

# CLIMATICA: UN SOFTWARE PER LA GESTIONE ED ELABORAZIONE DELLE INFORMAZIONI CLIMATICHE

## CLIMATICA: A SOFTWARE FOR CLIMATIC INFORMATION MANAGEMENT

Francesco Danuso \*, Mariangela Sandra

Dipartimento di Scienze Agrarie e Ambientali – Università di Udine, via delle Scienze 208, 33100 Udine (Italy)

\* Corresponding author, tel. +39 432 558614, fax +39 432 558603, e-mail francesco.danuso@uniud.it

Ricevuto 5 ottobre 2005, accettato 28 aprile 2006

### Riassunto

Viene presentato *CLIMATICA*, un sistema informatico per la gestione dei dati climatici mensili, giornalieri e orari. Il software dispone di funzionalità diverse tra cui la possibilità di controllo e validazione dei dati, la conversione del formato delle date e delle unità di misura delle variabili, la produzione di indici statistici. Consente di effettuare operazioni generali di gestione sui dataset climatici (fusione per righe o per colonne, riduzione della dimensionalità di un dataset con calcolo di statistiche e tabelle), sui metadati (inserimento e modifica), sulle variabili (cancellazione, generazione e ricalcolo) e sui record (ricerca, ordinamento, cancellazione, inserimento). Sono disponibili inoltre procedure per l'interpolazione e ricostruzione dei dati mancanti come la regressione lineare multipla, la regressione non-lineare, la creazione di reti neurali e le "moving statistics" (media mobile, ecc.). *CLIMATICA* implementa anche il generatore climatico *Climak* con il quale è possibile stimare i parametri climatici a partire dai dati storici e generare variabili meteo per la valutazione dei rischi o la creazione di scenari climatici. Le operazioni di gestione dei dataset possono essere effettuate sia attraverso finestre di dialogo sia con l'uso di comandi. E' disponibile l'automazione delle procedure ripetitive attraverso la creazione di *script*. *CLIMATICA* presenta caratteristiche di estrema flessibilità che ne permettono un impiego personalizzato in relazione a specifiche esigenze.

**Parole chiave:** software, dati climatici, qualità del dato, database management, generatore climatico.

### Abstract

*CLIMATICA, a software for management of climatic data is presented. The software has different features like the possibility of control and validation of data, the conversion of date formats and of the units of the variables, the calculation of statistical indexes. CLIMATICA can perform management operations on climatic dataset (fusion of more datasets by rows or by columns, reduction of the dataset dimensionality with estimation of statistics), on metadata (input and modification), on variables (deletion, generation and re-calculation) and on records (search, sorting, deletion and input). There are also procedures for interpolation and reconstruction of missing data using multiple linear regression, non-linear regression, neural networks and moving statistics (moving average, etc.). CLIMATICA also includes the weather generator Climak for the estimation of climatic parameters from historical data and for the generation of meteorological variables to evaluate risks or to create climatic scenarios. The management of datasets can be made either by dialog windows or by commands, which also allows the automation of procedures. CLIMATICA presents a wide flexibility that permits to be adapted to specific requirements.*

**Keywords:** software, climatic data, data quality, database management, weather generator.

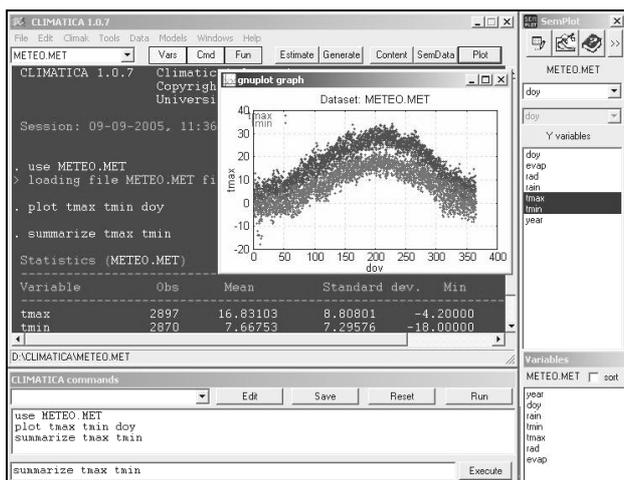
### Introduzione

Il crescente impiego delle informazioni climatiche nella modellistica e per la previsione e pianificazione degli agroecosistemi ha determinato la necessità di disporre di strumenti informatici integrati capaci di trattarne con facilità i dati al fine della loro esplorazione, validazione, ricostruzione e gestione. Tuttavia, al momento attuale, non sono numerosi i software disponibili che affrontano in modo sistematico ed integrato il problema. Come esempi, di notevole interesse, si ricordano tuttavia ClimLab2000 (Tanco *et al.*, 2000) e AnClim (Štěpánek, 2003). Tali software, però, sebbene inglobino numerose metodologie di analisi e gestione dei dati climatici, non offrono la possibilità di effettuare controlli sulla qualità del dato e, qualora questo sia mancante o anomalo, la

possibilità di ricostruirlo. E' noto inoltre che operare con dati climatici affetti da errori casuali o sistematici può portare ad errori anche consistenti sui risultati ottenuti con l'impiego dei modelli (Rana *et al.*, 2004).

