

Analysis of food web

Alberto Biliotti - Jacopo Manetti

May 2024

1 Introduzione

Le reti alimentari sono essenziali per comprendere le complesse interazioni tra gli organismi in un ecosistema e il flusso di materia ed energia attraverso di essi. Nel contesto degli ecosistemi del Sud della Florida, le reti alimentari rivestono un'importanza particolare, riflettendo le dinamiche ecologiche uniche di questa regione ricca di biodiversità. Questo progetto si propone di analizzare le reti alimentari di Florida Bay, concentrandosi sulle stagioni umide e secche, per cogliere la complessità dei processi ecologici in atto in uno degli habitat più significativi della regione.

2 Descrizione della rete e preprocessing

Il dataset utilizzato per questo studio comprende reti alimentari relative agli ecosistemi del Sud della Florida, focalizzandosi specificamente sul Florida Bay. Inizialmente, il dataset conteneva otto reti alimentari separate, rappresentative di quattro biomi durante entrambe le stagioni umide e secche. Tuttavia, per questo progetto ci siamo concentrati esclusivamente sul Florida Bay, in quanto questa rete presenta il maggior numero di nodi rispetto alle altre.

Entrambe le reti relative al Florida Bay, sia per la stagione umida che per quella secca, sono composte dallo stesso numero di nodi, pari a 128.

Ogni nodo all'interno delle reti alimentari del Florida Bay è caratterizzato da tre attributi distinti:

- Nome del componente: Questo attributo identifica il componente specifico dell'ecosistema rappresentato dal nodo.
- Bio-masse: Indica la quantità di materia organica presente in ciascun componente dell'ecosistema. Questo attributo fornisce informazioni cruciali sulla quantità di risorse disponibili e sulle dinamiche energetiche all'interno della rete.
- Tipo di classificazione: Questo attributo classifica il nodo in una delle seguenti categorie:
 1. Compartimento vivente/produttivo
 2. Altro compartimento
 3. Input
 4. Output
 5. Respirazione

Per garantire un'analisi centrata sulle interazioni tra gli organismi e non sulla loro contribuzione totale all'input/output di carbonio nell'ecosistema, abbiamo rimosso dalla rete tutti i nodi speciali che si riferiscono a compartimenti non organismi. Questo ci consente di focalizzarci sulle catene alimentari e sulle relazioni tra gli organismi all'interno dell'ecosistema.

Dopo questa selezione, sono rimasti 121 nodi da analizzare. La rete relativa alla stagione secca presenta 1799 archi, mentre quella relativa alla stagione umida ne presenta 1767. Questi dati forniscono una base solida per l'analisi delle reti alimentari del Florida Bay e per comprendere le dinamiche ecologiche in questo importante ecosistema del Sud della Florida.

3 Statistiche di rete

Per riepilogare le funzionalità delle reti alimentari del Florida Bay, sono state calcolate le seguenti statistiche di rete per entrambe le stagioni:

Statistiche di Rete	Stagione Secca	Stagione Umida
Densità	0.12	0.12
Reciprocità	0.004	0.004
Transitività	0.27	0.27
Odds Ratio Transitività	2.63	2.62
Assortatività della Biomassa	0.05	0.06

Table 1: Statistiche di rete per le stagioni secca e umida del Florida Bay

3.1 Densità

La densità della rete, pari a 0.12 per entrambe le stagioni, indica che solo il 12% delle possibili connessioni tra nodi sono effettivamente presenti. Questo suggerisce una rete relativamente sparsa, il che è comune nelle reti ecologiche, dove non tutti i componenti sono direttamente connessi.

3.2 Reciprocità

Il valore di reciprocità di 0.004 mostra che c'è una bassa probabilità che le connessioni tra i nodi siano reciproche. Questo è tipico delle reti alimentari, dove le relazioni tra preda e predatore non sono simmetriche.

3.3 Transitività

La transitività, pari a 0.27, indica una presenza moderata di triadi chiuse nella rete. Questo valore suggerisce che se un nodo A è connesso a un nodo B, e B è connesso a C, c'è una discreta probabilità che A sia anche connesso a C, riflettendo sottosistemi ecologici connessi tra loro in maniera più stretta.

3.4 Odds Ratio Transitività

L'odds ratio della transitività, circa 2.63 per entrambe le stagioni, conferma che le triadi chiuse sono più frequenti rispetto a quelle che ci si aspetterebbe in una rete randomizzata. Questo suggerisce che le reti alimentari del Florida Bay sono strutturate in modo che certi gruppi di nodi tendano a

formare cluster. La presenza di cluster trofici può indicare la presenza di micro-habitat o nicchie ecologiche ben definite all'interno dell'ecosistema, dove specie connesse condividono relazioni trofiche più intense.

3.5 Assortatività della Biomassa

L'assortatività della biomassa, pari a 0.05 nella stagione secca e 0.06 nella stagione umida, indica una debole tendenza dei nodi con biomassa simile ad essere collegati tra loro. Questa bassa assortatività è tipica delle reti ecologiche, dove le connessioni tra nodi sono spesso determinate da fattori trofici piuttosto che da somiglianze di biomassa.

3.6 Conclusioni

L'analisi delle reti alimentari del Florida Bay rivela che queste sono relativamente sparse, con una bassa reciprocità ma una moderata transitività. La bassa assortatività della biomassa suggerisce che le connessioni tra nodi sono più influenzate dalle relazioni trofiche piuttosto che da somiglianze nelle dimensioni dei nodi. Questi risultati forniscono una base solida per ulteriori studi sulle dinamiche ecologiche del Florida Bay, contribuendo a una migliore comprensione delle interazioni complesse all'interno di questo ecosistema.

4 Misure di Centralità

Le misure di centralità sono cruciali per comprendere l'importanza relativa dei nodi all'interno delle reti alimentari del Florida Bay. In particolare, abbiamo esaminato cinque misure di centralità: in-degree, out-degree, closeness, betweenness ed eigenvector. Queste misure offrono diverse prospettive sulla posizione e sul ruolo dei nodi nella rete.

