Estimation of coral growth parameters via Bayesian hierarchical non-linear models

Stima dei parametri di crescita di coralli tramite modelli bayesiani gerarchici non lineari

Crescenza Calculli, Barbara Cafarelli and Daniela Cocchi

Abstract In Ecology, the von Bertalanffy growth function (VBGF) is the standard model used to investigate the body growth of marine species. The parameters of this function are usually estimated by a classical method that might induce bias in the results: this method does not allow to distinguish the variability at individual or population level nor to take into account the contribution of environmental factors. A Bayesian hierarchical nonlinear model for estimating the VBGF parameters is proposed in order to overcome the limitations of the traditional method. The proposal improves both the statistical accuracy and the quantification of uncertainties affecting marine species growth. The proposal is assessed through a case study concerning two Mediterranean corals, *Balanophyllia europaea* and *Leptopsammia pruvoti*.

Riassunto *La funzione di von Bertalanffy (VBGF) è il modello standard usato in Ecologia per studiare la crescita degli individui nelle popolazioni marine. I parametri di questa funzione sono comunemente ottenuti tramite un metodo classico che può portare a errori delle stime, oltre a non permettere di modellare la variabilità disgiuntamente a livello di individuo o di popolazione, né di tenere in considerazione il contributo di fattori ambientali. Si propone un modello bayesiano gerarchico non lineare che permette di superare i limiti del metodo tradizionale e di ottenere un guadagno in termini di accuratezza statistica e di quantificazione delle componenti di incertezza che caratterizzano la crescita delle popolazioni marine. La proposta è valutata tramite un caso di studio riguardante due coralli mediterranei, Balanophyllia europaea e Leptopsammia pruvoti.*

Key words: Bayesian hierarchical models, Growth curves, Von Bertalanffy Growth Function

---

Crescenza Calculli
Dept of Economics and Finance, University of Bari, Italy, e-mail: crescenza.calculli@uniba.it

Barbara Cafarelli
Dept of Economics, University of Foggia, Italy e-mail: barbara.cafarelli@unifg.it

Daniela Cocchi
Dept of Statistical Sciences, University of Bologna, Italy e-mail: daniela.cocchi@unibo.it

1 Introduction

Marine biologists use individual demographic variables (age, oral disk length and body size) for modeling coral peculiarities, as well as their relationships with the environment. Individual coral age is usually obtained from coral body size using reliable growth models. The VBGF is a popular model for predicting the growth of marine organisms linking their lengths to ages by a non-linear relationship [4, 9]. For estimating the VBGF parameters, marine biologists use several methods that do not exploit statistical reasoning as they might. These methods are based on linear transformations of the VBGF, whose parameters are subsequently estimated by OLS. This approach has some limitations. Linear transformations are often proposed without accounting for measurement errors of observed data. Variability, at individual or population level, is neglected, inducing bias in parameter estimates and variance underestimation.

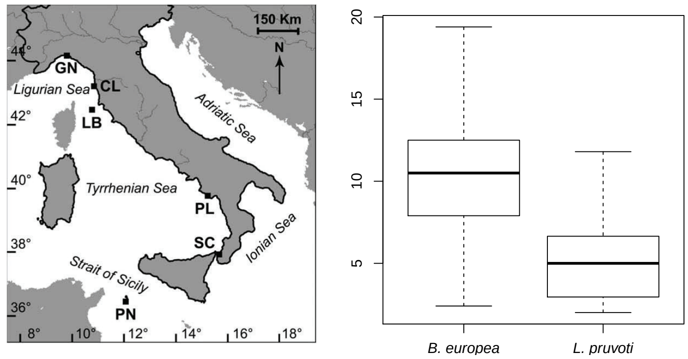
Environmental features at different sites are crucial in affecting coral characteristics. Unfortunately, these features are often unsuitable to this end, since they are measured as synthetic aggregates collected for other purposes: association of site-specific environmental measures to individual species data might be a forcing. The effort for considering environmental information, such as sea temperature, solar radiation, surface ocean acidification and anthropogenic stress, might be costly in terms of modeling and useless when interpreting results. For these reasons, a statistical model with random effects that avoid the explicit consideration of environmental covariates is suitable.

In particular, a feasible approach for estimation of coral growth parameters based on Bayesian hierarchical non-linear mixed effects models is proposed. Hierarchical modeling [2] is equivalent to handle VBGF in a two stage framework, accounting for two different sources of variability: the within-site and the between-site variations. The approach is applied to data of two solitary corals living in the Mediterranean sea, *Balanophyllia europaea* (*B. europaea*) and *Leptopsammia pruvoti* (*L. pruvoti*).

2 Motivating example

In this study, data on two species of solitary Mediterranean corals, *B. europaea* and *L. pruvoti*, are available. Both species live in rocky habitats and are widely distributed in the Mediterranean basin. Their growth and demographic peculiarities have been proved to be sensitive to changes in environmental conditions and, more generally, to the global warming [4, 9].

