conjuntos de datos seleccionados es una de las ventajas fundamentales del tratamiento electrónico de datos hidrológicos. Cuando disponen de sistemas eficaces de recuperación de datos, los hidrólogos y planificadores de recursos hídricos pueden concentrarse en el análisis de los datos y no perder tiempo en ubicar, comparar y procesar manualmente los datos. Un sistema completo de recuperación de datos presentará las características siguientes:

- a) gran diversidad de criterios de selección: normalmente, la selección se efectúa por tipo de variable, cuenca, estación, período de tiempo y valor (o intervalo de valores) de la variable. En particular, será posible seleccionar datos utilizando cualquier combinación de esos criterios;
- b) interpolación/agregación de datos en el tiempo y en el espacio: la más importante de estas opciones será posiblemente la interpolación de series cronológicas irregulares en series regulares, y la agregación de series cronológicas de intervalo corto en totales o promedios de período más largo, es decir, la conversión de valores horarios en diarios o de valores diarios en valores de 10 días.
Si se utilizase algún sistema de referencia geográfico/hidrológico, será posible también ajustar los datos espaciales;
- c) cálculo de estadísticas simples: deberá ser posible calcular estadísticas elementales para los períodos de registro seleccionados. En particular, valores totales (si los hubiere), valores medios, desviaciones típicas e intervalos de valores. Los sistemas estándar podrían ofrecer estadísticas más completas –correlaciones cruzadas, regresiones múltiples, análisis de probabilidad, etc.– o podría existir la posibilidad de transferir los datos seleccionados a un programa informático estadístico (o programa de usuario) como el que se describe más adelante;
- d) selección del formato de salida: esta función debería permitir extraer datos directamente en formatos (especificados) tabulares o gráficos, o para la creación de archivos de datos en formatos apropiados para su ulterior tratamiento. En este último caso, los conjuntos de datos recuperados pueden ser almacenados para introducirlos en un programa informático estadístico o en un programa de aplicación específico para el usuario. Un formato de salida puede ser

- apropiado para el intercambio de datos hidrológicos a escala nacional o internacional; y
- e) selección del dispositivo de salida: debería ser posible seleccionar fácilmente el dispositivo de salida. Como mínimo, debería disponerse de una impresora en línea, una unidad de presentación visual (VDU) y un archivo en disco y, si fuera posible, una trazadora de gráficos. Antes de transferir los datos a una cinta o disco flexible, suele grabarse éstos en disco duro y transferirse con un programa diferente que requerirá de cierto número de variables especificadas por el usuario.

Es importante que los datos recuperados, en particular los destinados a ser impresos en forma tabular, conserven sus claves y banderines relativos a su estado y fiabilidad (sección 9.3).

En el archivo descriptivo de la estación (sección 2.5.2) o en los catálogos de datos, el usuario dispondrá de información general sobre la fiabilidad general de los datos y/o sobre su falta de fiabilidad en períodos específicos.

Los datos pueden ser recuperados de tres maneras:

- a) recuperación de rutina: resúmenes y estadísticas de datos de estaciones producidos mensual y anualmente;
- b) recuperación especificada por el usuario: después de consultar anuarios hidrológicos o catálogos de datos, el usuario podrá solicitar la recuperación de datos mediante el formulario correspondiente, que será tramitado como un proceso por lotes normal. Se necesitarán para

ello operadores informáticos u otros técnicos que introduzcan las solicitudes mediante el programa de recuperación de datos. El formulario de solicitud permitirá escoger entre numerosos soportes de salida; y

- c) recuperación de datos en línea (interactivo): existen varios modos de especificación en línea de las operaciones de recuperación de datos que se examinarán más adelante, en razón de sus amplias posibilidades de utilización.

Como ya se ha observado en este capítulo y puede verse en la figura I.10.4, la existencia de una base de datos principal en línea permite la recuperación interactiva de datos. Sin embargo, salvo en los sistemas que contengan pequeñas cantidades de datos o tengan una gran capacidad de almacenamiento en disco, la mayor parte de la base de datos se almacenará fuera de línea. Así, el modo interactivo directo suele ser apropiado únicamente para recuperar cantidades limitadas de los datos más recientes. En algunos sistemas, el usuario distante puede enviar mensajes al operador informático para solicitar la preparación de un determinado volumen de la base de datos fuera de línea. Sin embargo, esas solicitudes pocas veces son inmediatamente atendidas, por lo que este método puede ser muy ineficaz en términos de utilización de terminales y en costos de comunicación.

Probablemente, el medio más eficaz para especificar una recuperación en línea es un proceso en dos etapas. En la primera etapa, el usuario utiliza un programa interactivo para especificar sus necesidades de recuperación; en la segunda etapa, la solicitud

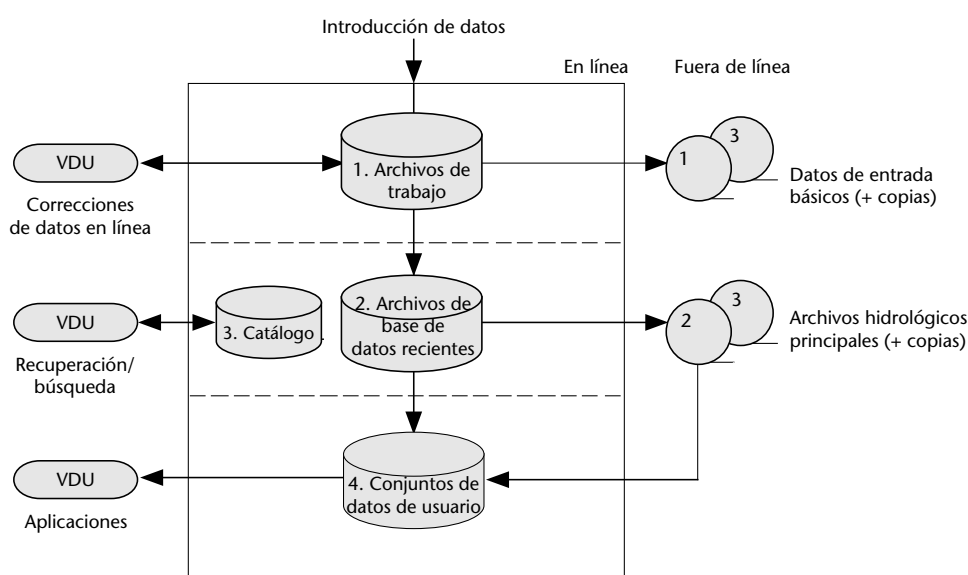


Figura I.10.4. Disposición de los conjuntos de datos en línea y fuera de línea

es tramitada automáticamente como un proceso por lotes, y la salida se obtiene posteriormente. El formato de una interfaz máquina/usuario interactiva se denomina sistema de menús. La recuperación de grandes cantidades de datos en modo de tarea por lotes es mucho más eficaz en términos de la capacidad de la computadora para asignar sus recursos, en particular para los datos contenidos en volúmenes fuera de línea.

