9 Concetti di base per lo sviluppo di modelli

Fino a questo punto abbiamo esaminato lo sviluppo di organismi in funzione di temperatura e luce, l'effetto dell'acqua e della radiazione ambientali sugli organismi, la "cattura" da parte delle piante di luce e carbonio e infine la radiazione nelle manti vegetali.
Astraendoci in certa misura dallo specifico processo in esame, in questo capitolo vedremo brevemente come può essere strutturato, dal punto di vista numerico, un modello di simulazione di un processo, riprendendo anche alcuni concetti esposti nel capitolo 3. La discussione che segue riprende la trattazione sui modelli fatta da Acock e Acock (1991).

Le interazioni tra fattori nei modelli

In situazioni in cui sono corrette le maggiori deficienze a carico di tutti i fattori che influenzano un processo, le interazioni tra fattori diventano importanti nel determinare l'effetto sulla variabile di risposta. Supponiamo di conoscere il livello ottimale, per diversi fattori, nel massimizzare una variabile di risposta, che indichiamo con Y. Se tutti i fattori meno uno, che indichiamo con x, sono a livelli ottimali e costanti, la risposta di Y al variare di x puo essere descritta come:

$$Y = Y_{\max} * f(x) \qquad [9.1]$$

dove $f(x)$ è una funzione che varia tra 0 e 1. Se vogliamo determinare l'effetto di diversi fattori, a livelli sub-ottimali, dobbiamo decidere come questi interagiscono nel determinare il valore di Y. Esaminiamo ora modelli additivi, motiplicativi e del fattore minimo.
Se scegliamo un modello additivo, dobbiamo anche determinare quanto ogni fattore contribuisca a determinare Y in rapporto agli altri. Studiamo adesso un ipotetico processo in cui la variabile di risposta, Y, sia funzione di tre variabili indipendenti, x_1, x_2 e x_3, che contribuiscono con un peso di 0.4, 0.4 e 0.2 (somma=1) rispettivamente[2]. L'equazione per il modello additivo in questo caso sarà:

$$Y = Y_{\max} * \{[0.4 * f(x_1)] + [0.4 * f(x_2)] + [0.2 * f(x_3)]\} \qquad [9.2]$$

Adesso supponiamo che il valore delle funzioni sia 0.5 per $f(x_1)$, 0.4 per $f(x_2)$ e infine 0.4 per $f(x_3)$. Secondo il modello additivo avremo:

$$Y = Y_{\max} * \{[0.4 * 0.5] + [0.4 * 0.5] + [0.2 * 0.4]\} = Y_{\max} * 0.48$$

Se invece utilizziamo un modello motiplicativo

$$Y = Y_{\max} * f(x_1) * f(x_2) * f(x_3) \qquad [9.3]$$

con gli stessi valori per le funzioni $f(x_1)$, $f(x_2)$ e $f(x_3)$, otteniamo:

---

[2] I diversi coefficienti possono essere stimati come valori di R^2 parziale, ad ognuno dei quali sono sommate quote proporzionali dell'errore (in modo che la loro somma sia 1), in un modello di regressione multipla, utilizzando dati sperimentali che associano a ogni valore misurato di Y (variabile dipendente) i valori delle funzioni $f(x_1)$, $f(x_2)$ e $f(x_3)$ per i corrispondenti valori dei tre regressori. La stima può però essere attendibile solo se non esiste alcuna correlazione tra x_1, x_2 e x_3; se e tanto più i parametri sono correlati, il modello stimato sarà instabile e quindi il valore del "peso" delle tre funzioni non avrà validità generale.