4.1 In-Degree Centrality

L'in-degree centrality misura il numero di connessioni dirette che arrivano a un nodo. Nella rete alimentare, questo rappresenta il numero di predatori che consumano una particolare specie.

- **Stagione Umida:** Crocodiles, Predatory Ducks, Raptors, Greeb, Scianids, Dolphin, Pompano, Loon, Big Herons & Egrets.
- **Stagione Secca:** Big Herons & Egrets, Raptors, Crocodiles, Dolphin, Predatory Ducks, Greeb, Scianids, Loon, Pompano.

Questi nodi sono generalmente grandi predatori o specie carismatiche con un impatto significativo sulla struttura della rete alimentare.

4.2 Out-Degree Centrality

L'out-degree centrality misura il numero di connessioni che partono da un nodo. Questo rappresenta il numero di specie che una particolare specie predatrice consuma.

- **Stagione Umida:** Herbivorous Shrimp, Predatory Shrimp, Pink Shrimp, Omnivorous Crabs, Bivalves, Detritivorous Gastropods, Goldspotted Killifish, Detritivorous Amphipods, Suspension Feeding Polychaetes.

- **Stagione Secca:** Predatory Shrimp, Herbivorous Shrimp, Pink Shrimp, Omnivorous Crabs, Bivalves, Detritivorous Gastropods, Goldspotted Killifish, Detritivorous Amphipods, Suspension Feeding Polychaetes.

Questi nodi sono spesso piccoli invertebrati che sono adatti ad essere consumati da molte altre specie, riflettendo la loro importanza come prede cruciali nella rete alimentare.

4.3 Closeness Centrality

La closeness centrality misura quanto un nodo è vicino a tutti gli altri nodi nella rete, indicando l'efficienza con cui una specie può influenzare l'intera rete alimentare.

- **Stagione Umida:** Rays, Snook, Jacks, Pompano, Other Snapper, Gray Snapper, Scianids, Spotted Seatrout, Red Drum.
- **Stagione Secca:** Rays, Snook, Jacks, Pompano, Other Snapper, Gray Snapper, Scianids, Spotted Seatrout, Red Drum.

Questi nodi sono spesso pesci predatori di media taglia che svolgono ruoli cruciali nelle dinamiche trofiche, con una rapida influenza su molte altre specie.

4.4 Betweenness Centrality

La betweenness centrality misura la frequenza con cui un nodo appare nei percorsi più brevi tra altri due nodi, indicando il ruolo di una specie come intermediaria o ponte nella rete alimentare.

- **Stagione Umida:** Predatory Gastropods, Echinoderma, Predatory Polychaetes, Other Cnidariidae, Other Demersal Fishes, Predatory Shrimp, Eels, Flatfish, Catfish.
- **Stagione Secca:** Predatory Gastropods, Echinoderma, Other Cnidariidae, Predatory Polychaetes, Other Demersal Fishes, Predatory Shrimp, Eels, Flatfish, Catfish.

Questi nodi rappresentano specie che connettono diversi compartimenti trofici, svolgendo un ruolo chiave nella stabilità e coesione della rete.

4.5 Eigenvector Centrality

L'eigenvector centrality misura l'influenza di un nodo non solo in base al numero di connessioni dirette, ma anche alla qualità delle connessioni, tenendo conto dell'importanza dei nodi a cui è connesso.

- **Stagione Umida:** Raptors, Crocodiles, Dolphin, Barracuda, Predatory Ducks, Loon, Pelican, Comorant, Big Herons & Egrets.
- **Stagione Secca:** Raptors, Crocodiles, Dolphin, Big Herons & Egrets, Barracuda, Predatory Ducks, Loon, Comorant, Pelican.

Questi nodi rappresentano grandi predatori con ampie reti di influenza, spesso al vertice delle catene alimentari.

4.6 Considerazioni sulle Differenze tra Stagione Umida e Secca

Analizzando le reti alimentari del Florida Bay nelle stagioni umida e secca, emergono alcune differenze interessanti:

- **In-Degree Centrality:** Le specie più centrali per in-degree rimangono simili tra le stagioni, ma l'ordine cambia leggermente. Ad esempio, i "Big Herons & Egrets" passano dalla prima posizione nella stagione secca alla nona nella stagione umida, suggerendo variazioni stagionali nella predazione o nella disponibilità di prede.
- **Out-Degree Centrality:** Le specie con il più alto out-degree rimangono le stesse in entrambe le stagioni, indicando che i piccoli invertebrati come i gamberi e i granchi mantengono un'alta vulnerabilità alla predazione tutto l'anno.
- **Closeness Centrality:** Le specie con alta closeness centrality non variano tra le stagioni, riflettendo la loro posizione stabile e centrale nella rete alimentare.
- **Betweenness Centrality:** Anche per la betweenness centrality, le specie chiave rimangono costanti tra le stagioni, suggerendo che il loro ruolo di connettori all'interno della rete alimentare è critico in ogni periodo.
- **Eigenvector Centrality:** Le specie con alta eigenvector centrality sono molto simili tra le stagioni, con leggere variazioni nell'ordine. Questo riflette che le specie più influenti tendono a mantenere il loro status indipendentemente dalla stagione.

Queste osservazioni indicano una relativa stabilità nella struttura delle reti alimentari del Florida Bay tra le stagioni umida e secca, con alcune variazioni nelle posizioni centrali delle specie. Questa stabilità potrebbe essere dovuta all'adattamento delle specie alle variazioni stagionali nell'ambiente e nella disponibilità di risorse.

4.7 Discussione e Conclusioni

Le misure di centralità forniscono informazioni complementari sulla struttura e la funzione della rete alimentare del Florida Bay. Ad esempio, specie con alto in-degree centrality, come i coccodrilli e gli aironi, sono cruciali per la rete poiché molti altri nodi dipendono da loro come predatori. Specie con alto out-degree centrality, come i gamberi e i granchi, sono importanti prede che influenzano molte altre specie.