Las consideraciones precedentes se refieren principalmente a la recuperación en línea de datos contenidos en inventarios hidrológicos. Sin embargo, la posibilidad de examinar los datos recopilados y almacenados para sistemas en tiempo real es posiblemente una necesidad más importante. Las opciones de recuperación van desde la búsqueda telemétrica de estaciones o grupos de estaciones hasta la representación gráfica y visualización de datos recientemente obtenidos y de predicciones recientes en el centro de tratamiento.

10.4 DIFUSIÓN DE DATOS

10.4.1 Consideraciones generales

Si no se utilizan, los datos carecen de utilidad; los datos hidrológicos solo adquieren valor cuando son analizados y utilizados en el marco de un proceso de planificación y adopción de decisiones de gestión hídrica. Para poder cuantificar el valor medio y la variabilidad (estacional e interanual) de cualquier variable hidrológica se necesitan registros de largo período y de buena calidad. Así, el período de datos típicamente necesario para obtener una estimación 'fiable' de la precipitación anual media es de 30 años, pese a lo cual, considerando la amplísima periodicidad de la precipitación observada en algunas partes del mundo, podría no ser suficiente. Además, ante la manifiesta evidencia de calentamiento mundial y cambio climático, los científicos e ingenieros necesitan registros de largo período para poder detectar y controlar tendencias de precipitación, flujos fluviales y recargas de agua subterránea, a fin de poder preparar planes de contingencia para hacer frente a la evolución de los recursos hídricos.

Para ser útiles, los datos de buena calidad deberán estar también disponibles para una amplia diversidad de usuarios. Los datos son frecuentemente obtenidos por organismos que son sus principales usuarios, ya que aquéllos son esenciales para su trabajo, por ejemplo en relación con el abastecimiento en agua, la irrigación o las instalaciones hidroeléctricas. Se trata, en la mayoría de los casos, de

organismos públicos. En la presente Guía no se aborda el tema de la comunicación interna de los datos en esos organismos. Se examinan en ella los medios que podría utilizar un usuario de datos hidrometeorológicos que no pertenezca al organismo de origen de los datos, suponiendo que éstos hayan sido introducidos en una base de datos apropiada como las descritas en secciones anteriores.

Los usuarios externos de datos hidrológicos podrían ser funcionarios de otros departamentos públicos, compañías de suministro de agua, empresas hidroeléctricas, consultorías de ingeniería o medio ambiente, centros académicos o investigadores. Los usuarios podrían ser muy numerosos, y sus necesidades muy diferentes, desde datos de un punto específico de un río a datos de una región, país o incluso grupo de países, en el caso de ríos transfronterizos.

Acceso a los datos

El acceso internacional a los datos meteorológicos e hidrológicos es un tema que interesa a la OMM y a sus países Miembros desde hace muchos años. Este interés culminó en 1995 con la adopción por el Duodécimo Congreso de la OMM de la Resolución 40 (Cg-XII), que establece la política y práctica de la OMM para el intercambio de datos y productos meteorológicos y afines, incluidas las directrices sobre las relaciones establecidas en las actividades meteorológicas comerciales. En 1999, en su siguiente reunión, el Decimotercer Congreso de la OMM adoptó la Resolución 25 (Cg-XIII) – Intercambio de datos y productos hidrológicos, que establece las políticas y prácticas para el intercambio internacional de datos y productos hidrológicos. Esta resolución proporciona un marco que facilita el acceso internacional a los datos y productos hidrológicos (OMM, 2001). El texto completo de la Resolución 25 (Cg-XIII) puede consultarse en la dirección http://www.wmo.ch/pages/about/Resolution25_es.html.

10.4.2 Catálogos de datos disponibles

La primera necesidad de los usuarios de los datos suelen ser mapas que indiquen la ubicación de todas las estaciones, junto con las correspondientes tablas de datos y el período de registro de cada una de ellas. Este tipo de información constituye la base de metadatos del conjunto de datos, y puede consistir en un conjunto de tablas de datos contenido en una base de datos informatizada (como las anteriormente descritas), o en una copia impresa en papel. Los usuarios pueden así acceder electrónicamente a la información en un portal de Internet, o buscarla en anuarios publicados en papel.

Tradicionalmente, la información se publicaba en anuarios, aunque en ocasiones se imprimían solo catálogos resumidos cada tres o cinco años, ya que las redes de datos son generalmente bastante estables. Este sistema es simple, pero puede resultar gravoso en costos de impresión, y los catálogos impresos podrían no ser fácilmente accesibles para todos los usuarios. Sin embargo, durante muchos años los catálogos impresos han sido el medio más eficaz para difundir información sobre los datos disponibles, y seguirán siendo eficaces en los países en que el acceso a Internet no esté muy extendido o no sea fiable.

Con todo, es probable que el medio de acceso más común a estos catálogos sea en el futuro un navegador web enlazado directamente a los metadatos. La ventaja radica en que todos los usuarios con acceso a Internet podrán acceder fácilmente a los mismos y no se necesitarán ya anuarios. El mantenimiento y actualización del sistema podrían ser también más fáciles.

La información suministrada para cada cuenca medida podrá consistir en:

- a) pormenores de la cuenca, por ejemplo, dimensiones, geomorfología, configuración del terreno, vegetación y usos de la tierra;
- b) zona climática y precipitación de lluvia y evaporación promedio anual en la cuenca;
- c) ubicación, tipo y calidad de la estación de aforo;
- d) información sobre posibles reglamentaciones o factores corriente arriba que pudieran complicar la utilización de los registros;
- e) período, completitud y calidad de los registros de flujo fluvial y calidad del agua (incluido el transporte de sedimentos); y
- f) ubicación de las estaciones meteorológicas en la cuenca o en sus inmediaciones y períodos de registro correspondientes.

Esta información se agrupa y analiza en tres apartados: información descriptiva, mapa de la cuenca y disponibilidad de datos.

Para ayudar a los usuarios a identificar las cuencas que respondan a sus necesidades, se proporcionará una descripción de las características de cada cuenca observada, junto con las características principales de las instalaciones de aforo, así como indicaciones sobre la calidad y fiabilidad del registro de flujo.

