

# IL TEMPO È UN DONO DI DIO: PERCHÉ SPRECARLO?

## PROPOSTA DI UN FORMATO COMUNE PER LA DIFFUSIONE DI DATI IN AMBITO AGROMETEOROLOGICO

Luigi Mariani, Tommaso Maggiore

*Università degli Studi di Milano, Dipartimento di Produzione Vegetale, via Caloria, 2 - 20123 Milano  
E\_mail: lugi.mariani@unimi.it*

### Abstract

La disponibilità dei dati meteorologici su supporto magnetico è una delle più importanti conquiste degli anni più recenti in quanto evita agli utenti trascrizioni dal formato cartaceo, noiose e foriere di errori. Tuttavia già si profila un problema di accessibilità ai nuovi formati elettronici, a causa della varietà dei formati dei dati che servizi operativi o strutture di ricerca propongono ai loro utenti. Esiste un modo per ovviare a tale problema? Questo il tema che ci si propone di discutere in questa presentazione.

### Il problema

Molte attività nel settore della meteorologia applicata comportano la preliminare assimilazione di dati provenienti da più fonti, su territori spesso anche vasti (Daley, 1991). Per inciso un'esauriente attività di raccolta rappresenta il presupposto per un controllo di qualità efficace (Mariani, 2006) e per una spazializzazione realistica dei dati, anche tramite metodi geostatistici (Saucier, 1989; Goovaerts, 1997).

In particolare in agrometeorologia si scambiano dati di tipo fisico e biologico (van Keule e Wolf, 1986) e tuttavia in questa sede ci limiteremo alle sole variabili fisiche atmosferiche, partendo dalla seguente constatazione: di fronte alla necessità di raccogliere dati provenienti da più fonti, l'utente dedica una quota sempre maggiore del tempo ad interventi sui dati stessi tesi a rendere fruibili dati che gli enti di servizio o di ricerca erogano su supporto magnetico utilizzando formati alquanto diversi fra loro. Gli interventi dell'utente consistono nello specifico nella stesura di software di trascodifica ovvero in attività di taglia-incolla classicamente svolte in ambiente excel.

Occorre premettere che la fornitura su supporto magnetico è comunque un enorme passo avanti rispetto ai dati distribuiti in formato cartaceo. Tuttavia i vantaggi di tale progresso vengono spesso vanificati dalla varietà dei formati di cui sopra, varietà che si traduce non solo in perdite di tempo ma anche nell'incremento della possibilità di errori.

La varietà dei formati è spesso frutto della giusta ambizione dei programmatori di adeguarsi alle esigenze del pubblico, il che spiega l'incapsulamento dei dati all'interno di "gabbie" complesse e a volte a geometria variabile, rispetto alle quali la messa a punto di una procedura software di lettura automatica si rivela spesso ardua, costringendo così a "pre-processing manuali".

Ad esempio lavorando in un'area del mondo che non è l'Italia, uno degli autori si è confortato di recente con files mensili raccolti da un gruppo di ricerca e che erano a geometria variabile; in sostanza le variabili di mese in mese venivano collocate in colonne sempre diverse e con header sempre diversi; addirittura ci si trovava ad un certo punto ad assistere al cambiamento del formato ora (UTC nei primi files, ora locale nei successivi). Capirete che di fronte ad un tale problema

l'unica soluzione si sia rivelata quella di operare manualmente sui files, omogeneizzandoli alla meglio in modo da renderli digeribili per una procedura automatica.

Un altro esempio è costituito da un lavoro di tesi che mira all'applicazione di modelli fenologici e di produttività della vite in Oltrepò Pavese. In questo caso si utilizzano dati provenienti da cinque reti diverse (Arpa Lombardia, Arpa Emilia Romagna, Regione Piemonte, Co.Pro.Vi., CNA-cma), con impieghi di tempo veramente rilevanti; e un problema analogo si è manifestato nell'attività di assimilazione di dati radiometrici e termometrici propedeutica al lavoro di Confalonieri et al (2008), presentato in questo convegno.

Si tenga conto che tale problematica interessa parecchie attività professionali: si pensi a esempio al settore degli studi professionali impegnati a livello nazionale o locale in attività di tipo agronomico o di analisi ambientale (analisi vocazionali, valutazione delle esigenze idriche delle colture finalizzate alla pianificazione irrigua, valutazioni d'impatto, ecc.) in cui le variabili chiave sono quelle meteo-climatiche.

Si deve peraltro rimarcare che la problematica descritta esplose inevitabilmente in una nazione in cui la frammentazione delle reti ha raggiunto livelli abnormi ed il livello d'attenzione alla qualità del dato è ancora in larga misura insufficiente.