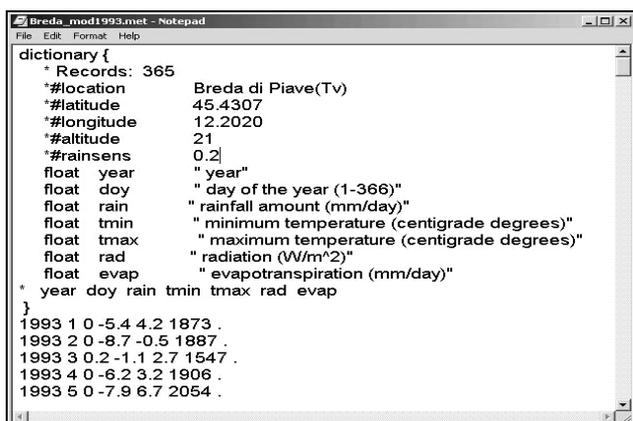
In questo contesto, è stato realizzato il software *CLIMATICA*, sviluppato presso il Dipartimento di Scienze Agrarie e Ambientali dell'Università di Udine.

*CLIMATICA* lavora in ambiente Windows, presenta un'interfaccia utente di tipo grafico (figura 1) ed è disponibile gratuitamente via Web. Offre la possibilità di effettuare controlli di qualità del dato, nonché di ricostruire dati mancanti con modalità diverse (reti neurali, statistiche mobili, regressione) a seconda delle necessità dell'utente. Ingloba il generatore stocastico del clima



**Fig. 1** – Ambiente di lavoro di *CLIMATICA*. Sono attivate la finestra principale con i risultati delle elaborazioni, la finestra dei comandi (in basso), la finestra per la produzione dei grafici (a destra), un grafico e la lista delle variabili presenti nel dataset corrente (in basso a destra)

**Fig. 1** – *CLIMATICA* framework. The shown dialogs are the main window with results of data processing, the command dialog (bottom), the plotting window (right) and the list of variables of the current dataset (bottom-right)



**Fig. 2** - Esempio di file meteo (.met) di *CLIMATICA*.

**Fig. 2** - Example of a *CLIMATICA* meteo file (.met).

*Climak* (Danuso, 2002) e presenta le stesse possibilità di gestione dati, analisi statistica, interpolazione e grafica dell'ambiente di modellazione SEMoLa (Danuso, 2003).

*CLIMATICA* è attualmente in fase di evoluzione sia per incrementarne che per migliorarne le funzionalità.

## Gestione dei dataset climatici

### Tipologie di file e operazioni sui dataset

*CLIMATICA* tratta dataset climatici in formato **csv** (Comma Separated Value) e in formato specifico standard (**dct**). I dataset di tipo standard possono avere estensione **dct** (dataset generico), **met** (file meteo storico), **red** (file contenente i parametri climatici stimati da *Climak* per un dato territorio) e **gen** (file di dati clima-

tici generati). Le ultime due tipologie di file sono create dal generatore climatico *Climak*, incluso in *CLIMATICA*. I dataset di dati climatici fanno riferimento ad una specifica stazione mentre quelli di tipo **red** riportano parametri climatici georiferiti per più stazioni di una regione.

I dataset in formato standard sono file di testo costituiti da un'intestazione e dai dati. Nell'intestazione sono descritte le variabili e vengono riportati i metadati (figura 2). La sezione dati del dataset corrisponde ad una tabella in cui le variabili climatiche sono riportate sulle colonne e le diverse osservazioni nel tempo sulle righe.

È possibile convertire i dataset dal formato standard al formato **csv** (utilizzabile direttamente dal foglio elettronico) ed eseguire operazioni di gestione sui dataset come il merge, l'append (per fondere assieme più dataset) e il collapse (per convertire il dataset corrente in un altro contenente medie, somme, mediane, ecc.). Viene elaborato un solo dataset (multiannuale) per volta, ma è consentito tuttavia riunire più dataset in un unico file.

### Trattamento dei metadati

I metadati di una stazione meteorologica (latitudine, longitudine, altitudine, ecc.) possono essere trattati da una specifica finestra di dialogo, operando modifiche sui nomi e sui valori dei metadati già esistenti o aggiungendone di nuovi, scegliendo tra quelli proposti.

L'impostazione dei metadati, nonostante sia opportuna, non è strettamente richiesta, anche se alcune procedure richiedono obbligatoriamente tali informazioni (ad esempio, la latitudine) per poter operare.

### Comandi e automazione delle procedure

Le elaborazioni possono essere ottenute sia utilizzando le finestre di dialogo sia impiegando i comandi. I comandi sono formati da una parola che identifica l'azione richiesta e da una serie di argomenti che indicano gli oggetti da elaborare (ad esempio, le variabili, i file) e le eventuali opzioni. I comandi attualmente disponibili sono riportati in tabella 1.

È possibile, operando con i comandi, memorizzarne delle sequenze (*script*) da poter ripetere successivamente in modo automatico. Gli *script* vengono salvati come file di testo con estensione **cmf**. A titolo d'esempio, si riporta uno *script* di comandi che consente di calcolare una somma termica con base 10 e di produrne il grafico relativo in funzione del giorno dell'anno (figura 3):

```
use Breda_mod1993.met
gen tmed=(tmin+tmax)/2
gen DST=(tmed-10)*abs(tmed>10)
gen ST=sum(DST)
plot ST doy
```

Breda\_mod1993.met è il nome del file meteo da impiegare; tmed, DST e ST sono variabili generate nel dataset corrente. doy è il giorno dell'anno, variabile già presente nel dataset.