En la tabla I.10.7 se sugieren varios conceptos, junto con la información correspondiente. En la práctica, es posible que no se disponga de todos los detalles, o que éstos no se correspondan con cada concepto

para una cuenca observada, pero se recomienda mantener el mismo formato en todos los casos.

En la figura I.10.5 se ofrece un ejemplo como complemento de la tabla I.10.7. Es útil disponer de un mapa de cada cuenca o grupo de cuencas. La escala del mapa estará adaptada a la información presentada. Cuencas representadas a escalas diferentes pueden requerir mapas de escalas diferentes. En un futuro próximo, toda la información necesaria para producir mapas de cuencas se conservará en SIG informáticos, y podrá ser consultada en escalas diferentes. En la tabla I.10.8 se describe la información que habrá de figurar en el mapa, y en la figura I.10.6 se expone un ejemplo básico de ella.

La página dedicada a los datos disponibles contendrá un resumen relativamente conciso y fácilmente actualizable de los datos de flujo fluvial, precipitación y calidad del agua. Estará basada en datos mensuales de flujo y precipitación, y en datos anuales de calidad del agua. En las cuencas con gran número de estaciones pluviométricas, no será práctico incluir un resumen para cada estación. Todas las estaciones, junto con su período de registro, figuran en el mapa descrito en la sección anterior, por lo que bastaría con limitar la disponibilidad de los datos a aquellos provenientes de pluviógrafos y de un conjunto seleccionado de estaciones de medición diaria de la precipitación. Las estaciones de largo período tendrían que ocupar varias páginas para permitir una legibilidad adecuada.

Se sugiere incluir la información de la tabla I.10.9 en la página sobre disponibilidad de datos.

10.4.3 Informes resumidos

Son numerosas las organizaciones que publican resúmenes de datos. Por ejemplo, promedios climatológicos, estadísticas de precipitación, estadísticas o registros de flujo fluvial, y registros o estudios de la calidad del agua.

Habitualmente, esas publicaciones contienen información sobre las estaciones, incluidos el número de estación, su latitud y longitud, los tipos de datos recogidos, ciertas particularidades del emplazamiento (nombre, nombre del río, coordenadas del cuadrículado, cuenca, etc.), el período de operación, el período de datos procesado, y cierto número de resúmenes instantáneos, diarios, mensuales y anuales (incluidos los valores mínimos, máximos y medios). Los datos se pueden presentar en forma de texto, adjuntos en una microficha o en formato informático, por ejemplo, en disco o CD-ROM.

Tabla I.10.7. Descripción del formato de un catálogo de datos

<i>Identificación</i>	<i>Descripción</i>
Nombre	Nombre del río, nombre y número de la estación
Cuenca	Nombre y número de la cuenca
Ubicación	Ubicación de la estación de medición en latitud y longitud, y en coordenadas del cuadrículado locales
<i>Información sobre la cuenca</i>	
Superficie de la cuenca	La superficie de la cuenca expresada en kilómetros cuadrados
Zonas climáticas	Clima en la cuenca, expresado en áreas bioclimáticas que reflejen la cantidad y pautas de la precipitación
Promedio de lluvia	Evaluación de las lluvias anuales medias en el centroide de captación y, en cuencas de gran tamaño, intervalo de valores de lluvia anual medio en toda la cuenca. Se indicarán las fuentes de procedencia de las imágenes
Evaporación en tanque	Evaluación de la evaporación en tanque anual media en el centroide de la cuenca. Se citarán las fuentes de procedencia de las imágenes
Geomorfología	Descripción de la topografía, del paisaje y de la geología de la cuenca observada
Topografía	Estimación cuantitativa de las proporciones de los principales accidentes del terreno en el interior de la cuenca
Vegetación natural	Descripción de la vegetación natural, obtenida de levantamiento
Deforestación	Proporción de la vegetación natural eliminada o sustancialmente alterada por efecto de actividades humanas intensivas. Se indicará la fuente y la fecha de las estimaciones de la vegetación eliminada.
Vegetación presente	Descripción de la cubierta vegetal actual en toda la cuenca, haciendo referencia a la fuente
Uso de la tierra	Comentario sobre los usos de la tierra. Se citará la fuente de proveniencia de la información, indicando si ésta proviene de observaciones in situ, de mapas de usos de la tierra rurales o de evaluaciones más detalladas
Regulación	Comentarios sobre incidencias corriente arriba que pudieran modificar el régimen de escorrentía. Se indicarán las posibles fuentes de información detallada
Comentario general	Cuando la estación no mida la escorrentía total en la cuenca o no sea posible corregir el registro para reflejar la regulación aguas arriba, se omitirán las características de la cuenca, sustituyéndolas por un comentario sobre las finalidades o funciones especiales de la estación
<i>Información sobre la estación de aforo</i>	
Período de registro	Mes y año de apertura y cierre de la estación de aforo. Cuando haya más de una estación que ha operado cerca del mismo río, se incluirá una referencia apropiada
Clasificación	Clasificación actual de la estación de aforo en la red hidrológica (por ejemplo, estación de proyecto o estación de red básica)
Equipo de medición	Descripción de los instrumentos de registro de altura del agua y de las características que controlan ésta en la estación de aforo. Convendría indicar los cambios introducidos en cualquiera de estos equipos durante el período de utilización
Registro de altura del agua	Promedio anual porcentual de los datos registrados y porcentaje de éstos que ha de ser interpretado durante el tratamiento (registros defectuosos)
Curva de gasto	Comentarios sucintos sobre el método utilizado y la calidad de la relación altura-caudal, junto con los valores de caudal máximo medidos. Siempre que sea posible, la proporción de flujo medido que representa el caudal máximo medido será un dato conocido
Medida de la sensibilidad	Se aportará algún indicador de la sensibilidad de la curva de gasto. El método preferido para indicar la sensibilidad es el porcentaje del volumen de flujo que sería posible medir dentro de un margen del 1, 2 o 5 por ciento, con un error de 1 mm en el registro de valores de altura. Obsérvese que este indicador está basado en la pendiente de la curva de gasto y de la curva de duración de flujo acumulativa. También puede definirse en términos de un error de 10 o 100 mm en la altura del agua

