

C.N.R. - PROGETTO FINALIZZATO "CONSERVAZIONE DEL SUOLO"
CONVEGNO SULL'AGGIORNAMENTO DELLE METODOLOGIE DI RACCOLTA
ED ELABORAZIONE DEI DATI IDROMETEOROLOGICI
ROMA, 29 E 30 NOVEMBRE 1978

UNA BASE DI DATI IDROLOGICI PER IL PROGETTO
FINALIZZATO "CONSERVAZIONE DEL SUOLO".

B. Biagi*, M. Morandi Cecchi**, O. Salvetti***

* Borsista CNR presso l'I.E.I. di Pisa

** I.E.I. del CNR di Pisa e I.S.I. dell'Università di Pisa

*** Laureando in Scienze dell'Informazione, lavoro tratto dalla
tesi di laurea

RIASSUNTO

Fattore essenziale per l'elaborazione dei dati di natura territoriale, perchè siano adatti all'applicazione di modelli matematici, è, sia la loro raccolta che la loro archiviazione e la loro gestione.

A tale scopo è stata creata una base di dati idrologici che si coordina ad una base di dati morfologici già realizzata per il progetto finalizzato "Conservazione del suolo".

L'esigenza di una raccolta organica di dati di tipo idrologico è sorta dalla necessità di omogeneizzare e di standardizzare il modo in cui si memorizzano tali dati al fine di agevolare l'archiviazione e la gestione da parte di varie unità operative del progetto che possono usufruire di questo strumento per le loro elaborazioni.

La tecnica seguita nella preparazione della base di dati viene descritta in dettaglio. La base di dati è di tipo gerarchico e gli elementi che la compongono sono:

- il bacino, ossia il nodo di massimo livello o radice dell'albero.
- le stazioni, cioè i nodi di livello immediatamente inferiore rispetto al bacino.

- il tipo di informazioni (afflussi, deflussi ecc.)
i cui nodi vengono a costi-
tuire le foglie dell'albero.

Si descrive poi il passaggio dal modello logico al modello fisico.

Si descrivono i formati con cui i dati possono essere caricati nella base di dati; il linguaggio di interrogazione della base di dati ed il loro interprete.

1. INTRODUZIONE

Dall'esigenza di realizzare una raccolta organica di dati di tipo idrologico è sorta la necessità di omogeneizzare e standardizzare la memorizzazione dei dati raccolti dalle Unità Operative del Progetto Finalizzato Conservazione del Suolo per agevolare l'archiviazione e la gestione mediante data-base e per consentire la possibilità di elaborare i dati secondo le esigenze delle varie Unità Operative.

A questo scopo si è preparata una scheda di bacino, di cui si allega una copia, che è stata inviata dalla direzione del Sottoprogetto agli organi di studio e di ricerca interessati a questo lavoro, per ottenere un quadro generale delle zone attualmente sotto osservazione per la registrazione di dati idrologici.

Sulla base di queste schede di bacino si è iniziato l'immagazzinamento e la gestione dei dati che sono pervenuti.

Le caratteristiche morfologiche e idrogeologiche generali suggeriscono che i dati vengano organizzati con una archiviazione suddivisa per bacini idrografici. Le entità che, immediatamente dopo il bacino, assumono importanza sono le stazioni di rilevamento dei dati, le quali, nella loro globalità, concorrono a definire le caratteristiche idrografiche di un bacino. La struttura logica che quindi più si adatta al trattamento di informazioni idrologiche è la struttura ad albero a due livelli, che nel nostro caso è estesa con ulteriori nodi aggiungendo ad ogni stazione i tipi di dati ad essa associati.

Tale gerarchia è stata fisicamente realizzata usando la struttura a lista, anche se, per particolari esigenze, è stata ampliata e trasformata, rendendo più complessi i rapporti fra le informazioni.

In realtà i dati che vengono memorizzati non sono solamente di natura idrologica, ma anche meteorologica e in genere di tutti

i tipi ritenuti importanti da coloro che fanno ricerca su questo argomento.

Oltre alle schede di bacino vengono archiviati i seguenti tipi di dati: afflussi, altezza di falda, deflussi, eliofania, evaporazione, evapotraspirazione, neve, portate di drenaggio, portate di sorgente, pressione atmosferica, radiazione globale, temperatura dell'aria, temperatura del suolo, trasporto solido, umidità dell'aria, umidità del suolo, vento.

2. LE STRUTTURE DEL DATA-BASE

Come è già stato accennato precedentemente, la struttura logica che lega in modo più naturale le informazioni di tipo idrologico è quella ad albero.

Gli elementi gerarchici che la formano sono:

- il bacino, ossia la radice dell'albero;
- le stazioni;
- i tipi di dati (afflussi, deflussi, ecc.).

Lo schema relativo alla struttura sopra descritta è quello riportato in fig. 1.

