

Breve riassunto

Dati entomologici di campo presentano comunemente eccessi di zero ed over-dispersione. Dall'utilizzo di diversi modelli statistici e sull'esperienza di due prove di campo in condizioni molto diverse, si è potuto verificare che per stabilire i possibili effetti ambientali come differenza fra due trattamenti sull'abbondanza di specie di artropodi, deve essere impostato considerando specificamente le caratteristiche della dispersione dei dati ottenuti. Una analisi preventiva del potere discriminatorio è sempre necessaria.

La natura dei dati entomologici di campo e implicazioni per le analisi statistiche

Salvatore Arpaia⁽¹⁾, Ferdinando Baldacchino⁽¹⁾, Paul W. Goedhart⁽²⁾, Martine Kos⁽³⁾, Joop J.A. van Loon⁽³⁾, Hilko van der Voet⁽²⁾

⁽¹⁾ENEA, Agenzia Nazionale per le Nuove Tecnologie, l'Energia e lo Sviluppo Economico Sostenibile - Laboratorio di Sviluppo Sostenibile della Produzione Primaria (UTTRI-SSPP), Centro Ricerche Trisaia, Rotondella (MT); ⁽²⁾DLO, Wageningen University and Research Centre, Plant Research International, Biometris, Wageningen, Netherlands; ⁽³⁾Laboratory of Entomology, Wageningen University, Netherlands

Introduzione

La raccolta di dati ecologici di campo rappresenta la base su cui vengono condotte analisi entomologiche per vari scopi (e.g. valutazione di impatti ambientali dovuti alle attività agricole, valutazione della resistenza varietale agli artropodi, ecc.). I metodi di analisi statistiche utilizzati andrebbero scelti in base sia alla natura dei dati stessi che alla specifica risposta attesa. Quando l'obiettivo dello studio intende valutare gli impatti ambientali di una pratica agricola sull'artropodofauna, una particolare attenzione va indirizzata al calcolo del potere discriminatorio del modello statistico (power analysis). Con l'obiettivo di ottenere un modello efficiente per la valutazione di impatti ambientali, in questo lavoro sono stati analizzati dati provenienti da campi di confronto varietale, per verificare il modello statistico più adatto a rappresentare la distribuzione di artropodi su colture erbacee e per stimare il numero di repliche necessarie per un esperimento con un sufficiente potere discriminatorio.

Materiali e metodi

Dataset n.1. Campo sperimentale (Metaponto (MT), IT) con 10 varietà di fragola, anni 2004 e 2005. Disegno sperimentale completamente randomizzato, con 6 repliche nel 2004 e 5 nel 2005. Ogni plot costituito da 10 piante, tutte campionate per monitorare la presenza di fitofagi tramite campionamento visivo. *Aphis gossypii* Glover è stato monitorato contando il numero di afidi da una foglia scelta a caso per ogni pianta; le osservazioni sono state eseguite settimanalmente dal 15 marzo 2004 al 31 maggio 2004 e dal 3 maggio 2005 al 20 giugno 2005. *Tetranychus urticae* Koch monitorato nel 2004 con la stessa metodologia e frequenza applicata per la popolazione afidica, mentre nel 2005 sono stati eseguiti solo due rilievi (16 e 29 maggio).

Dataset n.2. Campo sperimentale (Wageningen, NL) con 4 cultivar di cavolo bianco anni 2008 e 2009. in Olanda) Disegno sperimentale a blocchi randomizzati, con 8 repliche e campionato settimanalmente da inizio giugno a fine settembre, scegliendo casualmente 9 piante per plot. Il dataset, contenente dati relativi alla presenza di fitofagi, parassitoidi e predatori, è costituito dal numero di afidi (*Brevicoryne brassicae* L.), larve e pupe del lepidottero *Plutella xylostella* L., mummie e afidi parassitizzati dall'imenottero *Diaeretiella rapae* McIntosh, pupe di *P. xylostella* parassitizzata da *Diadegma semiclausum* Hellén, larve e pupe del sirfide *Episyrphus balteatus* de Geer, uova e pupe del neurottero *Chrysoperla carnea* Stephens.