607003 río Warren	Granja Wheatley		
Cuenca fluvial	Río Warren		
Ubicación	Latitud S 34°22' 14"	Cuadrícula AMG	N 6196500
	Longitud E 116°16' 34"		E50 433450
<i>Características de la cuenca</i>			
Superficie de la cuenca	2 910 km ²		
Zona climática	Clima mediterráneo; lluvias intermedias a escasas		
Promedio de lluvia	735 mm/año (rango 950 a 550)		
Evaporación en tanque	1 275 mm/año (rango 1 250 a 1 400)		
Geomorfología	Relieve escaso a moderado; meseta ondulada surcada por un valle principal, suelos de laterita bauxítica sobre rocas arqueanas graníticas y metamórficas		
Topografía	Unidades cartográficas; Atlas del suelo de Australia (Ref. 8) 16% – lateritas diseccionadas Ub90; paisaje ondulado con suelos moteados de amarillo y cadenas de montañas pedregosas 14% – llanuras pantanosas Cb43, Tf6; líneas de drenaje someras con arenas lixiviadas y suelos podzólicos 57% – meseta de laterita Cd22, Tc6; tierras altas de arena y gravilla de mineral de hierro sobre arcilla moteada 13% – valles en surco Tf6, Ta9; pendientes moderadas, suelos principalmente podzólicos, amarillos		
Vegetación natural	Unidades cartográficas; estudio de la vegetación de Australia Occidental (Ref. 1) 20% – terreno boscoso eMi; bosques de Eucalyptus calophylla, Eucalyptus wandoo en lateritas disectadas 70% – bosque eMc; bosques de Corymbia calophylla y Eucalyptus calophylla en meseta de laterita 10% – bosques bajos mLi; bosques de Melaleuca en llanuras pantanosas		
Deforestación	Aproximadamente un 40% de la superficie talada (solo un 27% en 1965)		
Uso de la tierra	Aproximadamente la mitad de la cuenca es un bosque estatal; las superficies deforestadas se destinan a la producción ovina y cerealera en la cuenca superior y a la producción bovina en los tramos inferiores		
Regulación	Pequeñas represas agrícolas en cursos fluviales secundarios		
<i>Datos sobre la estación de aforo</i>			
Período de registro	Desde mayo de 1970 hasta la fecha		
Clasificación	Red hidrológica – Cuenca hidrográfica primaria		
Equipo de medición	Servomanómetro L&S y registrador gráfico continuo hasta la fecha. Control de bancos rocosos para flujos medios; control del canal para flujos altos		
Registro de altura del agua	96,5% registrado, 7,6% con fallos		
Curva de gasto	Curva de gasto para flujos bajos a medios, aceptable debido a la naturaleza del control; para flujos medios a altos, de buena calidad, aunque teóricamente fuera del intervalo medido. Numerosas mediciones de caudal a 97,04 m ³ s ⁻¹ , que representa un 99% del flujo total del registrador		
Medida de la sensibilidad	99% del flujo < 1; 100% del flujo < 2		

Figura I.10.5. Ejemplo de una página de un catálogo de datos

Tabla I.10.8. Descripción de las características de los mapas

Identificación	Descripción
Límites de la cuenca	Escala y origen del mapa a partir del cual se han definido los límites de la cuenca
Líneas de corriente	El número de líneas de corriente que deberán incluirse estará en función de la superficie de la cuenca. Fuente de los datos sobre las líneas de corriente
Escala de la cuenca	Variable, en función del tamaño de la cuenca
Estaciones de precipitación de lluvia	Ubicación y número de estación, período de funcionamiento y tipo de pluviómetro, por ejemplo, de lectura diaria o totalizador
Isoyetas de precipitación de lluvia (opcional)	Valores anuales promedios de las isoyetas de lluvia en la cuenca, indicando la referencia de lluvia
Uso de la tierra (opcional)	Cuando proceda, convendría averiguar los contornos de los distintos usos de la tierra, por ejemplo, los contornos de los bosques, de los cultivos o de los núcleos urbanos

10.4.4 Anuarios

Los anuarios son un medio muy eficaz de difundir datos hidrológicos, aunque solo pueden contener ciertos tipos de datos. Así, mediante modernos registradores cronológicos o telemedidores de datos, se miden cada vez con mayor frecuencia los valores de precipitación de lluvia y los niveles fluviales (y, por consiguiente, los flujos) cada 15, 30 o 60 minutos, que representan entre 8 760 y 35 040 valores al año. No es práctico ni, generalmente, necesario publicar datos con tal grado de resolución, por lo que en los anuarios se publican generalmente tan solo los totales de precipitación

diarios, o incluso mensuales, y los flujos diarios medios.

Los datos sobre aguas subterráneas varían muy lentamente, y solo se controlan de manera intermitente, posiblemente una vez por semana o por mes. Por consiguiente, estos conjuntos de datos se pueden publicar completos. Otras variables climatológicas, como la temperatura, la velocidad del viento o la radiación, se publican únicamente, en muchos casos, en forma de medias mensuales.

En las figuras I.10.7 a I.10.13 se ofrecen ejemplos del contenido típico de un anuario, extraídos del National River Flow Archive (centro de datos de caudal fluvial) de Reino Unido.

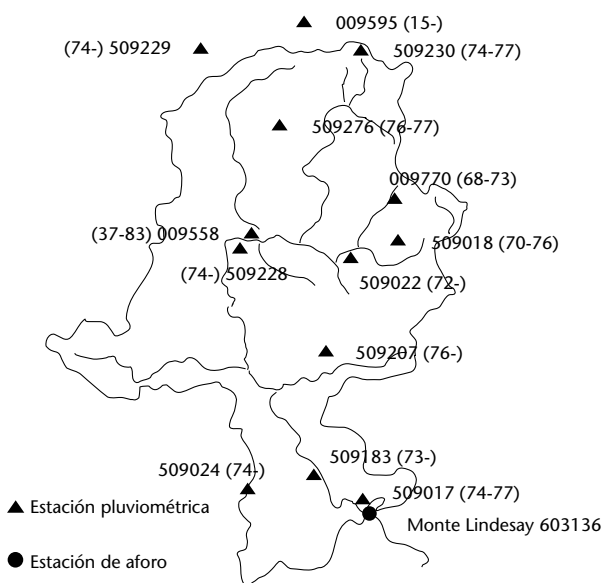


Figura I.10.6. Cuenca del río Dinamarca

10.4.5 Exportación de datos previa solicitud

Los datos hidrológicos nacionales se publican cada vez más por Internet, donde los usuarios pueden utilizar un mapa y una interfaz de diálogo gráfica para seleccionar las estaciones que les interesan y los tipos de datos que desean descargar. El acceso por Internet permite al usuario explorar el conjunto de datos y determinar los tipos de datos que aquél necesita de una estación o conjunto de estaciones.