Per costruire una struttura fisica flessibile ed efficiente, cioè che unisca una buona velocità di ricerca alla possibilità di aggiunte, modifiche e cancellazioni, si è pensato di connettere i vari elementi, per ogni livello, tramite dei puntatori e collegare analogamente anche elementi di diverso livello ma logicamente congiunti.

Al fine di non imporre limitazioni sulla quantità dei tipi di dati da inserire, si è introdotto un "livello di switch" immediatamente dopo il livello delle stazioni (fig. 2).

In questo modo diverse stazioni possono essere collegate a tipi di dati diversi e le nuove aggiunte o eliminazioni si riflettono esclusivamente sulla gestione delle varie liste.

Dal momento che nel processo di inserimento di informazioni nella struttura si deve logicamente procedere a ritroso dalla base al vertice della gerarchia, tale processo implica un modello logico del tipo

DATI → STAZIONE → BACINO

in cui le tre entità, pur distinte, diventano inseparabili, è in-

fatti impensabile supporre un inserimento di informazione dal vertice che si fermi ad un qualunque livello intermedio senza espandersi fino all'ultimo livello, cioè ai dati veri e propri.

I dati vengono a loro volta caratterizzati dai parametri "tipo" (aff., def. ecc.) e "scansione" (mensile, giornaliera, ecc.); è stato introdotto, quindi, un altro livello, dopo quello di "switch" contenente due tipi di informazione:

- scansione dei dati, per ogni cammino

BACINO → STAZIONE → TIPO DI DATI

- anni e mesi presenti BACINO → STAZ. → TIPO DI DATI

In questo modo ogni blocco di dati per un certo anno è individuato da una quadrupla:

BACINO, STAZIONE, TIPO, SCANSIONE

Una soluzione di questo tipo prevede un aggiornamento dinamicamente continuo della struttura, essendo impossibile anticiparne e definirne i contorni e che rende il sistema il più possibile autoregolato lasciando all'utente la massima libertà di gestione.

L'impostazione del data base è basata quindi sui seguenti due stadi di lavoro:

- 1) definizione e costruzione di una struttura base non modificabile dall'utente nei suoi aspetti fisici, ma accessibile per l'aggiornamento, la cancellazione, la correzione e la lettura di informazioni;
- 2) Costruzione di un software che modifichi la struttura base a seconda delle richieste dell'utente (linguaggio di interrogazione) senza che questo debba conoscere l'organizzazione fisica dei dati e le strutture di memorizzazione.

3. IL LINGUAGGIO DI INTERROGAZIONE

Mediante il linguaggio di interrogazione viene data la possibilità all'utente di ottenere:

- 1 - generalità sulle informazioni contenute nel data base
- 2 - inserzione di dati
- 3 - correzione di nomi o dati errati già memorizzati
- 4 - cancellazione di informazioni
- 5 - lettura di dati
- 6 - uscire dal sistema.

1 - GEN fornisce informazioni generali sulla quantità e il tipo dei dati presenti nel data base.

Le richieste possibili riguardano sia i dati veri e propri che le schede di bacino e per queste ultime è possibile specificare la richiesta di:

- dati generali sul bacino
- morfometria
- caratteristiche geolitologiche
- caratteristiche idrogeologiche
- caratteristiche pedologiche e vegetazionali
- sistemi di misura.

Per quanto riguarda invece l'insieme dei dati gestiti dal data-base si possono conoscere:

- nomi dei bacini, delle stazioni e tipi di dati che al momento attuale sono memorizzati;
- nomi di tutte le stazioni di un bacino;
- i tipi di dati e gli anni per ogni tipo per tutte le stazioni di un certo bacino
- i tipi di dati e gli anni relativi per stazioni di diversi bacini.

2 - INS permette l'inserzione dei dati nel sistema.

L'inserzione può avvenire, a scelta dell'utente, da schede, da terminale o da nastro (attualmente è possibile solo nel primo modo).

Il pacco di schede deve essere fornito di opportune schede di controllo per consentire la memorizzazione di blocchi non omogenei di dati.

3 - MOD permette l'accesso ad un ambiente nel quale è possibile la correzione di nomi e/o dati errati.

La modifica può essere effettuata da terminale o da schede: in entrambi i casi l'utente indicherà prima i dati di identificazione del record errato, quindi il record errato e in ultimo il record corretto.

4 - CAN permette di effettuare cancellazioni.

A causa della gerarchia con cui i dati sono legati, la richie-

sta di cancellazione di un bacino implica la eliminazione di tutte le stazioni, i tipi di dati e i dati stessi ad esso associati. Analogamente accade per la eliminazione di una stazione; se la stazione è l'unica di un bacino, verrà eliminato anche il bacino.

L'utente può anche cancellare solo un tipo di dati per un determinato bacino o un solo anno di dati.

5 - LET serve all'utente per ottenere l'informazione desiderata con varie possibilità di uscita.

L'utente può selezionare una tra le seguenti richieste:

- stampa di dati a terminale o su stampante, in formato originale o standard;
- trasferimento dei dati richiesti in un'area di lavoro dell'utente.

