

# Buone notizie per gli amanti del cocktail di gamberi

Giovanna Jona Lasinio

Dipartimento di Scienze Statistiche  
Università Sapienza di Roma

giovanna.jonalasinio@uniroma1.it

## 1 Introduzione

La pesca in Italia svolge un ruolo ancora oggi economicamente importante e per questa ragione viene costantemente monitorata da molti punti di vista e da numerosi enti scientifici. La dinamica di popolazione delle specie sfruttate, la loro distribuzione nel Mediterraneo al variare degli anni e gli effetti della pressione cui sono sottoposte, sono costantemente monitorate in base a regolamenti comunitari e a piani di raccolta dati gestiti dal Ministero delle Politiche Agricole, Alimentari e Forestali (MIPAAF). Al centro di questo grande movimento di informazioni si collocano i biologi della pesca in particolare coloro che si occupano della valutazione dello stato delle specie ittiche di interesse commerciale come il nasello (*Merluccius merluccius*), la triglia (*Mullus barbatus*) e il gambero rosa (*Parapenaeus longirostris*).

Al fianco dei biologi ormai costantemente sono presenti statistici che sviluppano strumenti di gestione dell'informazione atti a rispondere alle numerosissime domande, sia scientifiche, che gestionali, a cui i gruppi di lavoro sono chiamati a rispondere dalle autorità.

Personalmente collaboro, ormai da diversi anni, con il gruppo del professor Giandomenico Ardizzone<sup>1</sup>, con Paolo Sartor del CIBM<sup>2</sup> e con Francesco Colloca ora presso l'Istituto per l'Ambiente Marino Costiero del CNR di Mazara del Vallo (si vedano ad esempio i lavori [1] e [5]). Alcuni dei progetti nei quali abbiamo lavorato insieme sono stati: la realizzazione del sistema GIS-Pesca per il MIPAAF, l'individuazione delle aree di concentrazione dei giovanili<sup>3</sup> di specie

commerciali su scala nazionale (progetto NURSERY) e mediterranea (progetto europeo MEDISEH: Mediterranean Sensitive Habitats<sup>4</sup>). Negli studi realizzati insieme la domanda che ci siamo più spesso posti è stata:

in che modo varia nello spazio e nel tempo la densità<sup>5</sup> di individui distinti per età delle varie specie?

Questa domanda diventa cruciale quando si vogliono prendere decisioni per la conservazione del patrimonio ittico. Capire dove si concentrano i giovanili e i riproduttori di una specie può servire a decidere quali aree chiudere alla pesca in determinati periodi dell'anno. Analizzare come l'intero stock si modifichi nel tempo permette di capire quali specie sono in crescita o in diminuzione sia per cause naturali che per la pressione della pesca.

Il ruolo dello statistico in questo contesto non è semplice, per prima cosa deve capire di cosa si sta parlando, imparare un pochino di biologia marina, se non altro per riuscire a tradurre in modello statistico quanto i biologi vanno spiegando; a seguire deve vedersela con i dati a disposizione, prima compila una lista di "desiderata" e poi deve venire a patti con la realtà ovvero con ciò che è realmente disponibile.

Qui di seguito racconterò brevemente un caso di studio, ancora in corso di elaborazione, nel quale sono state studiate molte delle specie d'interesse commerciale e dato che amo le belle notizie, racconterò dell'unica specie che vede la propria consistenza in crescita, ovvero il *Parapenaeus*

---

dell'individuo. Usualmente si tratta di individui di piccole dimensioni non ancora in grado di riprodursi.

<sup>4</sup> <http://mareaproject.net/contracts/5/overview/>.

<sup>5</sup> per densità di individui si intende il rapporto tra numero di individui e area campionata espressa in chilometri quadrati.

---

<sup>1</sup> Università Sapienza di Roma

<sup>2</sup> Centro Interuniversitario di Biologia Marina ed Ecologia Applicata

<sup>3</sup> la definizione di *giovanile* di una specie viene di solito data in base alla taglia e alla maturità sessuale

*longirostris* (Figura 1) meglio noto come gambero rosa o ingrediente fondamentale del cocktail di gamberi.