**Tab. 1** - Comandi di *CLIMATICA***Tab. 1** - *CLIMATICA* commands

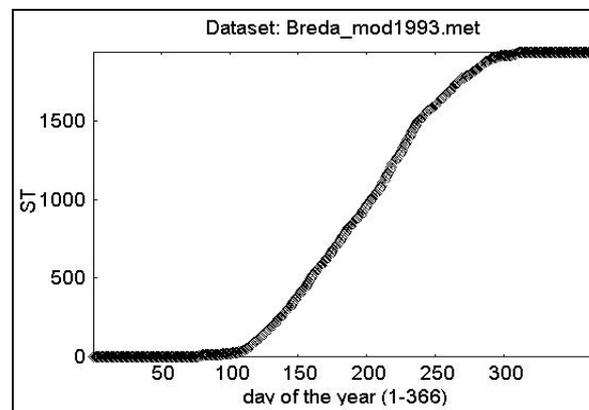
<b>Creazione di modelli empirici basati sui dati</b>	
net	crea, testa e salva reti neurali
regress	regressione lineare multipla
nlreg	regressione non lineare
<b>Gestione dei dataset</b>	
use	carica un dataset da file
merge	aggiunge colonne al dataset corrente
append	aggiunge righe al dataset corrente
collapse	genera un dataset ridotto di statistiche
save	salva il dataset corrente
<b>Visualizzazione dei dati</b>	
describe	visualizza i metadati del dataset caricato
header	elena, modifica e inserisce metadati
list	elena i valori delle variabili come tabella
listc	elena i dati come righe complete
<b>Gestione dei dati</b>	
drop	elimina variabili/osservazioni dal dataset
keep	conserva le variabili/osservazioni indicate
generate	calcola nuove variabili
replace	ricalcola intere variabili o singole osservazioni
rename	cambia il nome delle variabili
shift	fa slittare i dati di una variabile
sort	ordina le righe del dataset corrente
<b>Variabili ambiente</b>	
scalar	gestione di variabili ambiente scalari
string	gestione di variabili ambiente stringa
matrix	gestione di variabili ambiente matrici
<b>Statistiche</b>	
correlate	coefficienti di correlazione tra variabili
cumulate	calcola distribuzioni empiriche cumulate
movistat	statistiche mobili (media, mediana, ecc.)
rank	genera la variabile rango statistico
summarize	calcola statistiche descrittive
table	crea una tabella con statistiche
<b>Grafici</b>	
plot	crea grafici per visualizzare i dati
<b>Utilità</b>	
derive	differenziazione simbolica
dq	gestione della qualità del dato
eval	calcolo espressioni numeriche (calcolatrice)
for	super-comando per ripetere altri comandi
markest	stima matrice transizione catena di Markov
marksim	genera sequenze dati con catena di Markov
play	suona file MIDI musicali
sonify	sonifica i dati generando file MIDI musicali
<b>Guida</b>	
help	guida per l'uso di <i>CLIMATICA</i>

**Operazioni su osservazioni/variabili di un dataset**

Il software permette di operare sia sulle singole osservazioni che sulle variabili. E' possibile, ad esempio, eliminare o selezionare righe, cancellare o creare nuove variabili con espressioni matematiche definite dall'utente, sia impiegando una finestra di dialogo sia con comandi diretti.

Un aspetto particolare per l'uso del software riguarda i nomi delle variabili. *CLIMATICA* accetta per le variabili meteorologiche qualsiasi nome costituito da un massimo di dodici caratteri. Tuttavia, riconosce e tratta in modo

preferenziale i nomi standard riportati in tabella 2. E' utile, a tal proposito, di rinominare (*rename*) le variabili secondo i nomi standard, per semplificare e ottimizzare le procedure di elaborazione. Molte di esse, infatti, risultano agevoli o applicabili solo se le variabili meteorologiche presentano nomi riconosciuti.

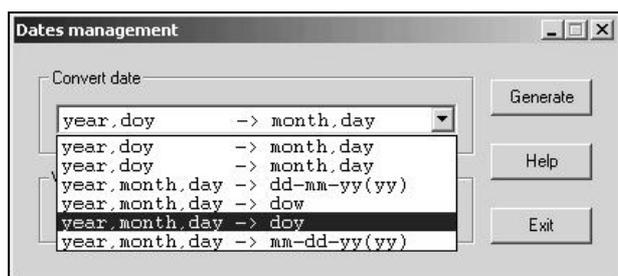


**Fig. 3** - Andamento dell'indice ST di somma termica  
**Fig. 3** - Trend of the temperature summation index ST

**Tab. 2** - Nomi delle variabili meteorologiche riconosciuti da *CLIMATICA* e tipo di rappresentazione interna (int=intero, 4 byte; float=doppia precisione, 8 byte; str##=stringa)

**Tab. 2** - Names of meteorological variables recognized by *CLIMATICA* and types of internal representation (int=integer, 4 byte; float=double precision, 8 byte; str##=string)

Variabile	Descrizione	Tipo
year	Anno	int
doy	Giorno dell'anno (1-366)	int
day	Giorno del mese (1-31)	int
month	Mese (1-12)	int
hhmm	Ora e minuto	str5
ymd	Anno-mese-giorno	str10
dmy	Giorno-mese-anno	str10
mdy	Mese-giorno-anno	str10
dow	Giorno della settimana	int
jdate	Giorno giuliano	int
dec	Codice decadico (1-36)	int
moonph	Fase lunare	int
rain	Precipitazioni	float
tmean	Temperatura media aria	float
tmax	Temperatura massima aria	float
tmin	Temperatura minima aria	float
rg	Radiazione solare globale	float
etr	Evapotraspirazione riferim.	float
winds	Velocità del vento	float
windd	Direzione del vento	float
rhmin	Umidità relativa minima	float
rhmean	Umidità relativa media	float
rhmax	Umidità relativa massima	float



**Fig. 4** – Finestra per la gestione delle date in *CLIMATICA*. A seconda del tipo di data già presente nel dataset, vengono proposte solo le conversioni ammesse

**Fig. 4** – *CLIMATICA* dialog for dates management. Only the allowed conversions, in relation to the types of date already present in the dataset, are proposed.

