The deep-water rose shrimp in the Ionian Sea: a spatio-temporal analysis of zero-inflated abundance data

D'Onghia G., Maiorano P., Carlucci R., Tursi A.,
Dipartimento di Biologia, Università degli Studi di Bari "Aldo Moro"
g.donghia@biologia.uniba.it

Pollice A., Ribecco N., Calculli C., Arcuti S.,
Dipartimento di Scienze Statistiche "Carlo Cecchi",
Università degli Studi di Bari "Aldo Moro"

Abstract: In the ecological field, the sampling of abundance data is often characterized by the zero inflation of population distributions. Constrained zero-inflated GAM's (COZIGAM) are obtained assuming that the probability of non-zero inflation and the mean non-zero-inflated population abundance are linearly related. Models of this class have been applied to a spatio-temporal case study concerning the deep-water rose shrimp, *Parapenaeus longirostris* (Lucas, 1846). Abundance data were collected during 16 experimental trawl surveys conducted from 1995 to 2010 in the Ionian Sea. The sampling design adopted was random-stratified by depth, with proportional allocation of hauls to the area of each depth range and geographical sector. Density index (N/km^2) and length (mm) were considered for each haul identified by time, depth, geographic coordinates and geographical sector.

Keywords: Zero-inflated data, COZIGAM, GAM, density, size, *Parapenaeus longirostris.*

1. Introduction

In the ecological field, the sampling of abundance data is often characterized by the zero inflation of population distributions. Many Mediterranean species show such a distribution due to their adaptation to the variable environmental conditions. One of these is the deep-water rose shrimp, *Parapenaeus longirostris* (Lucas, 1846), widespread throughout the whole Mediterranean Sea at depths between 20 and 700 m. The Ionian Sea is a basin where this shrimp represents the bulk of the catch due to the trawl fishing carried out on between shelf edge and upper slope. Aspects of the distribution and population biology of this shrimp are reported in D'Onghia et al. (1998) and Abelló et al. (2002). Its spatio-temporal distribution in the Ionian Sea for the period 1995-2010 has been studied and the relevant results have been reported in this paper.

	Density		*Length*	
parameters	*estimate*	*p-value*	*estimate*	*p-value*
intercept			14.248	<0.000
$depth_f$(0,200]	5.240	<0.000		
$depth_f$(200,500]	5.348	<0.000		
$depth_f$(>500)	3.5254	<0.000		
depth			0.029	<0.000
alpha	- 1.030	<0.000		
delta1	1.004	<0.000		
delta2	0.666	<0.000		
smooth terms	*df*	*p-value*	*df*	*p-value*
s(lon, lat)	28.855	<0.000	18.111	0.005
s(year)	8.507	<0.000	7.091	<0.000

Table 1: COZIGAM's estimates for the density index, GAM's estimates for the length.

2. Materials and Methods

Abundance data were collected during 16 experimental trawl surveys conducted from 1995 to 2010 in the Ionian Sea as part of the international MEDITS project funded by EC (Bertrand et al., 2000). The samples analyzed come from a total of 1052 hauls carried out during day-light hours between 10 and 800 m in the spring season (May-June). The sampling design adopted was random-stratified by depth, with proportional allocation of hauls to the area of each depth range and geographical sector. Density index (N/km^2) and carapace length (mm) were considered for each haul identified by time, depth, geographic coordinates and geographical sector.
A general approach to zero-inflated data modeling consists in assuming the response distribution as a probabilistic mixture of a zero and a non-zero generating process. Zero-inflated general linear models (ZIGLM) can be readily extended to include smooth effects of covariates giving rise to ZIGAMs. A constrained zero-inflated GAM (COZIGAM) is obtained assuming that the probability of non-zero inflation and the mean non-zero-inflated population abundance are linearly related. In this paper an analysis of the density index based on COZIGAM's is proposed. As *P. longirostris* carapace length is not affected by zero-inflation, given that no measurements are available when the density index is null, this variable is analyzed in the GAM's framework. The R libraries `COZIGAM` (Liu and Chan, 2010) and `mgcv` (Wood, 2006) were used for the data analysis.

3. Results

Preliminary exploratory data analysis (not reported) showed a discontinuous higher presence of zeroes and small density values at lower (shallower than 200 m) and higher (deeper than 500 m) depths. This lead to considering the factorization of the depth variable accordingly in the model for the density index. We propose the following specification for the mean of the log-Gaussian non-zero generating process:

$$\mu = s(\text{lon},\text{lat}) + s(\text{year}) + \text{depth}_f$$

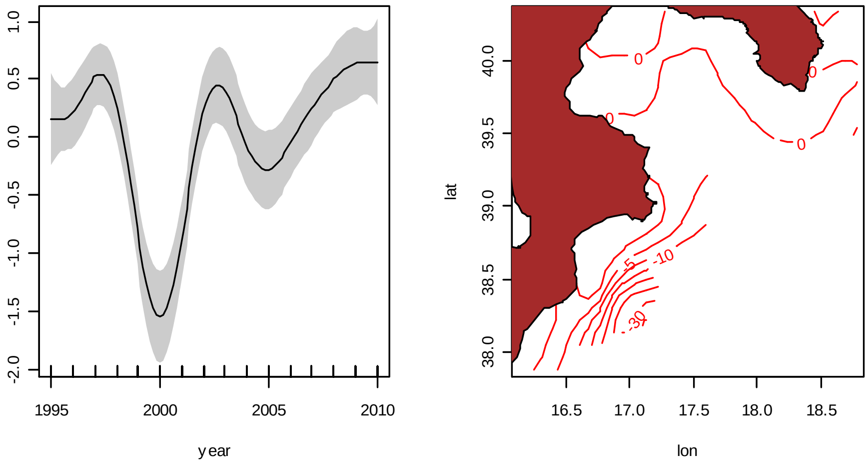

Figure 1: COZIGAM estimated effects of space and time for the *P. longirostris* density.