### Alcune possibili soluzioni

Da qui l'idea di avviare una discussione con i colleghi con i seguenti obiettivi:

1. verificare se il problema è effettivamente percepito come rilevante
2. verificare se e come sia possibile superarlo
3. verificare in particolare l'utilità di un formato di minima targato Aiam da utilizzare per l'interscambio dei dati..

Mentre i punti 1 e 2 vengono lasciati alla discussione che ci si augura ampia ed esauriente, per quanto riguarda il punto 3 un possibile formato potrebbe essere il seguente:

formato testo senza caratteri di controllo ad esclusione di eventuali tab

- gruppo data
- gruppo ora: minuti
- dato
- eventuale codice di validità (cv).

I dati mancanti dovrebbero essere sostituiti da un codice di sostituzione numerico efficace in relazione alla misura considerata (es: -9999 o 9999).

Esempio:

```
yyyymmdd hh.mm tx cv1 tn cv2 rr cv3  
20070301 00.00 24.5 1 16.6 1 0.0 1  
20070302 00.00 25.2 1 21.0 1 12.3 1  
20070303 00.00 9999 1 9999 1 9999 1
```

Valide alternative sono costituite da formati il cui uso a livello internazionale è già consolidato. Fra questi si segnalano ad esempio:

- il formato Synop del WMO (<http://wxp.unisys.com/Appendices/Formats/SYNOP.html>)
- Il formato della banca dati internazionale ECAD (<http://eca.knmi.nl/dailydata>)
- Il formato di interscambio del progetto MAP (<http://www.map.meteoswiss.ch/>)
- Il formato di input del modello Cropsys. ([www.bsyse.wsu.edu/cropsyst/](http://www.bsyse.wsu.edu/cropsyst/))

Ognuno dei formati discussi presenta ovviamente pregi e difetti che andrebbero soppesati attentamente. Tuttavia un'ampliamento dell'uso dei dati da parte dell'utenza nazionale ed internazionale passa necessariamente attraverso la semplificazione e la normalizzazione.

## I metadati

Da non trascurare è inoltre il cruciale tema dei metadati. Ai dati di ogni stazione occorrerebbe infatti che venissero affiancate non solo informazioni sulla qualità degli stessi (es: codici di validità, particolarmente importanti per i dati più "delicati" come quelli radiometrici, igrometrici o pluviometrici) ma anche le informazioni in merito a storia, localizzazione attuale e passata, tipologia degli strumenti, interventi di controllo effettuati, ecc.

Anche a livello di metadati sarebbe più che mai auspicabile la definizione di un formato comune che comprenda ad esempio:

- documentazione fotografica su stazione e sito (foto in primo piano e foto "di ambiente")
- collocazione su idonea base cartografica
- tipo di strumenti e relative caratteristiche metrologiche
- note storiche in merito a collocazione passata, modifiche subite dalla strumentazione, ecc.
- note sullo stato di manutenzione (frequenza dei controlli, problemi riscontrati, ecc.).

## Conclusioni

Come si può rilevare da queste brevi note la tematica è ampia e meritevole di un approccio ragionato e coordinato, anche in virtù del fatto che è ipotizzabile che normative di qualità analoghe a quelle in vigore per la strumentazione analitica vedano presto la loro estensione al settore delle misure meteorologiche.

Occorre infine sottolineare che non dev'essere ambizione dell'Aiam quella di sostituirsi ai servizi in quello che è un ambito di loro competenza, con tutti gli onori e oneri relativi. Pertanto sarebbe più che mai auspicabile che un "formato italiano comune" nascesse da un dibattito fra i servizi (e, ultimata la stesura di questo abstract, ci sorge il dubbio che magari una tale soluzione sia già stata trovata da qualche tavolo di lavoro interservizi e che noi non se ne sia a conoscenza, il che a maggior ragione giustifica una discussione sul tema).

## Bibliografia

- Confalonieri et al., 2008. Sviluppo di un prototipo per le previsioni di resa a larga scala basato sul modello di simulazione warm e su dati di concentrazione di azoto nella pianta derivati da satellite, Aiam 2008*
- Daley R., 1993. Atmospheric data analysis, Cambridge Univ. Press, 456 pp*
- Goovaerts Pierre, 1997. Geostatistics for water resources evaluation, Oxford Univ.Press, 483 pp*
- van Keulen H, Wolf J. (Eds), 1986. Modelling of agricultural production, weather, soils and crops, Pudoc, Wageningen, 479 pp*
- Mariani L., 2006. Alcuni metodi per l'analisi delle seriestoriche in agrometeorologia, Rivista Italiana di Agrometeorologia 48 - 56(2) 2006, 48.*
- Saucier W.J., 1989. Principles of meteorological analysis, Dover, 438 pp*