Le specie con alta closeness centrality, come le razze e i robali (snook), possono diffondere rapidamente effetti attraverso la rete, mentre quelle con alta betweenness centrality, come i gasteropodi predatori e gli echinodermi, agiscono come punti di snodo critici. Infine, specie con alta eigenvector centrality, come i rapaci e i delfini, sono altamente influenti e connesse a molte altre specie importanti.

In generale, i risultati ottenuti sono in linea con le aspettative ecologiche, dove grandi predatori e prede generaliste giocano ruoli chiave nella stabilità e funzionamento della rete alimentare. Questi risultati evidenziano l'importanza di preservare queste specie per mantenere l'integrità e la resilienza dell'ecosistema del Florida Bay.

5 Modelli random per grafi

In quest'ultima parte cerchiamo di sfruttare i modelli a nostra disposizione per cercare di descrivere per quanto possibile il meccanismo di generazione dei dati partendo da modelli più semplici con assunzioni più stringenti e andando verso modelli con assunzioni meno stringenti.

Si avvisa il lettore che tutte le immagini a coppie mostrate di seguito sono disposte in modo tale da rappresentare, a sinistra, la rete per la stagione secca e, a destra, la rete per la stagione umida.

5.1 Simple Random Graph Model

Il primo modello che utilizzeremo è il Simple Random Graph Model il quale come intuibile dal nome è il modello più semplice ma anche quello che presenta le assunzioni più stringenti. Infatti esso assume che ogni variabile dicotomica Y_{ij} relativa all'arco tra il nodo i ed il nodo j sia indipendente e identicamente distribuita rispetto ad ogni altra variabile relativa a un arco come una distribuzione binomiale con parametro per la probabilità di successo p comune a tutti gli archi della rete.

$$Y_{ij} \sim \text{Binomiale}(p) \quad \forall ij \in V$$

Utilizzando la formulazione del “Exponential Random Graph Model” si ha che la funzione di verosimiglianza ossia la probabilità di osservare la matrice di adiacenza y dato il modello è uguale a:

$$P(Y = y) = k(\theta)^{-1} \exp(g(y)' \theta)$$

Dove $g(y)$ ossia il vettore delle statistiche sufficienti contiene solo la somma di tutti i nodi della rete e il vettore di parametri θ contiene il parametro associato a tale statistica mentre $k(\theta)$ è una costante di normalizzazione.

Per adattare tali modelli alle nostre due reti abbiamo utilizzato su R il comando “ergm” dell'omonima libreria passando in ingresso la formula contenente la rete e la keyword “edge” per indicare che l'unica statistica di nostro interesse era il numero di archi della rete. I risultati dell'esecuzione per entrambe le reti sono riportati di seguito:

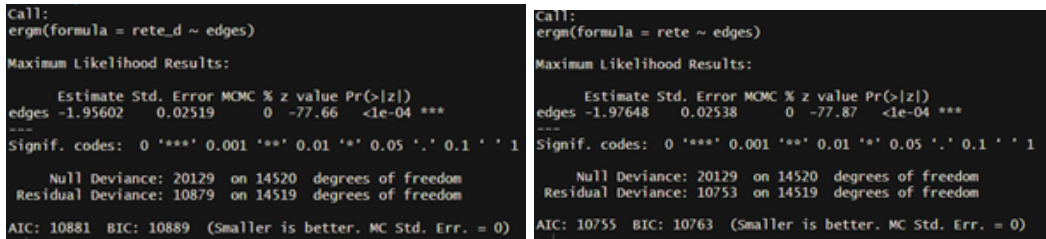


Figure 1: Risultati dell'esecuzione del Simple Random Graph Model

Notiamo come in entrambi i casi il parametro risulti essere estremamente significativo e molto simile per le due reti. Il segno negativo indica che la densità degli archi nella rete è minore di 0.5; se fosse pari a 0.5 tale parametro sarebbe infatti pari a 0. Possiamo riottenere il valore esatto della densità degli archi nelle reti che è pari al valore di p introdotto prima utilizzando la formula:

$$p = \frac{e^{\mu}}{1 + e^{\mu}}$$

Dove μ rappresenta il parametro associato al numero di archi. Possiamo infatti osservare tale risultato nella seguente immagine:

```
> exp(coef(mod0_w))/(1+exp(coef(mod0_w)))
edges
0.1216942
> exp(coef(mod0_d))/(1+exp(coef(mod0_d)))
edges
0.1238981
```

Figure 2: Densità degli archi nelle reti

Tale modello per quanto semplice non è appropriato per descrivere le nostre reti come possiamo osservare dal confronto tra le statistiche delle nostre reti in particolare la transittività e la deviazione standard del “in-degree” e del “out-degree” con quelli delle mille reti che abbiamo simulato da tali modelli:

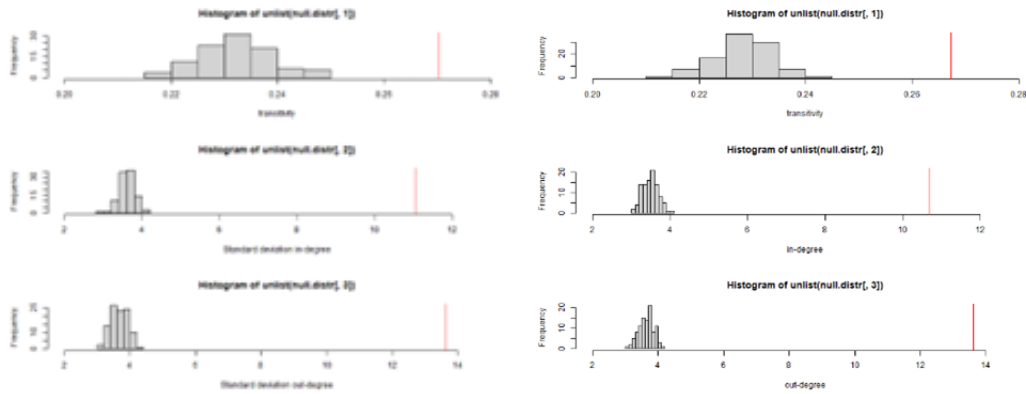


Figure 3: Confronto tra le statistiche delle nostre reti e quelle simulate

Notiamo come i valori delle statistiche delle nostre reti risultino essere parecchio distanti da quelli delle reti simulate dimostrando così l’inappropriatezza del “Simple Random Graph Model”.