$$Y = Y_{\max} * 0.5 * 0.5 * 0.4 = Y_{\max} * 0.08$$

Nel caso infine del modello del fattore limitante, il modello è:

$$Y = Y_{\max} * \mathrm{MIN}\left[f(x_1), f(x_2), f(x_3)\right] \qquad [9.4]$$

il risultato, con gli stessi valori per le tre funzioni, sarà:

$$Y = Y_{\max} * \mathrm{MIN}(0.5, 0.5, 0.4) = Y_{\max} * 0.4$$

I tre tipi di modelli esaminati possono fornire risposte molto diverse. In questo esempio il modello additivo e quello del fattore limitante producono risultati simili, ma il problema di fondo del modello additivo, a parte gli aspetti puramente numerici della stima dei coefficienti (cfr. con nota a piè della pagina precedente), risiede nel fatto che è difficile attribuirgli un significato agronomico. Il modello moltiplicativo può sottostimare il valore di *Y*, con errori crescenti all'aumentare del numero dei parametri. Il modello del fattore limitante tende probabilmente a sovrastimare leggermente *Y*, ma concettualmente il modello è facile da capire. Poniamo *Y* sia la sostanza secca di una pianta e che il fattore limitato maggiormente rispetto all'ottimale sia la luce: se la crescita è limitata dalla mancanza di luce la pianta cresce meno, e quindi richiede meno acqua e azoto. Quindi acqua e azoto, che sarebbero limitanti in piena luce, non hanno effetti negativi sull'accrescimento in quanto ne è richiesta una quantità minore. Questo tipo di approccio per la costruzione del modello ci porta a studiare in condizioni controllate, quali quelle di camere termostatate, un fattore alla volta, mantenendo gli altri fattori a livelli ottimali.
Le generalità esposte sui tipi di modelli non escludono che, a seconda dei casi, sia più opportuno usare un tipo di modello piuttosto che un'altro. Nei due esempi che seguono vedremo come i modelli più adeguati risultino in un caso il modello del fattore limitante, nell'altro quello moltiplicativo.

<u>esempio: tasso di decomposizione di residui colturali</u>
da prove di laboratorio evidenziamo in fig. 9.1 il tipo di risposta al tempo della percentuale di decomposizione dei residui:

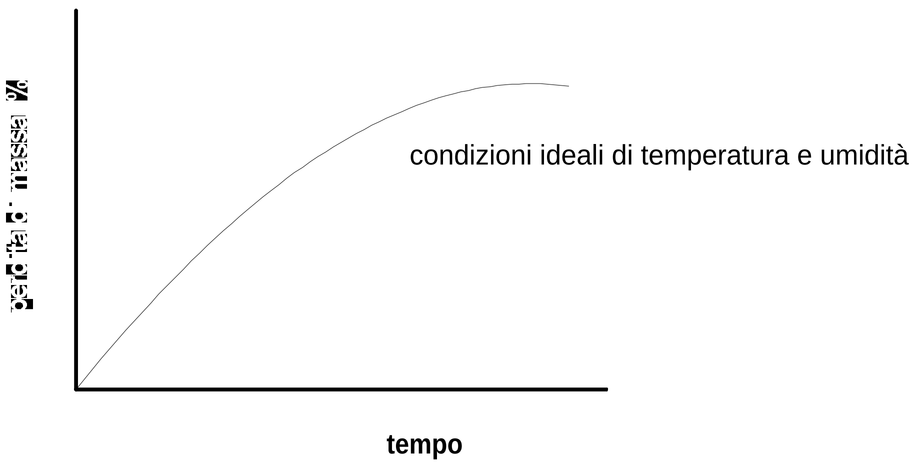

Figura **9.1** Decomposizione di residui vegetali come funzione del tempo in condizioni ideali di temperatura e umidità

Questo tipo di risposta si riferisce però a condizioni ottimali di temperatura e umidità nel suolo. Sia la temperatura che il contenuto idrico del suolo sono però variabili; dobbiamo quindi modellare la risposta ai due fattori inglobando i risultati in un unico modello, che permetta di stimare il tasso di decomposizione dei residui per ogni passo d'integrazione.
Sempre in condizioni di laboratorio, studiamo la risposta della decomposizione dei residui a temperatura e acqua separatamente (rispettivamente, grafici 9.2A e 9.2B):