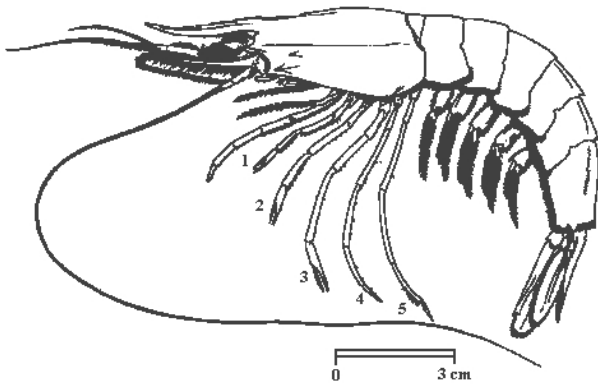


Figura 3 Gambero Rosa o *Parapeneus longirostris*

## 2 Dati disponibili e dati necessari

Cosa occorre per studiare esaurientemente il comportamento spazio temporale di specie ittiche? Prima di tutto occorre conoscere il numero di individui, corredato da loro peso e lunghezza, in un numero “elevato” di punti nelle aree di interesse. Poi occorrono informazioni sulla configurazione dell’area (batimetria, inclinazione del fondale etc), informazioni sull’ambiente marino (temperatura a varie profondità e in vari periodi dell’anno, salinità, nutrienti etc.), sulle correnti, sui fenomeni oceanografici che si verificano e/o si possono verificare (ad esempio fenomeni di risalita di acque profonde) e soprattutto, occorrono informazioni sulla pressione della pesca (o sforzo di pesca) nell’area. In realtà l’Italia è in una situazione molto migliore di tanti paesi europei rispetto alla disponibilità di dati sullo stock ittico. Infatti è dal 1994<sup>6</sup> che l’Italia conduce campagne di pesca mirate allo studio delle specie ittiche d’interesse commerciale nei paesi del Mediterraneo partecipando al progetto MEDITS<sup>7</sup>; in questo progetto si adotta un preciso disegno campionario basato sulla profondità, si utilizzano pescherecci muniti degli stessi attrezzi per la pesca a strascico e la campagna di pesca viene condotta più o meno sempre nello stesso periodo dell’anno.

<sup>6</sup> Dati affidabili sono in realtà disponibili dall’anno successivo

<sup>7</sup> Mediterranean International Bottom Trawl Survey, <http://www.sibm.it/SITOS/%20MEDITS/principalem%20edits.htm>.

Il lavoro dietro la produzione dei dati sugli stock ittici è enorme, basti pensare che è necessario per ogni pescata misurare, pesare e identificare ogni singolo pesce, crostaceo o cefalopode catturato e questo viene fatto a mano in almeno un centinaio di punti per ciascuna area.