5.2 Simple Random Graph con attributo relativo alla biomassa della specie

Dopo essere partiti controllando il modello più semplice lo complichiamo aggiungendoci le statistiche relative agli effetti principali (main effect) e all’omofilia dell’attributo nodale relativo alla biomassa della specie per osservare se tale attributo ha un qualche tipo di influenza sugli archi delle nostre reti. Per fare ciò andremo quindi a definire un modello con tre statistiche ossia il numero di nodi, gli effetti principali e l’omofilia.

Un valore significativo del parametro relativo agli effetti principali indicherebbe che valori elevati della biomassa di una specie abbiano influenza positiva o negativa in base al segno sugli archi

della rete e quindi sulla capacità di una specie di predare o essere predata da altre specie. Un valore significativo del parametro relativo all'omofilia indicherebbe invece che nodi con valore della biomassa simile tendano a predarsi tra loro.

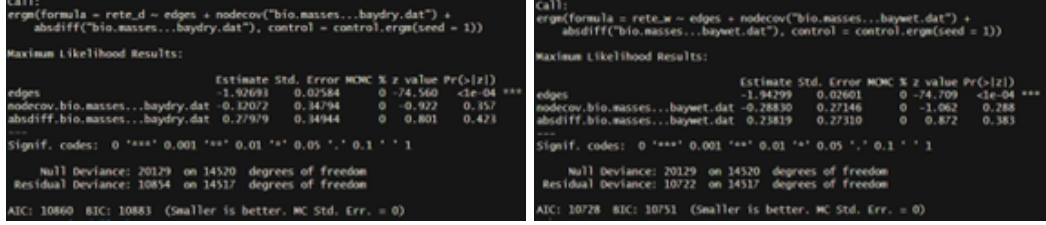


Figure 4: Risultati dell'applicazione del modello con attributo relativo alla biomassa della specie

Come possiamo notare dalle immagini sopra riportate che mostrano i risultati dell'applicazione di tale modello alle nostre reti, entrambi i parametri relativi alle statistiche della biomassa risultano essere non significativamente diversi da zero. Possiamo quindi affermare che in questo caso la biomassa di una specie non abbia alcuna influenza sulle relazioni della nostra rete ossia sul fatto che un organismo possa essere predato o predare un'altra specie. Non prenderemo più in considerazione quindi tale attributo nei modelli successivi.

5.3 Non-Homogeneous Simple Random Graph Model

Il prossimo passo consiste nel permettere a ciascuna variabile Y_{ij} relativa ad un arco della nostra rete di avere una probabilità di successo p distinta in base alle caratteristiche del mittente e del destinatario. In pratica associamo a ciascuna variabile dicotomica Y_{ij} una probabilità di successo p_{ij} distinta che dipende dalla propensione dei nodi mittente e destinatario di inviare o di ricevere relazioni. Nel nostro caso utilizzando questo modello la probabilità di un organismo di nutrirsi di un altro dipenderà dalla sua propensione a predare un qualsiasi altro organismo e dalla propensione della potenziale preda ad essere predata. Avremo quindi che:

$$Y_{ij} \sim \text{Bernoulli}(p_{ij}) \quad \forall ij \in V$$

Dove p_{ij} è modellata come:

$$p_{ij} = \frac{e^{\mu+a_i+b_j}}{1 + e^{\mu+a_i+b_j}}$$

In cui a_i è il parametro che modella l'effetto del nodo i come mittente mentre b_j è il parametro relativo all'effetto di j come destinatario. Come prima μ è il parametro che indica la tendenza relativa della rete a presentare un arco tra due nodi. Avremo quindi un parametro relativo alla propensione dell'organismo a predare una specie (effetto destinatario nel nostro caso) ed uno relativo alla propensione di un organismo ad essere predato (effetto mittente nel nostro caso) per ciascun organismo presente nella rete tranne per il primo organismo che viene preso come riferimento. Avremmo quindi in totale 241 parametri nel modello considerando anche il parametro μ .

Il problema è che in entrambe le reti alcuni organismi non si nutrono di nessun altro organismo come per esempio batteri, piante o alghe le quali possono ricavare energia direttamente dalla fotosintesi. Inoltre altri organismi non vengono predati da nessun altro organismo probabilmente perché tali organismi rappresentano le specie principalmente animali in cima alla catena alimentare come coccodrilli o altri grandi predatori. Queste due condizioni portano i parametri del modello

riguardanti l'effetto mittente o destinatario a seconda del caso ad assumere valori infinitamente negativi nell'implementazione del metodo "ergm". Per questo non possiamo utilizzare questo modello per descrivere la nostra rete in maniera adeguata e dobbiamo scartare l'ipotesi di usare parametri relativi all'effetto mittente o destinatario dei nodi nei prossimi modelli.

5.4 Dyad Independence/P1-model

Il prossimo modello che andremo a considerare come intuibile dal nome rilassa le assunzioni dei precedenti modelli assumendo che ciascuna diade ossia ciascuna coppia di variabili (Y_{ij}, Y_{ji}) relativa ai due archi di verso opposto presenti tra i nodi i e j sia indipendente da tutte le altre con una distribuzione multinomiale tale che:

$$D_{ij} \sim Multinomial(1, \pi_{ij})$$

Con parametro π_{ij} contenente la probabilità associata a ciascuna delle quattro possibili configurazioni delle diadi. Tale modello ci consente di considerare la reciprocità della rete che nel caso di una catena alimentare come abbiamo osservato anche nelle sezioni precedenti è praticamente inesistente. Però affinché il modello risulti identificabile abbiamo dovuto stimare un modello chiamato "P1 model" che pone ulteriori assunzioni come il fatto che il parametro relativo alla reciprocità è unico per ogni modello. Infine, a differenza della specificazione modello P1 abbiamo dovuto rimuovere dal modello gli attributi relativi all'effetto mittente e destinatario dei nodi per quanto visto nella sezione precedente. La funzione di verosimiglianza del modello che andremo a stimare sarà seguendo la notazione ERGM pari a:

$$P(Y = y) = k(\theta)^{-1} \exp(g(y)' \theta)$$

In cui le statistiche sufficienti contenute in $g(y)$ saranno quindi il numero totale di relazioni della rete ed il numero di relazioni reciprocate mentre in θ troveremo i parametri associati a tali statistiche. I risultati che abbiamo ottenuto adattando tale modello con il metodo "ergm" alle nostre reti sono riportati di seguito:

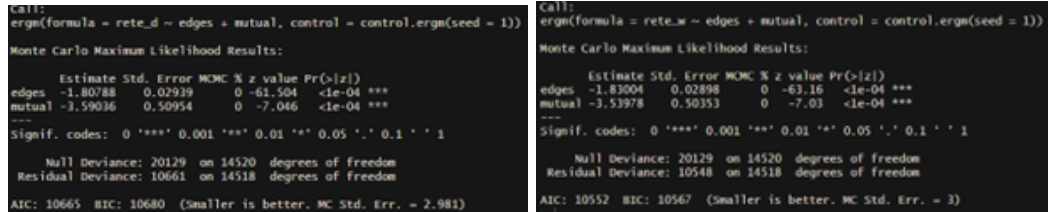


Figure 5: Risultati dell'applicazione del modello P1

Notiamo che come atteso il parametro associato alle reciprocità risulta essere negativo e molto significativamente diverso da 0. Questo dimostra che le reti hanno una forte tendenza a non avere reciprocità ossia è estremamente improbabile che un predatore venga predato dalla sua stessa preda. I valori dei due indici AIC e BIC invece non sembrano decrescere di molto, segno che l'adattamento del modello ai nostri dati non è molto migliorato rispetto al Simple Random Graph Model. Inoltre, simulando mille reti notiamo come questi due modelli non siano ancora in grado di rappresentare adeguatamente le statistiche di transitività e deviazione standard dell'in-degree e dell'out-degree delle reti come osserviamo dai grafici sotto riportati.

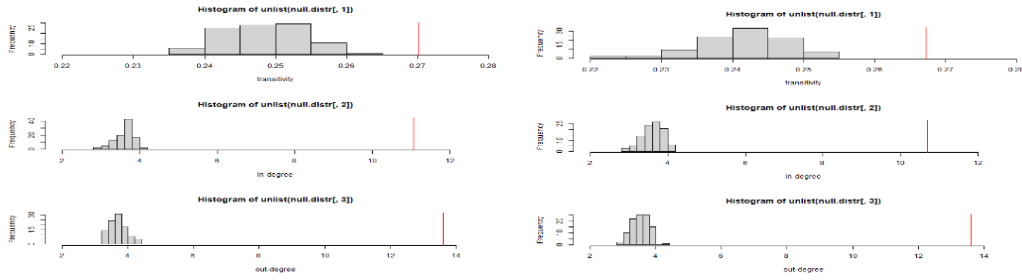


Figure 6: Confronto tra le statistiche delle nostre reti e quelle simulate

Anche in questo caso, la linea rossa che rappresenta il valore osservato della statistica nella rete è ben distante dai valori predetti dalle mille simulazioni.

5.5 Markov Graph Model

Il prossimo modello si basa sull'assunzione che due variabili Y_{ij} e Y_{kl} sono condizionatamente dipendenti tra loro se i rispettivi archi condividono un nodo ossia:

$$Y_{ij} \not\perp Y_{kl} \quad \text{se} \quad \{ij\} \cap \{kl\} \neq \emptyset$$

In pratica grazie al teorema di Hammersley-Clifford e quindi alle configurazioni della rete che creano cricche (clique) nel grafo delle dipendenze sappiamo che questo modello avrà come statistiche sufficienti anche il numero di 2-stelle, il numero di triangoli e il conteggio di ogni stella di grado fino a $n - 1$ dove n è il numero di nodi nel grafo. Solitamente tuttavia si tende a considerare le stelle di grado al massimo 3. Quindi nel nostro caso utilizzando la notazione del ERGM per la funzione di verosimiglianza del modello si avrà:

$$P(Y = y) = k(\theta)^{-1} \exp(g(y)' \theta)$$

Dove $g(y)$ rappresenta il vettore delle statistiche sufficienti contenente per l'appunto anche il numero di 2-stelle, il numero di triangoli e il conteggio di ogni stella di grado fino a $n - 1$ oltre al numero di relazioni reciprocate ed il conteggio degli archi che avevamo visto nel precedente modello. Purtroppo questo modello rischia di degenerare molto facilmente in quanto piccoli cambiamenti nei valori dei parametri del modello portano a enormi cambiamenti nei valori attesi delle statistiche della rete secondo il modello. Questo accade perché il modello pone quasi tutta la massa di probabilità in poche delle possibili configurazioni della rete che di solito corrispondono con il grafo vuoto in cui non si osservano archi o il grafo completo in cui ogni possibile arco tra due nodi è presente. Sfortunatamente questo è il caso anche per il nostro modello e per questo il metodo “ergm” di R non è in grado di adattare questo modello alle nostre reti come possiamo notare dai seguenti messaggi di errore ottenuti dopo quasi un'ora di attesa:

```
Iteration 11 of at most 60:
Error in ergm.MCMC(init, s, s.obs, control = control, verbose = verbose, :
Unconstrained MCMC sampling did not mix at all. Optimization cannot continue.
In aggiunta: Messaggio di avvertimento:
In ergm.MCMC.sample(s, control, theta = mcmc.init, verbose = max(verbose, :
Unable to reach target effective size in iterations allotted.

Error in ergm.MCMC(init, s, s.obs, control = control, verbose = verbose, :
Unconstrained MCMC sampling did not mix at all. Optimization cannot continue.
In aggiunta: Messaggio di avvertimento:
In ergm.MCMC.sample(s, control, theta = mcmc.init, verbose = max(verbose, :
Unable to reach target effective size in iterations allotted.
```

Figure 7: Messaggi di errore dell'adattamento del Markov Graph Model

Infine, spesso viene inserito nel modello al posto delle singole statistiche relative ai conteggi di ogni grado delle k -stelle una singola statistica relativa alle k -stelle alternanti per combattere il problema della degenerazione del modello tuttavia nel nostro caso neanche inserendo tale statistica siamo riusciti a adattare dei modelli alle nostre reti per lo stesso problema di prima.