*CLIMATICA* permette anche di convertire facilmente il formato delle date come, ad esempio, da anno-mese-giorno a giorno dell'anno (figura 4) e di trasformare le unità di misura per le diverse variabili meteorologiche, suggerendo quelle possibili.

### Controllo della qualità dei dati climatici

Il controllo della qualità dei dati consiste nella loro valutazione in relazione alla validità del contenuto informativo ai fini del loro impiego. Tale operazione, preliminare e fondamentale per ogni elaborazione successiva, viene ottenuta attraverso l'impiego di procedure diverse che l'utente può selezionare dal menu **Tasks**.

In bibliografia sono presenti numerosi studi sulla qualità del dato (ad esempio, Reek *et al.*, 1992; Meek *et al.*, 1994; Lanzante, 1996; Pavan *et al.*, 2003; Song *et al.*, 2004) che forniscono metodologie per la valutazione della qualità. Tuttavia, le classificazioni adottate risultano poco standardizzate, eterogenee e con terminologie spesso non chiaramente definite. I concetti che più comunemente ricorrono relativamente alla qualità dei dati sono (in inglese): good, suspect, missing, wrong, invalid, anomalous, outlier. Si è ritenuto quindi necessario operare una sistematizzazione di tali concetti, in relazione all'esigenza di conservare e utilizzare l'informazione dei dati stessi.

In *CLIMATICA* il dato meteorologico viene classificato in base a due criteri:

1. **indipendenza** o originalità dell'informazione, per cui il dato può essere classificato come P (primario), C (calcolato o derivato), oppure M (mancante);
2. **qualità** dell'informazione: B se ritenuto valido, S sospetto di invalidità, W non valido.

Le codifiche adottate per codificare i dati in relazione ai criteri di indipendenza e qualità dell'informazione sono riportate in tabella 3.

Si fa notare come il concetto di "suspect" sia simile a quello di "outlier", molto diffuso in statistica. Non viene dato un giudizio di merito, ma semplicemente si dichiara che il dato o non è stato valutato o è distante dal centroide della distribuzione e quindi sembra non appartenere alla stessa.

La tipologia di dato (misurato/ricostruito) ed il suo grado di attendibilità sono mantenuti, per ciascuna variabile

**Tab. 3** – Codifiche per la qualità e indipendenza dei dati.

**Tab. 3** – *Data coding for independence and quality.*

Codice	Criterio di indipendenza
P	Dato primario, informazione indipendente, valore di input, misura diretta. E' il contrario di "calcolato".
C	Dato calcolato, ricostruito, informazione dipendente. Alternativo a "primario".
M	Dato mancante. Indicato con un punto. Viene automaticamente trascurato nei calcoli.

Codice	Criterio di qualità
G	"Good": dato controllato e che ha superato tutti i test di qualità oppure dichiarato "valido" dall'utente.
S	"Suspect": dato non controllato oppure controllato ma che non ha superato almeno un test e quindi è sospettato di essere invalido. Viene utilizzato nelle procedure di calcolo come i dati G, ma con emissione di "warning".
W	"Wrong": dato aberrante, invalido, sbagliato. Un dato può assumere la condizione di "wrong" solo per l'esplicita dichiarazione dell'utente. Viene automaticamente trascurato nei calcoli.

climatica, attraverso l'impiego di una variabile di supporto (`_dq`) gestita automaticamente dal comando `dq` che permette la sua creazione o cancellazione, nonché l'elencazione dei dati per tipologia qualitativa e l'attribuzione diretta del codice di qualità da parte dell'utente. La variabile `_dq` è costituita da una stringa di caratteri formata da tanti caratteri quante sono le variabili del dataset. Ciascuno di questi caratteri viene codificato per gli otto bit che lo compongono in modo da mantenere le informazioni sulla qualità del dato e sulla tipologia dei controlli che ogni dato ha subito.

### Tipologie di controllo

Esistono tipologie di controllo per il dataset (verifica dei nomi riconosciuti e della presenza di dati mancanti), per le date (correttezza e ordine cronologico) e per i dati. Le tipologie di controllo per i dati sono classificate in cinque categorie:

1. controllo di range (*range check*): basato sul range ammesso per la variabile;
2. controllo di step (*step check*): basato sulla variazione ammessa tra un dato e il seguente nella successione temporale di ciascuna variabile;
3. controllo di persistenza (*persistence check*): basato sul numero massimo ammesso di valori uguali consecutivi nella serie temporale;
4. controllo di coerenza (*consistency check*): basato sulla probabilità che i residui dato-modello appartengano ad una distribuzione comune. I residui possono essere calcolati sia con un modello bivariato sia multivariato, basato sullo stesso dataset (coerenza interna) o basato su altri dataset di stazioni vicine (coerenza esterna). Il modello può anche tenere conto degli aspetti di autocorrelazione o di correlazione incrociata (*cross-correlation*);

5. controllo di omogeneità (*homogeneity check*): tende a valutare la presenza di discontinuità nella sequenza temporale dei dati, introdotte spesso da fatti comuni come il cambiamento del sensore o lo spostamento della stazione. Anche in questo caso ci può essere sia un controllo interno che esterno.