Algunos sistemas permiten seguidamente al usuario descargar los datos seleccionados directamente a su PC, mientras que en otros casos hay que solicitar electrónicamente los datos en el sitio web. Una buena razón para no permitir que los usuarios descarguen todos los datos que solicitan es el volumen que pueden llegar a alcanzar éstos; además, con

Tabla I.10.9. Descripción de la(s) página(s) indicativa(s) de la disponibilidad de los datos

Identificación	Descripción
Datos de flujo	Registros disponibles y calidad de los registros expuestos con claridad mes por mes
Datos de precipitación de lluvia	Registros disponibles y calidad de los registros expuestos con claridad mes por mes respecto del principal pluviógrafo y de los pluviómetros de lectura manual. El período de registro abarcado podría limitarse al período abarcado por la estación de aforo de caudales, por razones prácticas
Calidad del agua	Número de muestras analizadas anualmente, clasificadas por tipos de análisis claramente definidos. Por ejemplo: <ul style="list-style-type: none"> a) muestras sometidas a un análisis básico únicamente (parámetros de conductividad, pH, temperatura del río, color, o turbidez); b) muestras analizadas para determinar los iones principales; c) muestras analizadas para determinar los nutrientes; y d) muestras analizadas para determinar los metales pesados u otros componentes vestigiales

algunos proveedores la lentitud de la transferencia puede ser inaceptable, especialmente cuando se utilizan módems lentos. Por esa misma razón, puede ser problemático enviar los datos por correo electrónico, dadas las limitaciones impuestas por algunos portales al tamaño de los correos electrónicos.

En muchos casos, será preferible que los usuarios presenten una solicitud de datos en el sitio web y que los datos les sean enviados en CD o sean

insertados en un sitio FTP, desde donde podrán descargarlos.

Los datos podrían estar disponibles gratuitamente en el sitio web, en particular cuando el usuario pueda descargarlos directamente. En algunos casos, sin embargo, es posible que los usuarios deban pagar gastos de preparación para cubrir el costo de personal que trae aparejada la preparación del CD. Aunque algunos usuarios pueden objetar este pago, el precio suele estar justificado, ya que el organismo suministrador ha de rendir cuentas a sus financiadores y a sus directores. El importe cobrado por los datos permitiría cubrir, por lo menos, parcialmente las necesidades de financiación del organismo pero, sobre todo, significa que su trabajo es valorado por los usuarios o clientes externos.

El National Water Archive (centro de datos del agua) de Reino Unido (<http://www.nwl.ac.uk/ih/nwa/index.htm>), es un buen ejemplo de sistema de recuperación de datos por Internet, así como el sitio web para datos hidrológicos del USGS (<http://waterdata.usgs.gov/nwis/>), o los datos de los proyectos WHYCOS (<http://medhycos.mpl.ird.fr/> y <http://whycos.org/cms/content/data-and-products>).

10.4.6 Formatos de intercambio de datos

No hay actualmente normas aplicables a los formatos de intercambio de datos hidrológicos. Las únicas normas existentes son los estándares de facto establecidos por los registradores de datos y programas informáticos de bases de datos más habituales. Los formatos actuales de intercambio de datos se incluyen generalmente en una de las dos categorías siguientes: archivos en forma de texto o XML.

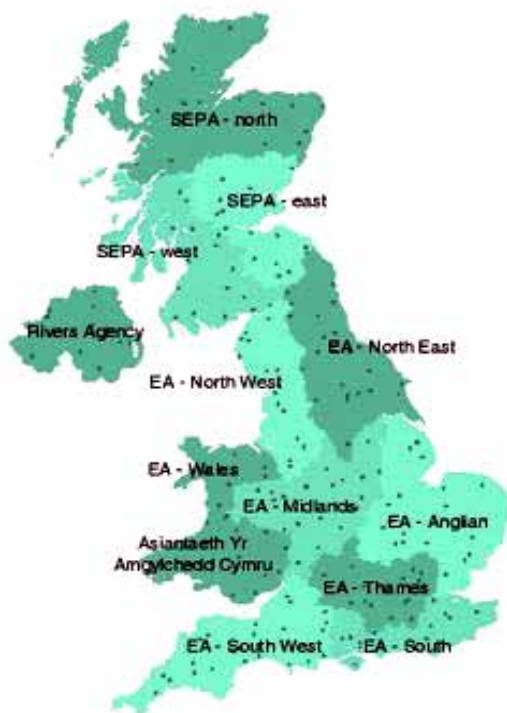


Figura I.10.7. Mapa de estaciones de aforo de Reino Unido conservado en el National River Flow Archive

039002 1996 Río Táchesis en la esclusa de DayAutoridad de medición: EA
Primer año: 1938Redícula de referencia: 41 (SU) 568 935
Nivel de estación (m OD): 45,80Cuenca de drenaje (km²): 3 444,7
Altura máxima (m OD): 330**Caudal promedio diario medido (en metros cúbicos por segundo)**