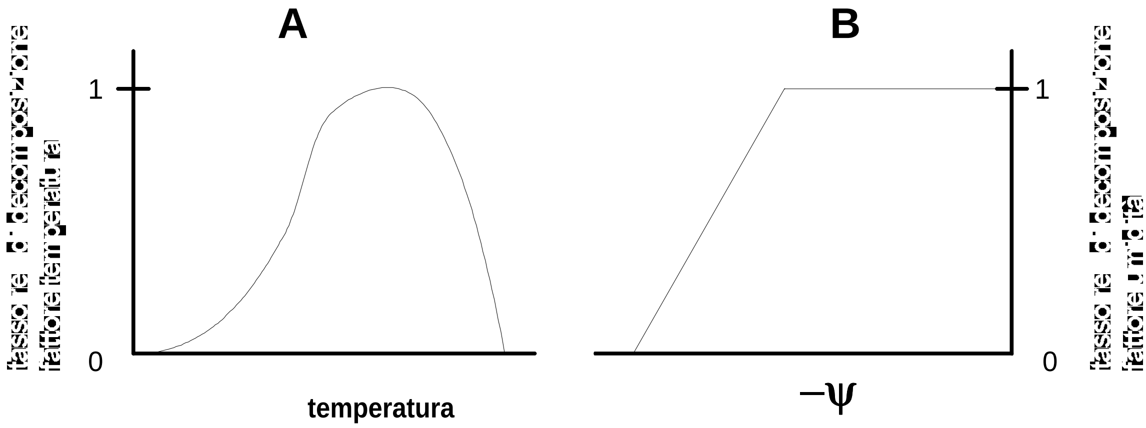

Figura **9.2** Tasso relativo di decomposizione di residui vegetali in funzione di temperatura (A) e dell'umidità (B)

In questo caso il modello appropriato è quello del fattore limitante; ad ogni passo d'integrazione il tasso massimo di decomposizione viene moltiplicato per il tasso relativo minimo tra quelli corrispondenti ai fattori temperatura e umidità.

<u>esempio: infestazione fungina su foglie di patata</u>
da prove di laboratorio evidenziamo in fig. 9.3 il tipo di risposta al tempo della percentuale di infestazione di foglie di patata:

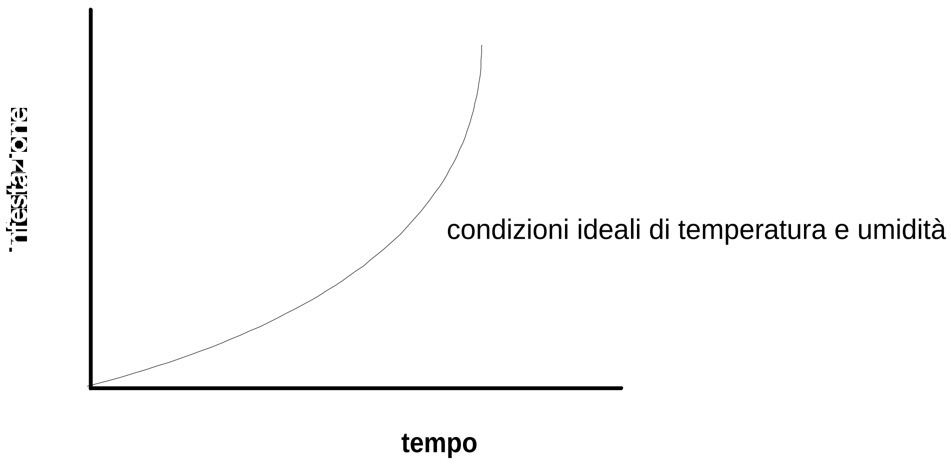

Figura **9.3** Infestazione fogliare per patata come funzione del tempo in condizioni ottimali di temperatura e di bagnatura fogliare.