Al momento della creazione della variabile `_dq`, i dati non mancanti contenuti nel dataset vengono sempre considerati primari (P) e sospetti (S) per ciascun tipo di controllo. I test automatici di controllo, eventualmente applicati, possono eliminare o meno la qualifica di dato "sospetto" per range, step, persistenza, coerenza o omogeneità. La qualifica di "sospetto" viene però mantenuta anche con un solo test fallito. Se tutti i test vengono superati, la qualifica diviene G (buono). Solo l'utente, in seguito ad un esame diretto dei dati, ha la facoltà di dichiarare invalido (W) un dato sospetto. L'utente può anche decidere d'imperio di dichiarare buono (G) un dato rimasto sospetto, a motivo di qualche test non superato. All'utente rimangono quindi la facoltà e la responsabilità di riclassificare i dati come in relazione alla loro validità o invalidità.

I dati mancanti possono essere ricostruiti secondo diverse modalità. Dopo la ricostruzione i dati vengono classificati come calcolati (C) e possono subire gli stessi controlli di qualità dei dati primari (P).

I codici di qualità vengono automaticamente trattati da *CLIMATICA* durante tutte le fasi di elaborazione come, ad esempio, durante la fusione di più dataset o la creazione di nuove variabili. Operando secondo un principio di cautela, quando si effettua il calcolo di nuove variabili, i dati generati avranno una codifica qualitativa pari alla peggiore qualità tra quelle dei dati impiegati nel calcolo.

Nel calcolo di indici statistici o di modelli di regressione o di reti neurali, i dati mancanti (M) e invalidi (W) vengono trascurati come non esistenti. I dati sospetti (S) vengono impiegati normalmente, ma il programma avverte l'utente che si stanno impiegando dati sospetti.

Attualmente sono disponibili procedure per controllare il numero di dati mancanti e il nome delle variabili riconosciute (**check - dataset**), per normalizzare il simbolo del dato mancante e per controllare le date presenti nel file meteorologico di lavoro (**check - dates**). Il dialogo ottenibile dalla voce di menu **range-step-persistency** permette di effettuare i relativi controlli. I controlli di coerenza e di omogeneità vanno effettuati, al momento, impiegando i comandi di base (tabella 1) con procedure manuali, anche automatizzabili tramite *script*.

### Calcolo di variabili e indici statistici

*CLIMATICA* permette di analizzare statisticamente i dati climatici con l'ausilio di procedure statistiche, come il calcolo di media, deviazione standard, minimo e massimo, con approccio convenzionale o *biveight* (figura 5). Possiede inoltre un ambiente di calcolo che dispone di numerose funzioni matematiche (scalari e matriciali), nonché di funzioni per la generazione di numeri casuali (da distribuzione uniforme, normale, beta, gamma, esponenziale, triangolare, lognormale, logistica, normale bivariata, Gumbel I, Gumbel II e Weibull).

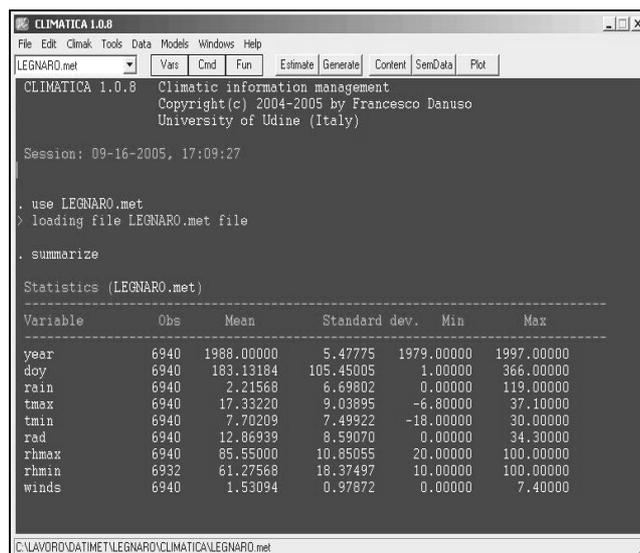


Fig. 5 – Esempio di analisi statistica con *CLIMATICA*: vengono calcolati il numero di osservazioni, la media, la deviazione standard, il minimo e il massimo per ogni singola variabile

Fig. 5 – Example of statistical analysis with *CLIMATICA*: number of observations, mean, standard deviation, minimum and maximum are calculated for each variable

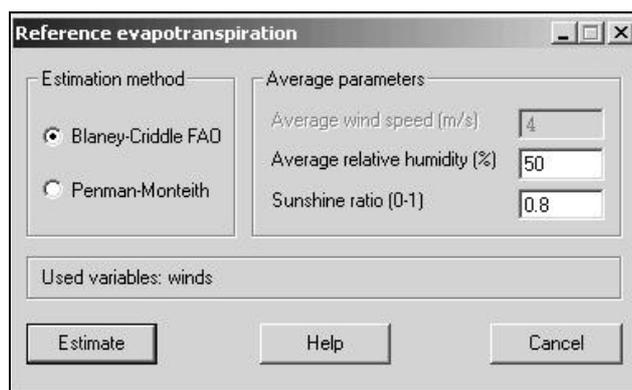


Fig. 6 - Finestra di dialogo per il calcolo dell'evapotraspirazione di riferimento. E' possibile scegliere tra la formula di calcolo di Blaney-Criddle-FAO, che richiede temperatura, velocità media del vento, umidità relativa media dell'aria e l'elofania relativa (qualora queste ultime due variabili non siano presenti nel dataset iniziale, vengono utilizzati dei valori costanti modificabili dall'utente) e quella di Penman-Monteith, che utilizza temperatura, radiazione globale, velocità del vento ed umidità relativa media dell'aria