5.6 Social Circuit Model

Per risolvere i problemi riscontrati con il precedente modello utilizziamo l'assunzione di dipendenza condizionata parziale che afferma che due variabili rappresentanti due archi nel grafo possono essere condizionatamente dipendenti in base all'esistenza di altri archi nel grafo. In particolare noi andremo ad adattare un modello che utilizza l'assunzione di dipendenza dei circuiti sociali che afferma che due variabili che rappresentano due archi distinti Y_{ij} e Y_{hl} sono condizionatamente dipendenti se $Y_{ih} = Y_{jl} = 1$ oppure se $Y_{il} = Y_{jh} = 1$ ossia se ciascuno dei due nodi dell'arco (ij) è collegato ad un nodo diverso dell'arco (hl). In pratica dovremo aggiungere al modello statistiche sufficienti riguardanti i k -triangoli ed i cammini 2- k . Tuttavia facendo ciò ricadremo nuovamente negli stessi problemi di degenerazione del modello visti prima per questo dovremo invece aggiungere al modello le statistiche alternanti per i k -triangoli e per cammini 2- k . I risultati dell'applicazione di tale modello sono di seguito riportati:

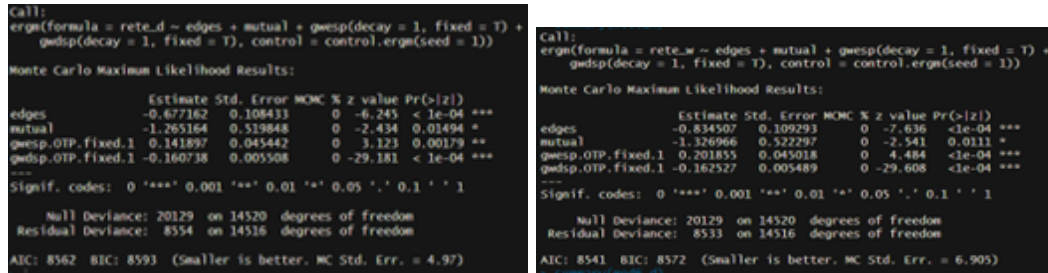


Figure 8: Risultati dell'applicazione del Social Circuit Model

Notiamo come tutti i parametri risultino essere molto significativamente diversi da zero tranne per il parametro associato alla reciprocità delle reti che diventa solo significativamente diverso da zero adottando la soglia del 5. Inoltre pur non essendo le nostre reti di tipo sociale questo modello risulta più coerente con i nostri dati sia secondo l'indice AIC che secondo il BIC che risultano nettamente inferiori rispetto ai modelli visti in precedenza. Anche osservando le statistiche delle reti simulate dai due modelli notiamo che almeno per quanto riguarda la deviazione standard dell'in-degree e la transitività della rete riguardante la stagione umida tali modelli sembrano molto più adeguati alle nostre reti.

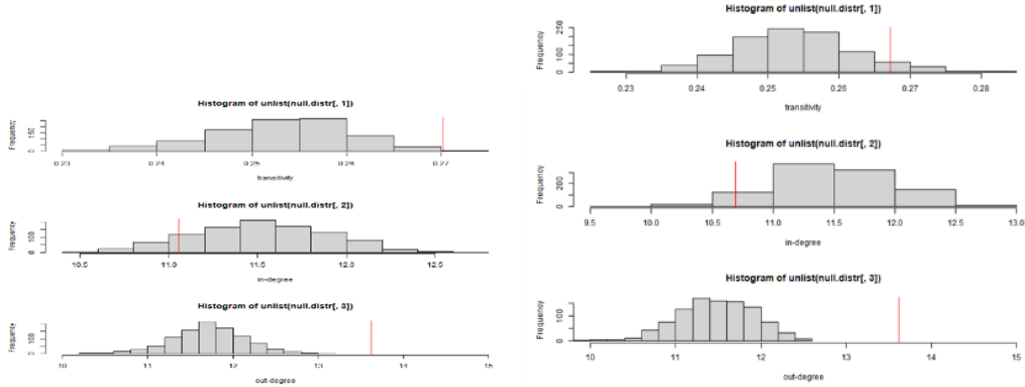


Figure 9: Confronto tra le statistiche delle reti simulate e le reti osservate

Anche andando ad osservare il p-value si osserva che i valori della deviazione standard dell'in-degree di entrambe le nostre reti e la transività della rete riguardante la stagione umida non siano significativamente più estreme rispetto alla distribuzione di tali statistiche nelle reti simulate anche usando una soglia del 5% per il rifiuto dell'ipotesi.

6 Stochastic Block Model

L'ultimo modello che vogliamo adattare alle nostre due reti è lo “Stochastic Block Model” che suddivide l'insieme dei nodi della rete in blocchi di nodi che si comportano in maniera simile ossia che si connettono in maniera simile con i nodi di un altro blocco. Nelle nostre due reti questo ci permette di dividere gli organismi in gruppi che tendono a predare o ad essere predati dagli stessi gruppi di organismi.