Día	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Sept.	Oct.	Nov.	Dic.
1	56,600	25,800	42,600	24,800	20,400	11,000	5,380	3,970	3,730	4,120	3,520	10,000
2	57,500	24,200	40,400	21,800	23,200	8,850	5,490	3,710	3,050	3,960	4,020	9,970
3	60,600	24,200	39,100	21,100	19,300	9,250	5,320	3,640	2,770	3,550	4,060	11,100
4	56,300	22,600	38,600	20,800	13,100	9,160	5,140	3,630	2,700	2,860	7,280	12,600
5	54,700	22,400	37,700	21,100	13,900	9,130	5,170	3,430	3,120	3,090	6,880	15,400
6	54,900	24,000	33,800	20,800	15,500	8,860	6,050	3,200	3,100	3,060	8,280	13,500
7	63,400	24,000	33,700	20,500	15,500	10,100	5,640	3,060	3,060	3,040	5,620	11,200
8	77,900	23,600	33,500	20,100	16,400	14,600	5,040	2,750	2,980	3,300	5,350	9,310
9	101,000	30,700	35,000	20,100	15,500	11,800	5,270	2,760	2,970	4,110	6,450	9,320
10	109,000	53,600	35,900	21,500	14,200	11,200	5,410	3,720	2,880	3,250	5,810	8,870
11	97,600	57,200	33,300	22,500	14,100	9,190	3,870	4,850	2,990	2,980	3,210	8,810
12	84,100	81,600	34,600	22,500	14,200	8,880	5,500	4,320	3,220	3,520	4,470	8,840
13	77,800	99,800	36,000	48,400	14,100	8,410	4,730	4,220	3,250	3,430	5,480	8,340
14	69,400	90,200	35,200	40,600	13,800	7,440	3,870	3,790	3,020	3,130	4,590	8,140
15	59,300	64,100	32,000	31,500	13,900	5,880	3,940	3,680	2,940	3,170	4,340	8,500
16	54,800	53,900	31,100	26,400	13,300	6,240	3,810	3,380	2,840	3,220	5,240	8,020
17	50,400	48,400	30,600	24,700	13,200	6,020	3,820	3,340	3,180	3,420	6,390	8,430
18	46,500	48,400	25,500	21,100	13,200	5,980	3,700	2,850	2,500	4,2900	5,900	8,020
19	45,600	47,700	27,300	23,900	14,100	5,990	2,650	2,840	2,620	4,030	10,700	10,500
20	44,300	40,700	26,700	24,400	13,900	5,970	3,210	2,940	2,910	3,210	11,600	12,900
21	41,100	38,500	26,800	23,500	13,900	5,930	3,720	3,170	2,900	3,680	11,400	15,000
22	37,400	37,100	27,800	25,200	13,400	6,130	3,350	3,340	2,850	3,500	10,700	14,300
23	37,900	37,200	33,900	35,100	14,000	5,990	3,260	5,340	2,850	3,380	9,080	12,100
24	38,400	59,000	36,400	42,800	19,500	5,660	3,210	6,820	3,150	3,410	8,330	10,100
25	37,600	95,400	32,500	26,800	15,300	5,510	3,200	7,110	3,660	3,010	11,900	10,700
26	33,700	92,500	34,200	25,000	15,700	5,480	3,270	4,790	3,750	3,530	11,900	10,100
27	32,900	73,700	45,900	22,000	13,100	4,660	3,430	3,800	3,300	4,120	12,600	9,530
28	26,800	59,100	33,300	21,700	12,800	4,970	4,000	3,960	3,340	3,870	10,300	8,340
29	26,500	43,600	32,300	20,600	12,800	5,570	5,480	3,210	3,370	4,670	10,100	8,720
30	26,300		26,300	20,500	11,100	5,520	4,620	3,590	3,350	4,560	9,390	9,030
31	26,100		25,500		11,000		3,950	4,180		3,240		8,830
Promedio	54,400	49,750	33,470	25,390	14,750	7,646	4,339	3,851	3,078	3,539	7,496	10,270
Mínimo	26,100	22,400	25,500	20,100	11,000	4,660	2,650	2,750	2,500	2,860	3,210	8,020
Máximo	109,000	99,800	45,900	48,400	23,200	14,600	6,050	7,110	3,750	4,670	12,600	15,400
(Total mensual millones de m ³)	145,70	124,70	89,64	65,82	39,52	19,82	11,62	10,32	7,98	9,48	19,43	27,52
Escorrentía (mm)	42	36	26	19	11	6	3	3	2	3	6	8
Lluvia (mm)	42	63	36	49	36	21	34	53	22	50	79	29

Estadísticas de datos mensuales del registro precedente (octubre de 1938 a diciembre de 1995)

Flujos medios	Prom.	56,450	56,680	44,600	30,650	20,140	14,260	8,397	7,073	8,666	14,960	30,750	45,570
	Min. (año)	6,252	5,548	5,619	4,255	2,854	1,504	0,401	0,290	1,740	2,782	3,751	5,308
		1976	1976	1976	1976	1976	1976	1976	1976	1959	1959	1990	1975
	Máx. (año)	133,600	120,800	163,200	85,060	61,140	41,560	48,810	18,690	38,640	74,570	128,100	128,700
		1939	1977	1947	1951	1983	1955	1968	1977	1946	1960	1960	1960
Escorr.	Prom.	44	40	35	23	16	11	7	5	7	12	23	35
	Min.	5	4	4	3	2	1	0	0	1	2	3	4
	Máx.	104	85	127	64	48	31	38	15	29	58	96	100
Lluvia	Prom.	68	47	53	47	58	54	53	64	62	64	70	73
	Min.	13	3	5	4	7	5	5	3	5	6	8	16
	Máx.	132	135	152	99	131	124	117	149	129	163	178	316

(continúa)

(continuación)

Resumen

Factores que afectan a la escorrentía

	En 1996	En el registro, antes de 1996	1996 Porcentaje antes de 1996	\$ Abstracción para suministro público \$ Reducción del flujo por abstracciones industriales y/o agrarias \$ Aumento por producción de desechos
Flujo medio	18,070	28,050	64	
Media anual mínima		10,100	1973	
Media anual máxima		51,290	1960	
Media mensual mínima	3,078	Sep. 0,290	Agosto de 1976	
Media mensual máxima	54,400	Enero 163,200	Marzo de 1947	
Media diaria mínima	2,500	18 de sep. 0,050	7 julio de 1976	
Media diaria máxima	109,000	10 de en. 349,000	17 marzo de 1947	
10% de excedente	44,770	67,810	66	
50% de excedente	9,412	15,940	59	
95% de excedente	2,959	3,181	93	
Total anual (millones de m ³)	571,40	885,20	65	
Escorrentía anual (mm)	166	257	65	
Lluvia anual (mm)	514	713	72	
	{promedio de lluvias en 1961–1990 (mm)}	690}		

Figura I.10.8. Ejemplo de tabulaciones del anuario del National River Flow Archive