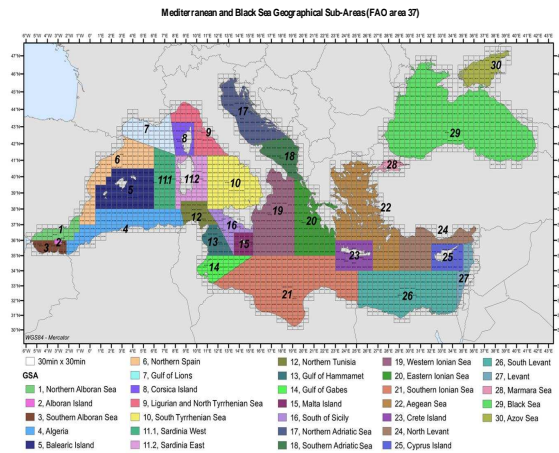


Figura 4 Suddivisione FAO-GFCM del Mediterraneo

Le informazioni morfologiche ed ambientali sono meno ricche, le temperature a diverse profondità sono disponibili solo su scale spaziali molto ampie e sono calcolate tramite modelli matematici, per tanto viene di solito utilizzata la sola temperatura superficiale che, come molti altri indicatori, è reperibile dalle rilevazioni satellitari (temperatura superficiale, salinità e clorofilla, quest’ultima solo dal 2002), batimetria e inclinazione del fondale sono disponibili dal database del progetto Marspec<sup>8</sup>. Purtroppo mancano un buon indicatore dello sforzo di pesca e informazioni sui fenomeni oceanografici. Questo lavoro ha riguardato il Tirreno settentrionale o GSA9 secondo la suddivisione GFCM<sup>9</sup> (General Fisheries Commission for the Mediterranean) dell’area Mediterranea (Figura 2), in quest’area come nelle altre, tutte le informazioni richiamate sopra hanno spesso risoluzioni spaziali diverse, quindi il primo problema da affrontare è l’allineamento dei dati nello spazio. In questo caso abbiamo scelto di allineare tutto alla griglia di rilevazione delle temperature che ha celle di circa 1km di lato per un totale di 61679 celle nell’area d’interesse. L’allineamento è stato realizzato tramite interpolazione spaziale (kriging) [3]. I dati delle

<sup>8</sup> <http://www.marspec.org/>

<sup>9</sup> maggiori dettagli sono disponibili sul sito <http://www.gfcm.org/gfcm/en>.

campagne MEDITS utilizzati partono dal 1994 fino al 2010, con 153 siti campionati (cale) dal 1994 al 2001 e 120 cale dal 2002 al 2010. Per ciascuna specie si è lavorato distinguendo per età dell'individuo (giovani e riproduttori), di conseguenza il numero di campioni con valore nullo è spesso più del 50% del totale annuo. Per il gambero al contrario delle altre specie rilevate, si nota un aumento costante del numero di campioni non nulli (giovani più riproduttori) e questo in tutta l'area considerata, passando dal 29.41% di cale positive del 1995 al 66.67% del 2010, aumento che corrisponde anche ad una crescita della consistenza complessiva dello stock.

### 3 I modelli

Per cercare di capire meglio cosa stesse succedendo al gambero abbiamo lavorato su due aspetti in parallelo, da un lato abbiamo costruito un modello che permettesse di descrivere e mappare la densità di individui e dall'altro un modello analogo che permettesse di prevedere (e quindi mappare) la probabilità di presenza del gambero (sia riproduttori che giovani). La classe dei modelli a cui abbiamo fatto riferimento è quella dei modelli additivi generalizzati (GAM) come proposti in [4], seguendo l'implementazione proposta nella libreria mgcv da Wood [6]. Dopo un'accurata selezione i modelli scelti sulla base sia delle capacità predittive che di adattamento, utilizzano entrambi come variabili esplicative il tempo (anno), le coordinate spaziali espresse in coordinate metriche (UTM), e la temperatura minima superficiale del secondo trimestre dell'anno di rilevazione; la probabilità di presenza dipende anche dalla profondità e dall'inclinazione del fondale. La scelta dei GAM non è ottimale per quanto riguarda la mappatura della densità di individui, questa viene meglio rappresentata con un modello più complesso che però richiede tempi di calcolo molto più lunghi. Il GAM dall'altra parte ci ha permesso di rappresentare le relazioni non lineari tra le variabili esplicative e la risposta, sia che fosse densità di individui che presenza/assenza, in modo efficace e computazionalmente efficiente. Quest'ultimo aspetto è di grande importanza quando pensiamo che la previsione a fini di mappatura, sia per la densità di individui che per le probabilità di presenza, andava condotta su 17 griglie di 61679 celle, quindi su 1048543 punti!

Se indichiamo con  $Z$  la densità di individui complessiva (giovani e riproduttori), il modello scelto è il seguente

$$\log(Z_{ij} + 1) = \beta_0 + s(j) + s(X_{ij} + Y_{ij}) + s(temp_{ij}) + \varepsilon_{ij} \quad (1)$$

dove  $\beta_0$  è un termine costante,  $s(\cdot)$  una funzione spline [2] di argomento una o più variabili esplicative,  $X_{ij}$  e  $Y_{ij}$  sono le coordinate del punto  $i$  ( $i = 1, \dots, n_j$ ) nell'anno  $j$  ( $j = 1994, 1995, \dots, 2010$ ),  $temp$  è la temperatura di cui si è già detto e  $\varepsilon_{ij}$  è un termine di errore normale con media nulla e varianza che varia al variare dell'anno. Il modello ci racconta come la relazione tra densità di individui e il tempo sia non lineare e crescente (figura 3 (b)), mentre quella con lo spazio, sempre non lineare (figura 3 (a)), evidenzia come la specie sia stanziale (la relazione non varia nella forma nel tempo) e infine come preferisca zone con temperature minime più elevate.