Fig. 6 - Dialog for calculation of the reference evapotranspiration. It is possible choosing between Blaney-Criddle-FAO formula, which requires temperature, wind speed, mean relative humidity and the sunshine ratio (if these variables are not in the dataset, constant values inserted by the user are adopted) and Penman-Monteith, which uses temperature, solar radiation, wind speed and mean relative humidity of air

Sono disponibili funzioni specifiche per il calcolo di: i) radiazione solare globale giornaliera al suolo con il metodo di Bristow-Campbell (1984) e con quello di Donatelli-Bellocchi (Donatelli e Bellocchi, 2001; Donatelli *et al.*, 2004; Donatelli *et al.*, 2006), ii) *wind chill* temperatura, iii) fotoperiodo, iv) temperatura diurna e notturna,

v) evapotraspirazione di riferimento calcolata, a scelta, con la formula di Blaney-Criddle-FAO o con quella di Penman-Monteith (Allen *et al.*, 1998), vi) codice decadico e vii) fase lunare. Tali variabili possono quindi essere facilmente aggiunte al dataset originale. Altri modelli per la stima dell'evapotraspirazione di riferimento, come quelli di Hargreaves-Samani (1982) e Priestley-Taylor (1972), saranno implementati successivamente. In figura 6 viene riportata la finestra di dialogo per la stima dell'evapo-traspirazione di riferimento. CLIMATICA include il software gratuito Gnuplot, che consente ulteriori elaborazioni grafiche per l'analisi e la presentazione dei dati.

### Ricostruzione dei dati

Dopo aver identificato i valori mancanti e i dati anomali può essere necessario operare una ricostruzione dei dati al fine di permettere l'impiego del dataset in altre procedure di calcolo come, ad esempio, nei modelli di simulazione. In CLIMATICA sono disponibili alcune opzioni per la ricostruzione dei dati mancanti o anomali, come il calcolo della media mobile, la regressione lineare multipla e quella non lineare. Esiste anche la possibilità di addestrare reti neurali personalizzate nonché di impiegare il generatore climatico *Climak* per la generazione di serie di dati climatici.

Come esempio, si illustra l'impiego di due metodi (regressione multipla e rete neurale) per la ricostruzione dei dati mancanti di umidità relativa media in un dataset di dati rilevato a Udine dall'OSMER-ARPA nel periodo 1992-2004.

Nelle figure 7 e 8 sono riportate le relazioni tra i valori di umidità relativa media dell'aria misurati e quelli calcolati, per lo stesso periodo, con le due metodologie: regressione lineare multipla e rete neurale rispettivamente. Il dataset di partenza, utilizzato per la ricostruzione, conteneva alcuni valori mancanti di umidità relativa media (*rhmean*). La presenza di tali valori è stata evidenziata attraverso il controllo **check - dataset**. Con il comando "list if *rhmean*=" si sono individuate le righe con i dati mancanti. Le variabili da impiegare come predittori sono state scelte sulla base della loro correlazione con *rhmean*. È stata quindi calcolata la matrice di correlazione al fine di scegliere le variabili "predittive", utilizzando il comando *correlate*. Queste variabili sono state poi utilizzate sia per stimare il modello di regressione (con il comando *regress*), sia per istruire una rete neurale (dalla finestra di dialogo **neural network**). Le variabili di input utilizzate sono state *doy* (giorno dell'anno), *tmax* (temperatura massima), *tmin* (temperatura minima), *rg* (radiazione solare globale), *winds* (velocità del vento).

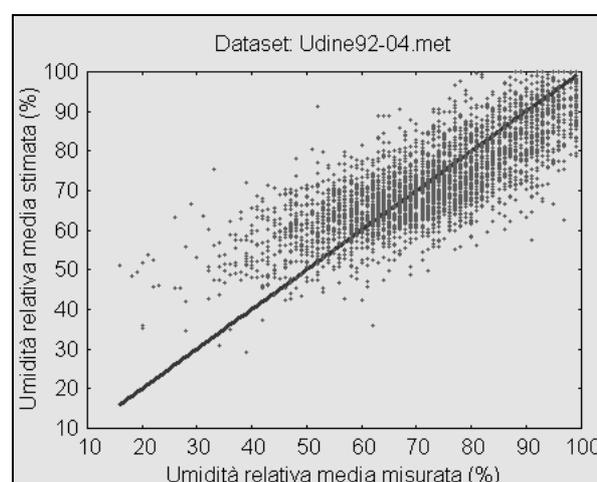
In tabella 4 sono riportate le statistiche per la valutazione del grado di adattamento dei valori stimati ai valori misurati. Si evidenzia chiaramente come la rete neurale abbia mostrato un miglior comportamento, anche considerando la distribuzione dei residui, la pendenza della retta di regressione e il termine noto. Di conseguenza, la ricostruzione dei dati di umidità relativa media è stata effettuata con i valori ottenuti dalla rete neurale con il

```
comando: replace rhmean=rhmeanNet if
rhmean=.
```

**Tab. 4** – Statistiche del confronto tra valori di umidità media giornaliera misurati e calcolati con regressione e con rete neurale. Vengono riportati la pendenza (b) e l'intercetta (a) della retta di regressione lineare.