and assume that the smooth spatial effect and the temporal one have a different importance on the non-zero inflation probability, according to the following proportionality constraint:

$$logit(p) = \alpha + \delta_1\, s(\text{lon,lat}) + \delta_2\, s(\text{year}) + \text{depth}_f$$

In Tab. 1 we report the estimates of the COZIGAM model effects for the density index. The estimated effects of the three depth levels agree with the observed data. The estimates of δ_1 and δ_2 have significantly positive values, showing that the zero inflation probability decreases with the density value.
In Fig. 1, Left we report the estimate of the smooth temporal effect showing a severe drop of the density index in 2000. This was also expected according to the results of the preliminary exploratory data analysis which described a decreasing in the density index in the 1999-2001 years. The map of the spatial effect (Fig. 1, Right) reveals a wide distribution of the species along the Ionian arc with three main areas with a greater density.
A Gaussian GAM for the *Parapenaeus* carapace length is specified as follows:

$$\mu = intercept + s(lon,lat) + s(year) + depth$$

In this case the exploratory data analysis shows a continuous linear relation with the depth, leading to consider the unfactorized variable within a linear term; carapace lengths increase with deeper sea beds (Tab. 1). The smooth estimated temporal effect (Fig. 2, Left) has an opposite behavior with respect to the density index, showing a peak in 2000. Also a second peak in 2005 and a drop in 2008 are noticeable. The former could be due to a lower density of greater specimens, the latter could be in relation to an increase of juveniles in the sampled population. The map of the spatial effect shows greater sizes in the Gallipoli (Apulia) and Roccella Ionica (southern Calabria) fishery districts (Fig. 2, Right).

4. Concluding remarks

These results confirm the knowledge on density and size distribution of *P. longirostris* in the Ionian Sea (D’Onghia et al., 1998; Abello et al. 2002) revealing geographic and temporal effect both on density and size. The increasing density together with the

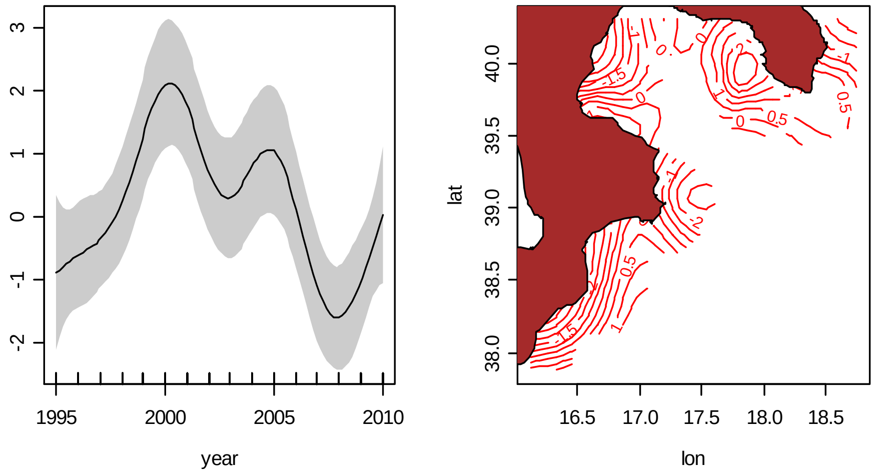

Figure 2: GAM estimated effects of space and time for *P. longirostris* carapace length.

decreasing size observed in some years could be related to the increase in the recruitment detected for the deep-water rose shrimp. This will require further investigation in order to identify the environmental variables affecting the changes observed in the species distribution.

References

Abelló P., Abella A., Adamidou A., Jukic-Peladic S., Maiorano P., Spedicato M. T. (2002) Geographical patterns in abundance and population structure of Nephrops norvegicus and Parapenaeus longirostris (Crustacea: Decapoda) along the European Mediterranean coasts, *Scientia Marina*, 66 (Suppl. 2), 125-141.

Bertrand J.A., Gil de Sola L., Papaconstantinou C., Relini G., Souplet A. (2000) An international bottom trawl survey in the Mediterranean: the MEDITS programme, *IFREMER Actes de Colloques*, 26, 76-93.

D'Onghia G., Matarrese A., Maiorano P., Perri F. (1998) Valutazione di Parapenaeus longirostris (Lucas, 1846) (Crustacea, Decapoda) nel Mar Ionio. *Biologia Marina Mediterranea*, 5 (2), 273-283.

Liu H., Chan K. S. (2010) Robust Introducing COZIGAM: An R Package for Unconstrained and Constrained Zero-Inflated Generalized Additive Model Analysis, *Journal Statistical Software*, 35(11), 1-26.

Liu H., Ciannelli L., Decker M. B., Ladd C., Chan K. S.. (2010) Nonparametric Threshold Model of Zero-Inflated Spatio-Temporal Data with Application to Shifts in Jellyfish Distribution, *Journal of Agricultural, Biological, and Environmental Statistics*.

Wood S. N. (2006). *Generalized Additive Models: An Introduction with R*, Chapman & Hall/CRC.