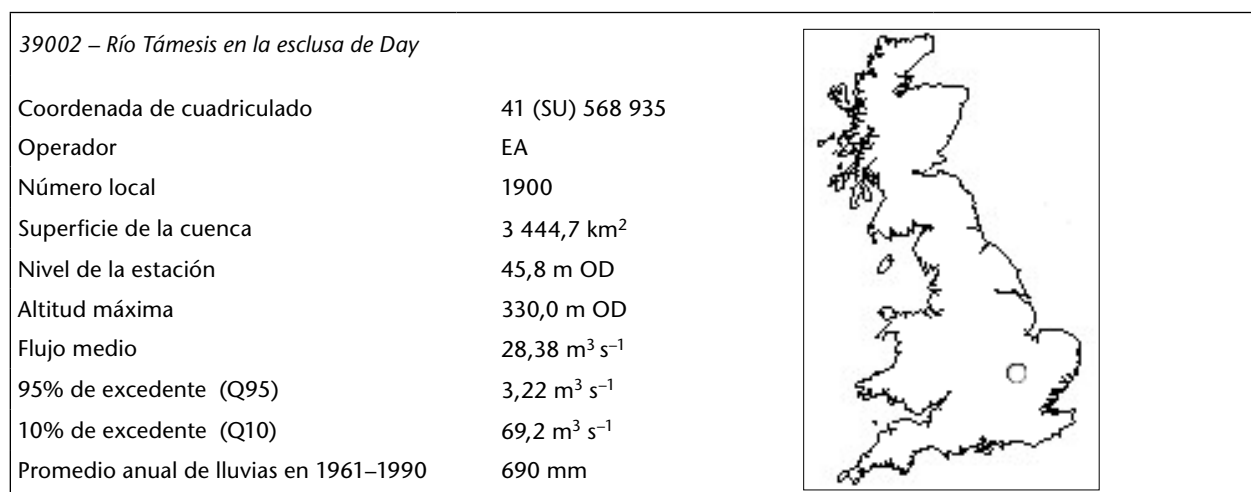


Figura I.10.9. Características de una estación

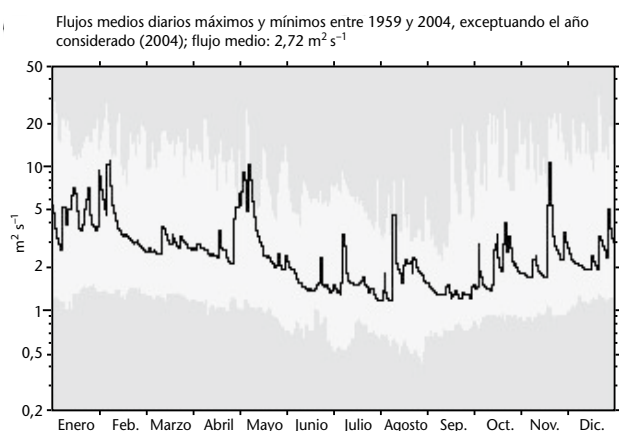


Figura I.10.10. Muestra de hidrograma de flujos diarios medidos

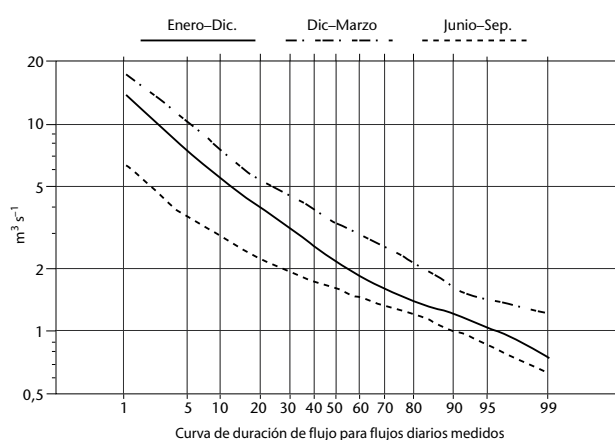


Figura I.10.11. Curva de caudales clasificados para flujos diarios medidos

Descripción de la estación

Vertedero ajustable de plancha delgada (anchura: 5,48 m), más 15 compuertas radiales, que reemplazó una concatenación de compuertas radiales y rectangulares en 1969. Fórmulas de gasto basadas en mediciones – calibración del nivel de descarga para flujos > 70 m³ s⁻¹; rebose por encima de 100 m³ s⁻¹. Flujos diarios naturalizados disponibles en el punto de referencia (iguales a los flujos medidos hasta 1973) – reflejan solo las pérdidas en la estación hidroeléctrica de Didcot. Flujos máximos en estudio.

Descripción de la cuenca

Geología mixta (aguas de cabecera de caliza oolítica, con arcillas Oxford inferiores). Predominantemente rural; desarrollo concentrado a lo largo del valle.

Factores que afectan a la escorrentía

- Escorrentía reducida por abstracción para suministro hídrico
- Escorrentía incrementada por vertidos de desecho
- Escorrentía reducida por abstracción industrial/agrícola

Figura I.10.12. Ejemplo de metadatos

Archivos en forma de texto

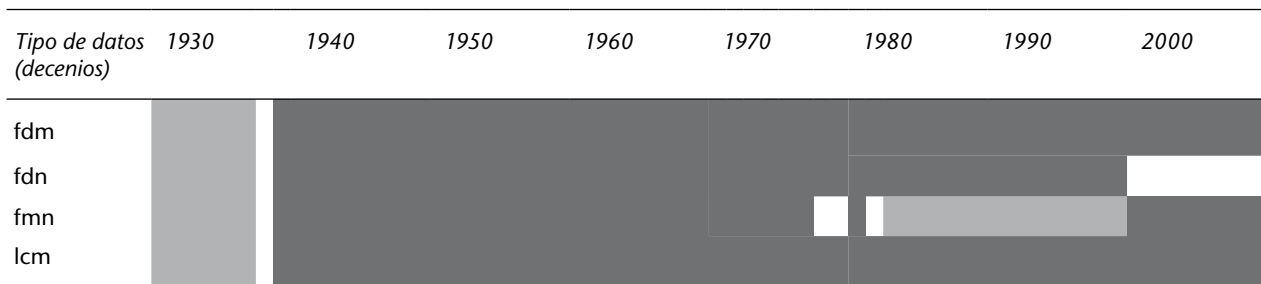
Los archivos de datos en forma de texto tienen la ventaja de que los usuarios pueden consultarlos fácilmente utilizando los programas informáticos más sencillos. En ellos, las series cronológicas pueden indicar las fechas, horas y valores en columnas, separadas por un delimitador que podría ser una coma (en cuyo caso se obtiene un archivo CSV, de valores separados por coma) u otro carácter o número de espacios fijo.

Formatos propios

La desventaja de los formatos en forma de texto es el tamaño del archivo resultante. Muchos sistemas informáticos utilizan formatos propios, que aprovechan la memoria mucho más eficazmente. Se obtienen así archivos más pequeños que ocupan menos espacio en los discos de computadora o en los soportes de transferencia de datos y que circulan más fácilmente por Internet. La desventaja, naturalmente, es que suelen ser necesarios programas especiales para leer los archivos.

XML

Se ha observado una cierta tendencia a identificar el formato de intercambio de datos con la adopción generalizada de XML, o Lenguaje Extensible de Marcado (Consortio Word Wide Web, 2004). XML es un formato estándar de archivos de datos, aunque en el nivel más general. Dada la creciente utilización de Internet en los años noventa, particularmente en los sectores de edición, fabricación y venta al por menor, se necesitaba una norma que permitiera a los programas informáticos leer archivos de datos y descifrar su contenido sin conocer el formato. En 1998, el consorcio World Wide Web (W3C) definió el lenguaje XML como plataforma y método que permitiera insertar datos estructurados en un archivo de texto. El formato XML es similar al lenguaje de etiquetado de hipertexto (HTML), que es el formato de archivo más común en Internet, y que permite visualizar el contenido de los archivos mediante una gran diversidad de programas informáticos. Pero, a diferencia del HTML, que describe el aspecto que presentará el contenido, el formato XML no presupone nada acerca de la presentación, y se limita a describir el contenido.