Su tali dataset sono stati calcolati i principali parametri di dispersione delle variabili considerate e applicati alcuni modelli di stima dei parametri di dispersione.

Risultati e discussione

I conteggi relativi ad afidi ed acari su fragola, hanno rivelato che le caratteristiche principali di queste distribuzioni sono: una notevole “over-dispersione” ed un notevole numero di zeri. La distribuzione di Poisson non riesce a modellizzare efficacemente il fenomeno di overdispersione né la mancata proporzionalità fra la varianza e la media (Goedhart et al., 2014). Un’analisi basata su una distribuzione lognormale ha il limite di un cattivo fit degli zeri. La distribuzione binomiale negativa offre caratteristiche intermedie fra una distribuzione di Poisson per medie basse e una distribuzione lognormale per medie alte. Anche l’analisi di dati provenienti dallo studio sul cavolo relativamente a diversi fitofagi e loro nemici naturali, ha mostrato che soltanto in 18 delle 280 date di campionamento (6%) in tutte le 32 parcelle sono stati registrati valori diversi da zero. Al contrario, in 87 casi (31%) tutte le 32 parcelle avevano conteggi a valore zero. In base alle caratteristiche dei fenomeni di dispersione registrati, si può evidenziare come il potere discriminatorio del modello statistico sia sensibile al valore della variabili rilevato, alla intensità del fenomeno di overdispersione e alla eventuale differenza tra le medie dei trattamenti da testare. Da un esempio teorico in cui si intende valutare un trattamento che fa registrare dei valori medi di una variabile doppi rispetto a quelli del testimone, si ricavano le stime del numero di repliche necessario, in base a 3 possibili modelli descrittivi, per raggiungere un potere discriminatorio dell’80% (Tab. 1). Una analisi preventiva del potere discriminatorio è necessaria. Per distribuzioni non-normali si può procedere per simulazioni, o ricorrere a un metodo pratico e più facile (Lyles et al., 2007).

Tabella 1. Numero di repliche necessarie per ottenere un risultato significativo ($\alpha=0,05$) con probabilità dell’80%, quando la tesi trattata ha valore medio doppio del controllo. μ = media, ω = fattore di overdispersione (approx. Varianza/media)

NegBinomiale	$\mu= 1$	$\mu= 2$	$\mu= 5$	$\mu= 10$	$\mu= 20$	$\mu= 40$
$\omega= 0.25$	29	21	13	10	9	9
$\omega= 0.50$	>40	27	21	19	17	16
$\omega= 1.00$	>40	>40	37	35	33	32
Over-Poisson	$\mu= 1$	$\mu= 2$	$\mu= 5$	$\mu= 10$	$\mu= 20$	$\mu= 40$
$\omega= 0.25$	32	22	13	10	10	9
$\omega= 0.50$	>40	27	21	18	16	16
$\omega= 1.00$	>40	39	32	32	28	27
Log-Normale	$\mu= 1$	$\mu= 2$	$\mu= 5$	$\mu= 10$	$\mu= 20$	$\mu= 40$
$\omega= 0.25$	36	26	17	13	12	11
$\omega= 0.50$	>40	37	30	27	24	23
$\omega= 1.00$	>40	>40	>40	>40	>40	>40

Conclusioni

L’applicazione di un modello statistico per dati entomologici in campo deve essere basata su un metodo che possa gestire sia gli eccessi di zero che l’over-dispersione. La caratteristica over-dispersione porta ad una necessità di aumentare in maniera notevole le repliche in caso di medie dei conteggi inferiori a 5-10. I dati da noi utilizzati indicano che un esperimento di campo per valutare l’abbondanza di specie non-bersaglio (es. nemici naturali) come differenza fra due trattamenti deve essere impostato considerando le caratteristiche di dispersione dei dati. Una analisi preventiva del potere discriminatorio è sempre necessaria.

Bibliografia

Goedhart PW, van der Voet H, Baldacchino F, Arpaia S (2014). *Ecology and Evolution*, 4: 1267–1283. <http://dx.doi.org/10.1002/ece3.1019>.

Lyles RH, Lin H-M & Williamson JM (2007). *Statistics In Medicine*: 6(7): 1632-1648.