La prima cosa da fare sarebbe selezionare il numero adatto di blocchi confrontando i valori del criterio ICL ottenuti da modelli con numero di blocchi differenti tuttavia il comando “estimateSimpleSBM” disponibile nel pacchetto “sbm” di R permette di fare questo confronto in automatico. Nel nostro caso abbiamo ottenuto i seguenti risultati:

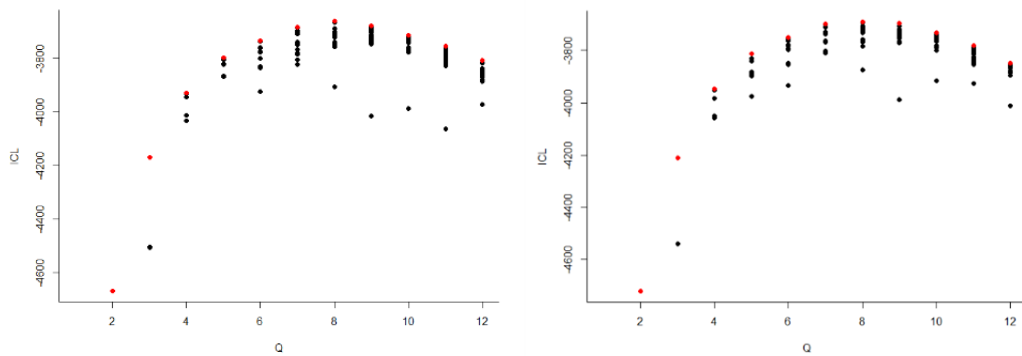


Figure 10: Valori del criterio ICL per differenti numeri di blocchi.

Notiamo come per entrambe le reti il numero di blocchi che porta ad ottenere valori di ICL migliori sia otto anche se di poco rispetto ad altri possibili valori. Deciso ciò è possibile osservare i risultati dell'adattamento di tale modello alle nostre due reti. Di seguito riportiamo i parametri stimati dal modello che indicano la probabilità attesa di osservare una relazione tra nodi appartenenti a due blocchi: in particolare le colonne indicano le probabilità di un organismo appartenente a tale gruppo di predare uno del gruppo appartenente alla riga corrispondente.

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]
[1,]	0.111	0.003	0.471	0.002	0.488	0.236	0.859	0.329
[2,]	0.103	0.040	0.679	0.001	0.387	0.283	0.074	0.004
[3,]	0.006	0.002	0.012	0.001	0.338	0.009	0.571	0.682
[4,]	0.230	0.303	0.068	0.053	0.011	0.210	0.030	0.005
[5,]	0.008	0.001	0.014	0.000	0.068	0.006	0.029	0.508
[6,]	0.010	0.001	0.014	0.000	0.141	0.042	0.072	0.313
[7,]	0.001	0.000	0.001	0.000	0.001	0.001	0.002	0.073
[8,]	0.000	0.000	0.001	0.000	0.001	0.001	0.001	0.056

Figure 11: Parametri stimati del modello per le due reti.

Notiamo subito come ci siano alcuni gruppi che tendono a non predare alcun altro gruppo di organismo in particolare il quarto. A tale gruppo appartengono, se ne osserviamo la composizione sotto riportata identica per entrambe le reti, principalmente batteri o alghe che non necessitano di nutrirsi di nessun altro organismo per poter sopravvivere.

[1]	"2um Spherical Phytoplankt"	"Synedococcus"	"oscillatoria"	"Small diatoms (<20um)"
[5]	"Big Diatoms (>20um)"	"Dinoflagellates"	"other Phytoplankton"	"Benthic Phytoplankton"
[9]	"Thalassia"	"Halodule"	"Syringodium"	"Drift Algae"
[13]	"Epiphytes"	"Free Bacteria"	"water Flagellates"	"water ciliates"
[17]	"Benthic Flagellates"	"Benthic ciliates"	"Meiofauna"	"Green Turtle"

Figure 12: Composizione del quarto gruppo per entrambe le reti.

Un altro gruppo che risulta interessante è l'ottavo che tende a predare quasi tutti gli altri gruppi tranne quelli composti di microorganismi o organismi appartenenti sempre a tale gruppo. Questo gruppo potrebbe essere definito come quello dei superpredatori: principalmente animali che si nutrono di numerose altre specie e difficilmente vengono predati da altri animali. Riportiamo di seguito la composizione di tale gruppo per entrambe le reti:

[1]	"Sharks"	"Tarpin"	"Grouper"	"Jackers"	"Barracuda"
[6]	"Barracuda"	"Loon"	"Green"	"Coonard"	"Big herons & egrets"
[11]	"Big herons & egrets"	"Predatory ducks"	"Raptors"	"Kingfisher"	"Crocodiles"
[16]	"Delphin"				

Figure 13: Composizione dell'ottavo gruppo per entrambe le reti.

Notiamo come atteso che tale gruppo è composto principalmente da predatori come squali, coccodrilli, delfini, rapaci, barracuda che per l'appunto sono spesso definiti come superpredatori, insieme però anche ad altri animali meno attesi come i tarponi, i pellicani o le anatre predatrici. Per quanto riguarda la dimensione dei blocchi non sembrano esserci blocchi estremamente più popolosi di altri ma sembra invece che tutti i gruppi abbiano dimensioni simili come possiamo osservare dalle proporzioni sotto riportate che corrispondono anche alla probabilità a priori di assegnare un organismo ad uno dei possibili blocchi:

[1]	0.10763347	0.09934553	0.14847048	0.16421137	0.17299043	0.07527693	0.11588193	0.11616995
[6]	0.10757037	0.13224396	0.09939813	0.16489748	0.14891548	0.10733794	0.10809115	0.13154549

Figure 14: Proporzioni dei blocchi per entrambe le reti.

Per osservare meglio i comportamenti degli organismi di ciascuno di questi gruppi riportiamo di seguito dei grafici che riassumono i rapporti di predazione tra gli organismi in base al blocco di appartenenza.