In particolare possono essere richiesti:

- tutti i dati relativi a un bacino;
- tutti i dati relativi a una stazione;
- tutti i dati di un certo tipo;
- tutti i dati di un certo anno.

6 - EOS permette di interrompere il dialogo utente-sistema e tutto viene predisposto per una nuova chiamata del sistema.

4. DESCRIZIONE DEI PROGRAMMI

Il sistema precedentemente descritto è stato realizzato per l'elaboratore IBM 370/158 del Centro di Calcolo del C.N.R. di Pisa (C.N.U.C.E.).

I programmi sono gestiti da una procedura CMS-EXEC che costituisce un intermediario tra l'utente e l'esecuzione delle sue richieste. I dati sono stati suddivisi in archivi parziali omogenei ognuno dei quali ha una sua precisa identificazione che viene gestita dal sistema operativo VM-CMS.

Tutti i files di dati sono memorizzati su dischi magnetici direttamente collegati con l'elaboratore per ottenere la massima velocità di ricerca delle informazioni.

Il linguaggio di programmazione usato è il PL/1 che permette di creare files ad accesso diretto per una migliore e più ve-

loce gestione del sistema.

I programmi realizzati sono:

- un programma PL/1 che crea ed inizializza la struttura gerarchica del data base;
- una procedura CMS/EXEC e un insieme di programmi PL/1 che costituiscono il software del data base per la gestione e l'utilizzo dei dati, il linguaggio di interrogazione, gli aggiornamenti ecc.

E' importante notare che, per poter essere memorizzati, i dati devono essere trasferiti dal loro supporto originale di registrazione su un supporto compatibile col calcolatore e per facilitare il compito di caricamento dei dati da parte dell'utente, sono state ridotte al minimo le restrizioni riguardo ai formati di scrittura dei dati; infatti ogni scheda perforata deve contenere necessariamente solo l'indicazione relativa all'anno e al mese a cui i dati si riferiscono, mentre viene lasciata completa libertà per quanto riguarda il formato di scrittura dei dati stessi.

Il programma provvede poi conoscendo la scansione e il tipo dei dati, a calcolare il loro numero e quindi a memorizzarli opportunamente. Inoltre sarà messa a disposizione dell'utente una piccola guida che gli permetterà di conoscere le unità di misura dei dati, i loro tipi, le schede controllo necessarie per il caricamento e contenente una descrizione delle varie possibilità di gestione dei dati, cioè lettura, inserimento, modifica e cancellazione dei dati.

Per ogni tipo e scansione dei dati è stata anche compilata una lista di formati "standard" per facilitare le successive elaborazioni da parte dell'utente; mediante una opportuna chiamata è prevista la possibilità di ottenere i dati secondo i formati standard oppure nei formati originali.

Qui di seguito in Tabella 1 si dà un esempio per gli afflussi in modo del tutto analogo si codificano i dati per gli altri casi elencati nell'introduzione.

Colonne	Informazioni	Formato
1-4	anno	I4
5-6	n.mesi	I2
7-8	dati	F6.1
1-4	anno	I4
5-6	mese	I2
7-80	dati	F5.1
1-4	anno	I4
5-6	mese	I2
7-8	giorno	I2
9-80	dati	F4.1
1-4	anno	I4
5-6	mese	I2
7-8	giorno	I2
9	AoBoC	
10-80	dati	F4.1
1-4	anno	I4
5-6	mese	I2
7-8	giorno	I2
9	A,B,C,D,E,F	
10-80	dati	F4.1
1-4	anno	I4
5-6	mese	I2
7-8	giorno	I2
9	A+S	
10-80	dati	F4.1

(II scheda del mese=mese+12)

(II del giorno=giorno+31)

a seconda che si tratti della prima, seconda o terza scheda per quel giorno.

BIBLIOGRAFIA

- 1 Morandi Cecchi M. Costruzione di banca dati ed elaborazione delle misure idrologiche. Convegno su metodologie di acquisizione dati e apparecchiature per la misura di grandezze idrologiche e marittime. Portici 10 Giugno 1977.
- 2 Monaci I. Un sistema informativo precursore di un Data Base per bacini montani e per bacini a morfologia parzialmente matura. Tesi di laurea, anno accademico 1975/76, I.S.I., Facoltà di S.M.F. e N., Università degli Studi di Pisa.
- 3 Martin J. Computer Data-Base organization Englewood Prantice Hall 1975.
- 4 Date C. J. An introduction Data-Base Systems. Addison Wesley (1975).
- 5 Appleton D. S. What data base is'n't Datamation Vol. 23 pag. 85-92 (1977).
- 6 Fimeron T. R., Henry G. S. Structured analysis for data base design Datamation Vol. 23, n. 11 (1977).
- 7 Curtice R. M. Integrity in data base systems Datamation Vol. 23 pag. 64 (1977).