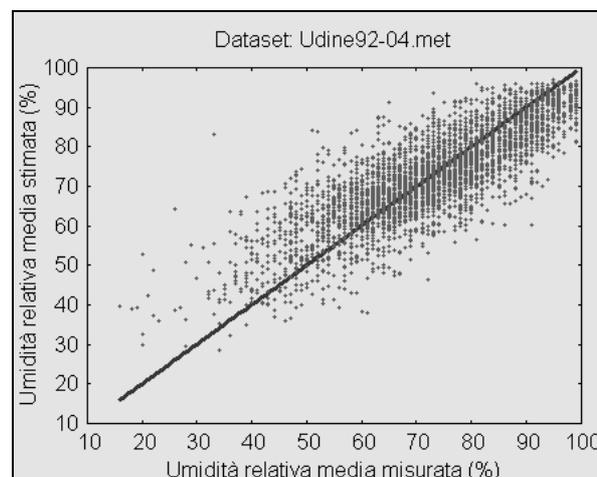
**Tab. 4** – Statistics for comparison of measured and estimated air mean relative humidity, obtained by regression and neural network. Slope (b) and intercept (a) of the linear regression estimated/measured are reported.

	R <sup>2</sup>	RMSE	a	b
<b>Regressione</b>	0.609	9.08	28.2	0.60
<b>Reti neurale</b>	0.700	7.78	23.7	0.68



**Fig. 7** - Relazione tra i valori di umidità relativa media misurati e quelli calcolati con il modello di regressione multipla

**Fig. 7** - Relationship between measured values of mean air relative humidity and those rebuilt by a multiple regression model



**Fig. 8** - Relazione tra i valori di umidità relativa media misurati e quelli calcolati con rete neurale

**Fig. 8** - Relationship between measured values of mean air relative humidity and those rebuilt by a neural network

### Generazione di dati climatici

*CLIMATICA* incorpora il generatore di clima *Climak* (Danuso, 2002), un modello stocastico in grado di generare dati climatici artificiali, sulla base dei parametri climatici stimati, per ciascuna località, partendo da dati storici. L'impiego del generatore climatico consente, tra l'altro, di facilitare la spazializzazione dei dati climatici in zone collinari e montane, lo studio degli effetti dei cambiamenti climatici e la valutazione dei rischi con il metodo Monte Carlo.

Le variabili che *Climak* può attualmente generare sono pioggia, temperatura minima e massima, radiazione solare globale ed evapotraspirazione di riferimento.

I parametri climatici stimati da *Climak* vengono accumulati in un file definito "regional database" (con estensione **red**) in cui, per ciascuna stazione meteorologica, sono riportate le coordinate geografiche, l'altitudine e i 142 parametri che caratterizzano il clima. Sulla base del file di parametri, *Climak* può generare sequenze indefinite di dati climatici giornalieri, sempre diverse, ma che riproducono le proprietà statistiche del clima.

rispetto alla serie completa, allo scopo di evidenziare possibili trend nei parametri climatici. Tale stima è stata eseguita sui dati storici rilevati ad Asiago (VI) dall'Osservatorio Astronomico, per un periodo di 45 anni (1957-2001). Sono state impiegate finestre di calcolo di 6 anni (figura 9a) e di 15 anni (figura 9b). A titolo di esempio, viene riportato uno dei parametri di *Climak*, il parametro A, che costituisce il valore di temperatura media annuale della località, calcolato per minime ( $T_n$ ) e massime ( $T_x$ ), sia per i giorni secchi che piovosi (dry, rainy). I trend di lungo periodo sono meglio evidenziati dalla finestra di maggiore dimensione e si rileva una tendenza costante all'incremento delle temperature medie annue.

### Conclusioni

Il software *CLIMATICA* nonostante sia ancora in una versione preliminare presenta diverse caratteristiche che ne permettono l'impiego nell'ambito degli studi climatici e del trattamento dei dati meteorologici, con particolare riferimento alla scala giornaliera. Altri sviluppi sono attualmente in corso e riguardano, in particolare, procedure per la produzione di statistiche mensili e annuali da dati giornalieri e di dati orari da dati giornalieri e, in *Climak* la generazione del vento (velocità e direzione) e dell'umidità dell'aria. Particolare interesse è rivolto anche alla creazione di moduli per il calcolo dell'intensità di pioggia oraria e al completamento dell'implementazione delle metodologie di valutazione della qualità del dato meteorologico.

Il software è scaricabile gratuitamente al sito [http://www.dpvt.uniud.it/~Danuso/docs/Climatica/Climatica\\_home.html](http://www.dpvt.uniud.it/~Danuso/docs/Climatica/Climatica_home.html).