Flujos diarios medidos (fdm): de 1938 a 2003
 Flujos diarios naturalizados (fdn): de 1938 a 2002
 Flujos mensuales naturalizados (fmn): de 1938 a 2002
 Lluvia de cuenca mensual (lcm): de 1938 a 2001

Figura I.10.13. Datos de flujo fluvial y de lluvia de la cuenca conservados en el National River Flow Archive

El contenido del archivo de datos se describe mediante etiquetas. Así, por ejemplo, en la tabla I.10.10 se indica el nombre y las coordenadas de una estación de aforo en XML.

Tabla I.10.10. Nombre y coordenadas de una estación de aforo en XML

```
<estacióndeaforo>
  <nombre>río Támesis en Wallingford
</nombre>
  <coordenadas>
    <unidad de distancia hacia el este=metros>
      461300
    </distancia hacia el este>
    <unidad de distancia hacia el norte=metros>
      189900
    </distancia hacia el norte>
  </coordenadas>
</estacióndeaforo>
```

Un programa informático que lea este código sabrá, sin conocimiento alguno de hidrología, que el archivo contiene información sobre una 'estación de aforo' que posee un 'nombre' que contiene el valor 'río Támesis en Wallingford' y que contiene 'coordenadas' con los atributos adicionales 'hacia el este', cuyo valor es 461300, y 'hacia el norte', con el valor 189900, ambos expresados en 'metros'. Los símbolos '<>' se denominan etiquetas, y cada par de etiquetas contiene un valor de datos, mientras que el texto contenido en las etiquetas describe los datos.

Las ventajas y desventajas del XML son objeto de amplios debates, pero pueden resumirse en términos simples.

Ventajas: capacidad para separar la forma del contenido y, por lo tanto, para aplicar rápidamente reglas de visualización diferentes a distintos archivos con un mismo formato. Los datos que pueden ser almacenados en un archivo, así como las reglas correspondientes, se pueden indicar explícitamente, y el programa puede utilizar esa indicación para validar los archivos de datos durante su lectura. Los archivos también permiten también efectuar búsquedas eficaces de su contenido.

Desventajas: los archivos de texto no están comprimidos y son muy voluminosos. El formato XML no se inventó para describir series cronológicas de datos que podrían multiplicar por 10 el tamaño de los archivos, incluso con formatos de texto no comprimidos.

Una gran ventaja de XML es que permite la especialización en temas particulares. Las bibliotecas, por ejemplo, han definido un formato internacional para describir las etiquetas y unas reglas para almacenar información bibliográfica en XML. La existencia de esas normas implica que todas las bibliotecas pueden proporcionar datos que las demás bibliotecas podrán leer y comprender. Ese mismo proceso se está repitiendo en el sector más complejo de las ciencias medioambientales. Están apareciendo formatos XML para muy diversas aplicaciones, como la descripción de moléculas, o el Lenguaje de Modelización de Climatología. Los datos de SIG disponen ahora de una norma de amplio alcance basada en XML, llamada Lenguaje de Etiquetado Geográfico (GML), que permitirá la interacción entre mapas digitales de todo tipo de fuentes, y que podrá ser utilizada para la comunicación de datos espaciales. El formato GML es la gramática XML definida por el Open Geospatial Consortium (OGC) para expresar características geográficas (Cox y otros, 2004). Puede utilizarse como lenguaje para la modelización de sistemas geográficos, y como formato de intercambio abierto para transacciones geográficas por Internet.

Muchas de las definiciones de estos dialectos de XML están evolucionando todavía, por lo que habrán de ser utilizadas con precaución. Sin embargo, algunos lenguajes han sido satisfactoriamente definidos y reconocidos mediante una norma ISO. No se ha desarrollado todavía un dialecto XML para la hidrología, pero el Servicio Meteorológico Nacional de Estados Unidos ha establecido un consorcio sobre aplicaciones hidrológicas de XML y ha elaborado el borrador de un plan de aplicaciones hidrológicas en XML.

Referencias y lecturas sugeridas

- Consortio World Wide Web, 2004: *Extensible Markup Language (XML) 1.0*, tercera edición, W3C Recommendation, 4 de febrero de 2004 (<http://www.w3.org/TR/2004/REC-xml-20040204>).
- Cox, S., P. Daisey, R. Lake, C. Portele y A. Whiteside (eds.), 2004: *OpenGIS Geography Markup Language (GML) Implementation Specification Version 3.1.0*, Recommendation Paper, febrero de 2004, Referencia Open GIS Consortium, Inc. e ISO núm. OGC 03-105r1.
- Kilpatrick, M. C., 1981: *WATSTORE: A WATER Data STORAGE and RETRIEVAL System*, publicación de la Oficina Gubernamental de Impresión de Estados Unidos, 52, Departamento del Interior de Estados Unidos, Servicio Geológico de Estados Unidos, Reston, Virginia, 341-618.

- Maurer, T., 2004: *Globally Agreed Standards for Metadata and Data on Variables describing Geophysical Processes: A Fundamental Prerequisite for an Integrated Global Data and Information Infrastructure and Thus Improved Management of the Earth System for Our All Future*, Global Runoff Data Centre, Informe 31, octubre de 2004.
- Organización Meteorológica Mundial, 1981: *Case Studies of National Hydrological Data Banks: Planning, Development and Organization*, Operational Hydrology Report No. 17, WMO-No. 576, Ginebra.
- , 1983: *Guía de prácticas climatológicas*, segunda edición, OMM-N° 100, Ginebra.
- , 2001: *Exchanging Hydrological Data and Information: WMO Policy and Practice*, WMO-No. 925, Ginebra.
- y Organización de las Naciones Unidas para la Agricultura y la Alimentación, 1985: *Guidelines for Computerized Data Processing in Operational Hydrology and Land and Water Management*, WMO-No. 634, Ginebra.
- Woolf A., B. Lawrence, R. Lowry, K. Kleese van Dam, R. Cramer, M. Gutierrez, S. Kondapalli, S. Latham, D. Lowe, K. O'Neill y A. Stephens, 2006: "Data integration with the Climate Science Modelling Language", en *Advances in Geosciences*, 8, 83–90 (<http://www.copernicus.org/EGU/adgeo/8/adgeo-8-83.pdf>).
-