Anche in questo caso ci interessa modellare la risposta a temperatura e bagnatura fogliare per poter stimare l'infestazione in condizioni subottimali per il patogeno. Studiamo in laboratorio separatamente la risposta ai due fattori; i risultati sono riportati nella figura 9.4

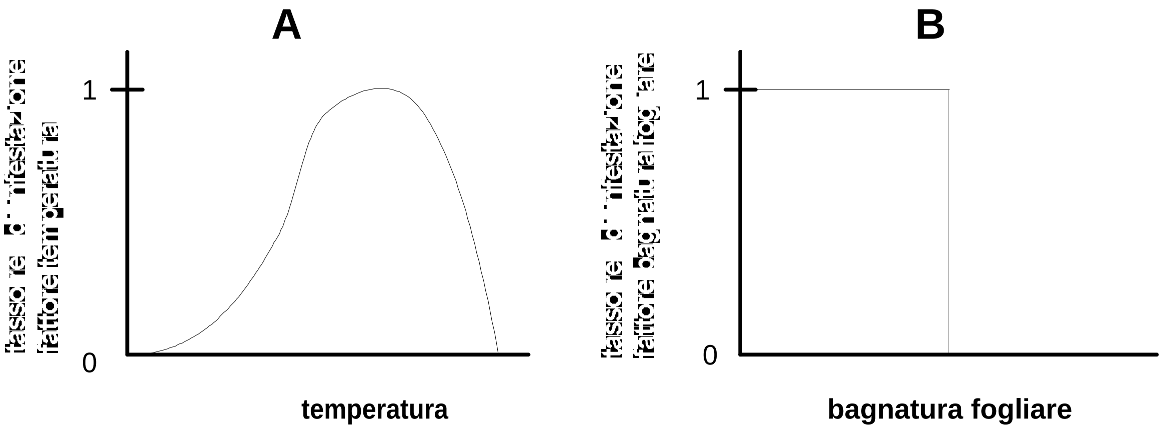

Figura **9.4** Tasso relativo di infestazione di foglie di patata in funzione di temperatura (A) e della bagnatura fogliare (B).

In questo caso il modello appropriato è quello moltiplicativo; ad ogni passo d'integrazione il tasso relativo di infestazione viene calcolato per la temperatura e per la bagnatura fogliare. Per bagnature fogliari al di sotto di un valore di soglia il valore risultante sarà positivo e pari a quello calcolato per la temperatura, al di sopra sarà uguale a zero.

Tipo di dati utilizzabili per sviluppare un modello

Essenzialmente, i dati che si rendono disponibili per sviluppare un modello provengono da esperimenti condotti o in campo oppure in condizioni controllate, come in camera di crescita, serre, o più in generale in laboratorio. Nel secondo caso si tenta di costruire un modello fisico semplificato del sistema in studio, e su di esso si apportano perturbazioni che consistono nel fare variare un fattore per volta, valutando il comportamento di una o più variabili di risposta al fattore in esame. Se si studia la risposta di una pianta in condizioni controllate, un problema minore è dato dal fatto che la pianta spesso cresce in modo differente rispetto al pieno campo. La differenza può essere chiaramente visibile o può risultare dalla misurazione di variabili fisiologiche (Hurd, 1969; Thorne, 1970; Rajan *et al.*; 1971; Bunce, 1977; tutti citati da Acock e Acock, 1991). I dati provenienti da condizioni controllate dovrebbero quindi essere utilizzati per stabilire la forma generale della risposta ai diversi fattori, mentre i dati provenienti dagli esperimenti in campo dovrebbero essere utilizzati per calibrare i parametri del modello, o per evidenziare che la semplificazione del sistema fatta per il suo studio in laboratorio è stata inadeguata. La calibrazione è però un processo estremamente pericoloso quando finisce per l'essere un "data fitting" (Penning de Vries, 1982), in quanto, se i parametri vengono fatti variare al di fuori dell'intervallo di variazione noto, il modello meccanicistico viene ad essere snaturato fino a diventare qualcosa di comparabile alle regressioni multiple. Se è richiesto un processo di calibrazione di questo tipo, ciò può voler dire che il modello in esame è inadeguato a simulare le

condizioni in studio, mentre può risultare perfettamente adeguato per altre; sul significato delle discordanze tra valori simulati e valori misurati torneremo con maggior dettaglio nel capitolo 13.
I dati di campo da soli non permettono invece, se non con rarissime eccezioni, di costruire modelli che abbiano una base meccanicistica.