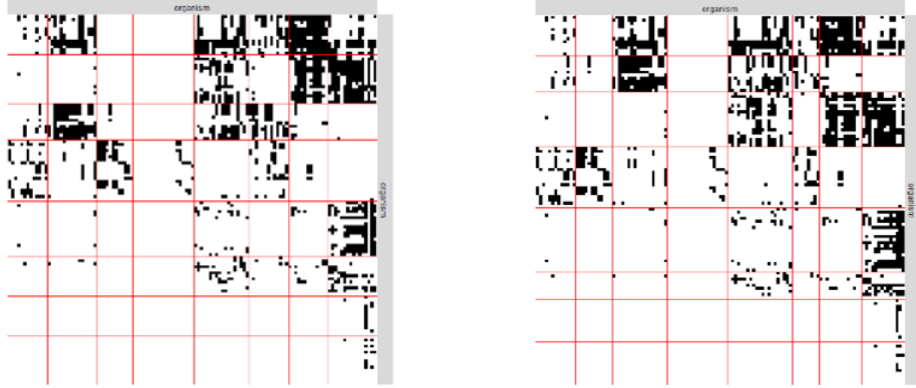


Figure 15: Grafico dei dati osservati (ordinati per blocco)

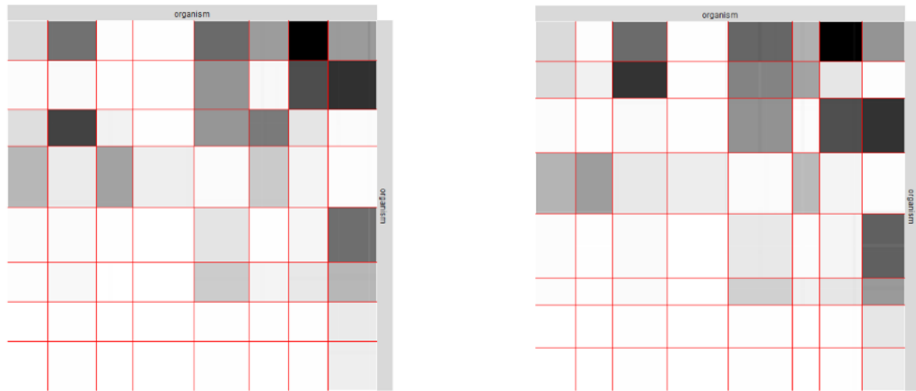


Figure 16: Grafico delle probabilità di connessione stimate tra i blocchi

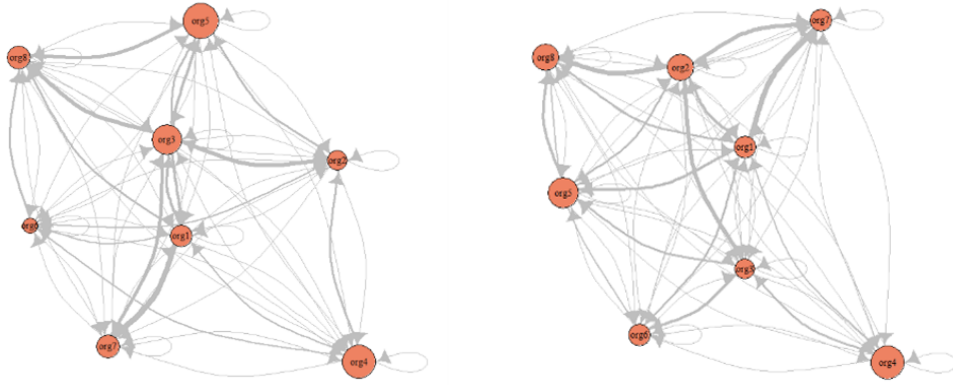


Figure 17: Grafico delle probabilità di connessione stimate (meso)

Notiamo in particolare dall'ultimo grafico che schematizza l'intensità delle relazioni tra nodi appartenenti ai vari gruppi come effettivamente l'ottavo gruppo riesca a predare buona parte degli altri gruppi mentre altri gruppi come il settimo sembrano sì predare molto anche loro ma principalmente organismi appartenenti al gruppo uno o due. Infine riportiamo anche il grafo delle due reti con i nodi colorati in base al gruppo di appartenenza:



Figure 18: Grafo delle due reti con i nodi colorati in base al gruppo di appartenenza.

Notiamo come effettivamente i nodi appartenenti al quarto gruppo sembrano più isolati rispetto agli altri nodi della rete in quanto essi tendono a predare meno rispetto agli altri organismi.

7 Conclusioni

Nel nostro lavoro abbiamo potuto osservare numerosi aspetti di rilievo delle nostre reti che potrebbero rivelarsi utili per preservare l'ecosistema delle baie della Florida meridionale. In particolare abbiamo osservato, grazie al bassissimo livello di reciprocità e al segno del parametro associato a tale statistica nei modelli, come sia estremamente raro osservare una coppia di organismi che si predano a vicenda. Invece, osservando la transitività, abbiamo osservato come non sia così raro che un predatore si nutra anche delle prede delle sue prede.

Inoltre, grazie ai valori di assortività ed alla non significativa differenza da zero dei parametri ad esso associati nei modelli, abbiamo osservato che l'attributo nodale relativo alla biomassa non sembri avere un'influenza rilevante sui rapporti di predazione tra gli organismi.

Abbiamo inoltre osservato, grazie ai bassi valori degli indici di centralità in generale, come non esistano organismi più importanti di altri all'interno della catena alimentare. Ciò significa che ciascun organismo è ugualmente importante all'interno della catena alimentare e che quindi è necessario proteggere tutti gli organismi viventi allo stesso modo, anche quelli che sembrano meno significativi come alghe o batteri.

Inoltre, abbiamo visto che il modello che si adatta meglio ad entrambe le nostre due reti è il “social circuit model” che tuttavia non è in grado di replicare adeguatamente tutte le statistiche della nostra rete nelle reti simulate.

Infine, abbiamo anche adattato il modello stocastico a blocchi ed abbiamo osservato che tale modello distingue gli organismi in otto blocchi con abitudini predatorie simili e che vengono predati dagli stessi gruppi di organismi. In particolare, abbiamo osservato un gruppo composto principalmente da microorganismi che sembrano non predare quasi nessun altro organismo ed un altro invece composto da superpredatori che si nutrono di quasi tutti gli altri tipi di organismi e che difficilmente vengono predati a loro volta.

Tutto questo dimostra come lo studio della catena alimentare attraverso tecniche di analisi e di modellazione delle reti possa rivelarsi utile per studiare, catalogare e proteggere gli organismi di un determinato ecosistema.