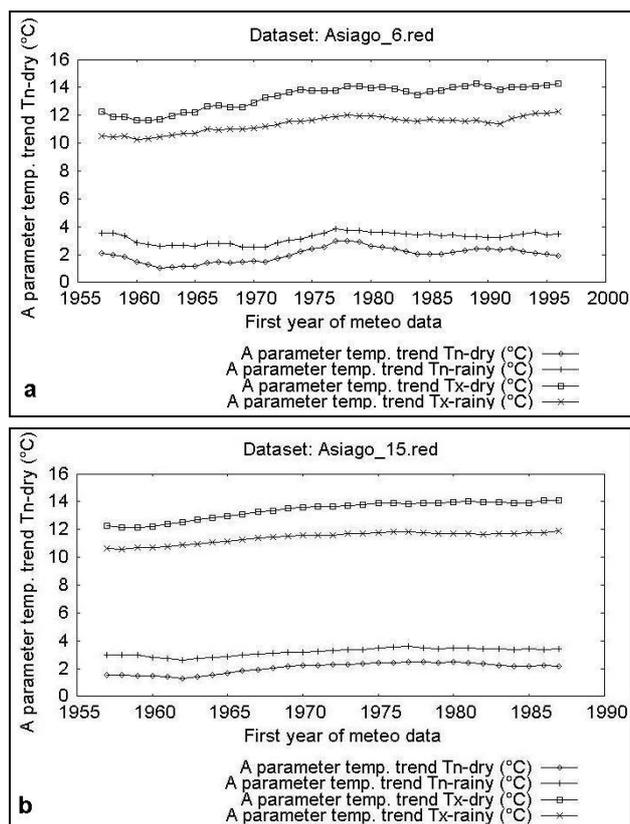
### Ringraziamenti

*CLIMATICA* è stato sviluppato con il contributo finanziario del MIUR, nell'ambito del progetto PRIN 2002 "Sviluppo di metodologie per il trattamento dei dati climatici a scala territoriale ai fini della valutazione dell'impatto dei prodotti fitosanitari sugli ecosistemi non bersaglio" (coordinatore Giuseppe Zanin) e nell'ambito del progetto SIPEAA del Mipaf "Strumenti Informatici per la Pianificazione Eco-compatibile delle Aziende Agrarie", pubblicazione n. 33 (coordinatore Marcello Donatelli).

Si ringraziano inoltre l'ARPAV del Veneto, l'OSMER-ARPA della regione FVG e l'Osservatorio Astronomico di Asiago per avere gentilmente resi disponibili i dati impiegati nel presente lavoro.

### Bibliografia

- Allen R.G., Pereira L.S., Raes D., Smith, M., 1998. *Crop evapotranspiration Guidelines for computing crop water requirements*. FAO Irrig. and Drain. Paper 56. Food and Agriculture Organization of the United Nations, Rome, pp. 300.
- Bristow K., Campbell G.S., 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric. For. Meteorol.*, 31, 159-166.
- Danuso F., 2002. *Climak, a stochastic model for weather data generation*. *Ital.J.Agron.*, 6: 57-71.
- Danuso F., 2003. *SEMoLa: uno strumento per la modellazione degli agroecosistemi*. *Atti XXXV Convegno SIA, Napoli, 16-19/9/ 2003*, 283-284.



**Fig. 9** - Stima dei parametri climatici di *Climak* per la località di Asiago a partire da un file meteo di 45 anni (1957-2001) con finestra di 6 anni (a) e di 15 anni (b)

**Fig 9** - Estimation of *Climak* climatic parameters for Asiago location from a meteo file of 45 years (1957-2001) with a window of 6 years (a) and of 15 years (b)

### Stima mobile dei parametri climatici

Un aspetto particolare di *CLIMATICA* è quello di permettere la stima dei parametri climatici del modello *Climak*, impiegando una "finestra mobile di annate", ridotta

- Donatelli M., Bellocchi G., 2001. Estimate of daily global solar radiation: new developments in the software RadEst3.00. 2nd International Symposium Modelling Cropping Systems, 16-18 July, Florence, Italy, 213-214.
- Donatelli M., Carlini L., Bellocchi G., 2004. GSRad, un componente software per la stima della radiazione solare. *Rivista Italiana di Agrometeorologia*, 1, 24-30.
- Donatelli M., Carlini L., Bellocchi G., 2006. A software component for estimating solar radiation. *Env. Modell. Softw. Vol. 21, 3*, 2006, 411-416.
- Gnuplot è disponibile al sito: <http://www.gnuplot.info>.
- Hargreaves G.H., Samani Z.A., 1982. Reference crop evapotranspiration from temperature. *Appl. Engrg. in Agric.*, 1(2):96-99.
- Lanzante J.R., 1996. Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology*, 16: 1197-1226.
- Meek D.W., Hatfield J.L., 1994. Data quality checking for single station meteorological databases. *Agricultural and Forest Meteorology*, 69: 85-109.
- Osservatorio Astron. Asiago <http://www.pd.astro.it/asiago/>
- Pavan V., Tomozeiu R., Selvini A., Marchesi S., Marsigli C., 2003. Controllo di qualità dei dati giornalieri di temperatura minima e massima di precipitazione. *Quaderno Tecnico ARPA-SIM*, 15/2003.
- Priestley C.H.B., Taylor R.J., 1972. On the assessment of the surface heat flux and evaporation using large scale parameters. *Monthly Weather rev.*, 100, 81-92
- Rana G., Rinaldi M., Introna M., 2004. Metodologie e algoritmi per il controllo di qualità di dati orari e giornalieri acquisiti da una rete agrometeorologica: applicazioni alla rete lucana. *Rivista Italiana di Agrometeorologia* 14-23.
- Reek T., Doty S.R., Owen T.W., 1992. A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *Bulletin American Meteorological Society*, 73: 753-761.
- Song F., Qi H., Weihong Q., 2004. Quality control of daily meteorological data in China, 1951-2000: a new dataset. *International Journal of Climatology* 24: 853-870.
- Štěpánek P., 2003. AnClim – software for time series analysis (for Windows). Department of Geography, Faculty of Natural Sciences, MU, Brno, <http://www.sci.muni.cz/~pest>.
- Tanco R.A., Berri G.J., 2000. CLIMLAB2000, Version 1.1.0. A Statistical Software Package for Climate Applications". IRI-TR-00/1, International Research Institute (IRI) for Climate Prediction, LDEO, Palisades, N.Y., 10964 USA.