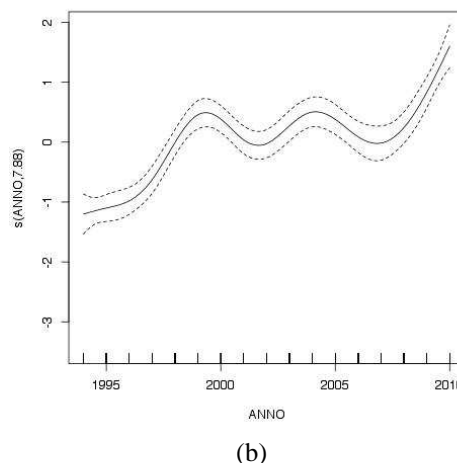
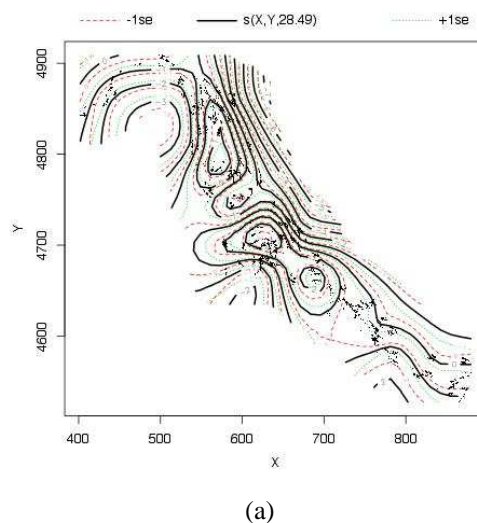


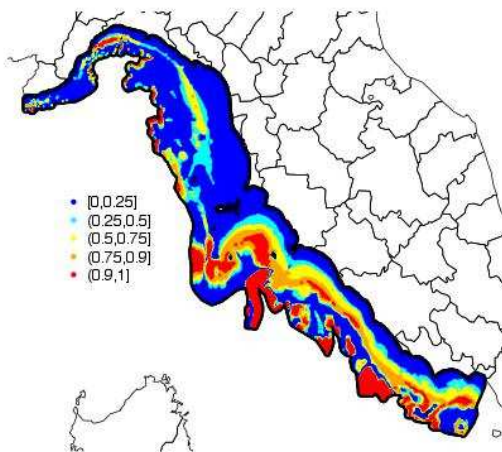
Figura 5 Effetto non lineare dello spazio (a) e del tempo (b) sulla densità di individui del gambero rosa stimati con il modello (1)

Il secondo modello riguarda le probabilità di presenza che sono rappresentate su scala *logit*, ovvero se  $p_{ij}$  è la probabilità che  $Z_{ij}$  sia maggiore di zero nel sito  $i$  nell'anno  $j$ , usando come variabile risposta  $\text{logit}(p_{ij}) = \log(p_{ij}/(1-p_{ij}))$  il modello scelto è:

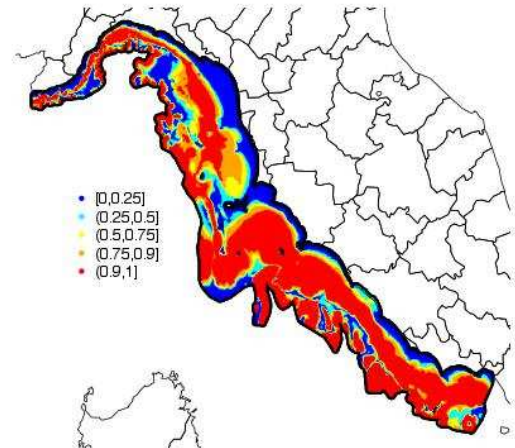
$$\text{logit}(p_{ij}) = \beta_0 + s(j) + te(X_{ij}, Y_{ij}) + s(\text{temp}_{ij}) + s(\text{prof}_i) + s(\text{incl}_i) + \varepsilon_{ij} \quad (2)$$

La probabilità di presenza per individui di qualsiasi età aumenta nel corso del tempo e con la temperatura minima più elevata, diminuisce in modo sostanziale oltre i 400 metri di profondità e su fondali con inclinazione (*incl*) oltre i 40 gradi<sup>10</sup>. La relazione con lo spazio evidenziata dal modello scelto è anch'essa non lineare e complessa<sup>11</sup>, ma stabile negli anni.

Costruendo la mappa stimata delle presenze (una presenza equivale a  $p_{ij} > 0.5$ ), possiamo vedere, confrontando con l'osservato, che il modello ha una leggera tendenza alla sovrastima delle presenze con un errore complessivo di classificazione sui 17 anni del 14% circa. Dalle figure 4 (a) e 4 (b) appare molto evidente come l'area in cui è probabile rinvenire gamberi (giovani o riproduttori) sia cresciuta molto dal 1995 al 2010 (zone in rosso indicano probabilità di presenza vicine a 0.9).



(a)



(b)  
Figura 6 Probabilità di presenza del gambero rosa nel Tirreno centro settentrionale negli anni 1995 (a) e 2010 (b)

Modelli analoghi costruiti per i soli giovanili del gambero mostrano anch'essi una crescita della probabilità di presenza in zone del Tirreno centro settentrionale in cui prima non erano presenti.

## 4 Conclusioni

In questo tipo di studi l'uso di modelli statistici appropriati permette di ottenere informazioni di grande rilievo per il decisore pubblico ed anche per lo scienziato. Dalle mappe costruite correttamente è possibile desumere in quali zone certe specie sono in crescita e in quali in diminuzione, eventualmente si prendono decisioni gestionali basandosi su di esse. Per lo scienziato le stesse mappe sono di grande interesse proprio perché forniscono importanti spunti di riflessione. Ad esempio la domanda che ora ci si pone rispetto al gambero rosa è come mai questa specie sia in crescita ed espansione nel Tirreno centro settentrionale. Un'ipotesi fatta è che stia aumentando la temperatura della zona. Come abbiamo però detto prima, le informazioni sulla temperatura non superficiale è di bassa qualità, da questa ciò che si desume va preso con molta cautela. Ad ogni modo sembra davvero che il *Parapenaeus longirostris* sia una specie che trae vantaggio da un innalzamento della temperatura globale ed in questo momento stiamo cercando di

<sup>10</sup> In linea con quanto noto della biologia della specie.

<sup>11</sup> infatti è definita come prodotto tensoriale di termini *smooth*.

convalidare questa ipotesi sempre attraverso l'uso di modelli statistici complessi.

Ma questa è ancora un'altra storia.

## 5 Ringraziamenti

Vorrei ringraziare Francesco Colloca e Gianluca Mastrantonio per le sempre utili discussioni e l'attenta lettura. Vorrei anche ringraziare Sandro Fassò per avermi convinto a scrivere queste pagine.

## 6 Riferimenti bibliografici

- [1] Colloca F., Bartolino V., Jona Lasinio G., Maiorano L., Sartor P., Ardizzone G.D.: Identifying fish nurseries using density and persistence measures. *Marine Ecology Progress Series* 381 ,287-296 (2009).
- [2] de Boor C.: *A Practical Guide to Splines* (revised edition), Springer Verlag (2001).
- [3] Gelfand A.E., Diggle P.J., Fuentes M., Guttorp P.: *Handbook of Spatial Statistics*, Chapman and Hall/CRC press (2010).
- [4] Hastie T. J., Tibshirani, R. J.: *Generalized Additive Models*. Chapman and Hall/CRC (1990)
- [5] Maiorano L., Bartolino V., Colloca F., Abella A., Belluscio A., Carpentieri P., Criscoli A., Jona Lasinio G., Mannini A., Pranovi F., Reale B., Relini G., Viva C., Ardizzone G.D.: Systematic conservation planning in the Mediterranean: a flexible tool for the identification of no-take protected areas. *Ices Journal of Marine Science*, 66, 137-146.10.1093/icesjms/fsn148 (2009).
- [6] Wood S.N.: Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 73 (1):3-36 (2011).