A dataset of 417 individuals combining 238 specimens of *B. europea* and 179 specimens of *L. pruvoti* is considered. In particular, specimens for the two species were collected during the same time interval (from 9$^{\text{th}}$ November 2003 to 30$^{\text{th}}$ September 2005) in the same sites as reported in Figure 1a [4, 8]. For each individual the main measurements available are the corallite length (L, in *mm*) which



(a) (b)

Fig. 1: (a) Site Locations: GN, Genova; CL, Calafuria; LB, Elba Isle; PL, Palinuro; SC, Scilla; PN, Pantelleria Isle. (b) Coral species length (in *mm*) distributions

represents the maximum axis of the oral disc and the ages in years obtained counting bands of the skeleton via computerized tomography scans (CTSs).

Table 1: Descriptive features the coral species and annual averages of environmental indicators per site. R: Solar Radiation (from 190 W/m^2); T: Sea Surface Temperature (from 18°C)

	GN		CL		LB		PL		SC		PN	
	B. europea	*L.pruvoti*	*B. europea*	*L.pruvoti*	*B. europea*	*L.pruvoti*	*B. europea*	*L.pruvoti*	*B. europea*	*L.pruvoti*	*B. europea*	*L.pruvoti*
n	42	30	34	29	34	30	54	30	32	30	42	30
Mean age (years)	7.4	6.2	5.5	4.1	4.6	8.1	6.9	5.6	6.2	7.8	5.2	7
Mean length (mm)	11.7	4.8	8.3	3.9	9	6.3	9.9	4.4	9.9	6.3	8.8	4.8
R (W/m^2), annual mean	166.95		170.07		172.74		181.48		187.31		192.95	
T (°C), annual mean	19.56		18.02		18.74		19.14		19.54		19.88	

Figure 1b shows differences in terms of length between the the two species. Summary features of the two species and the different environmental site conditions are reported in Table 1 where some differences among sites are appreciable. The measurement of environmental site conditions, even if in principle very stimulating, is currently performed under very general purposes, not directly linked to coral data collection. For this reason, they will not be explicitly considered in the following model.

3 Bayesian hierarchical VBGF modeling

Instead of the standard specification of the VBGF [5], an alternative parameterization [3] is followed to link the age and size of corals:

$$L(t) = L_\infty(1 - e^{-te^{c/L_\infty}}) \tag{1}$$

with

$$c = ln(k)L_\infty \tag{2}$$

where $L(t)$ is the individual length at age t, L_∞ is the asymptotic length and k is the rate at which growth approaches this asymptote [1]. Parameter c can be seen as the part of the length growth accountable for site-specific conditions. This parameterization allows to skip the interaction between coral asymptotic length and growth rate.

A suitable hierarchical model to link coral length and age is proposed among different alternatives [6]. For the i-th observation ($i = 1, \ldots, n_j$) in the j-th site, the coral length is modeled as $L_i \sim N(\mu_i, \tau^2)$, where τ^2 is the precision, with

$$\mu_i = L_{\infty_{h,j}}(1 - e^{-t_i e^{c_j/L_{\infty_{h,j}}}}) \qquad h = 1, \ldots, H \quad \text{and} \quad j = 1, \ldots, J \tag{3}$$

where H is the number of species, J is the number of sites and n_{jh} is the abundance of the h-th species in the j-th site. Since relevant environmental covariates are measured with different timing and reliability with respect to the coral dataset, the second stage of the hierarchy accounts for possible differences in the asymptotic length between sites by random effects, as follows

$$L_{\infty_{h,j}} \sim TN_{(0.1,\infty)}(0, 0.1). \tag{4}$$

The effects expressed in (4) allow different environmental site-specific features together with differences between the two species. Being L_∞ strictly positive, a left truncated normal distribution with large precision has been chosen [10] as the prior distribution for this parameter, while for the c parameter in (2) a vague prior distribution has been chosen. This parameter is assumed to capture the differential growth of species between sites and is specified as

$$c_j \sim N(0.001, 1.00e^{-3}). \tag{5}$$

The model implicitly includes environmental covariates such as the solar radiation, the sea surface temperature or the marine current through the c parameter.
Finally, to complete the Bayesian specification, $\tau \sim U(0, 10)$ has been chosen. Models are implemented by means of the JAGS software [11] via the `R2jags` package of `R` [12] providing a powerful and flexible tool for analyzing grouped data, such as repeated measures data and nested designs.

4 Results

The joint posterior distribution of model parameters is obtained using 72,000 iterations with 3 chains, discarding the first 12,000 of the burn-in phase of the algorithm. Chains were checked for convergence and reasonable mixing by graphical inspection of the trace plots and by means of the Gelman-Rubin convergence diagnostics [7]. Posterior distributions of parameters are summarized in Figure 2. To ease the interpretability of results, distributions are reported distinguishing among sites. Each site panel contains graphical representations of the posterior distributions for the random effect and the asympotic length for the two species.

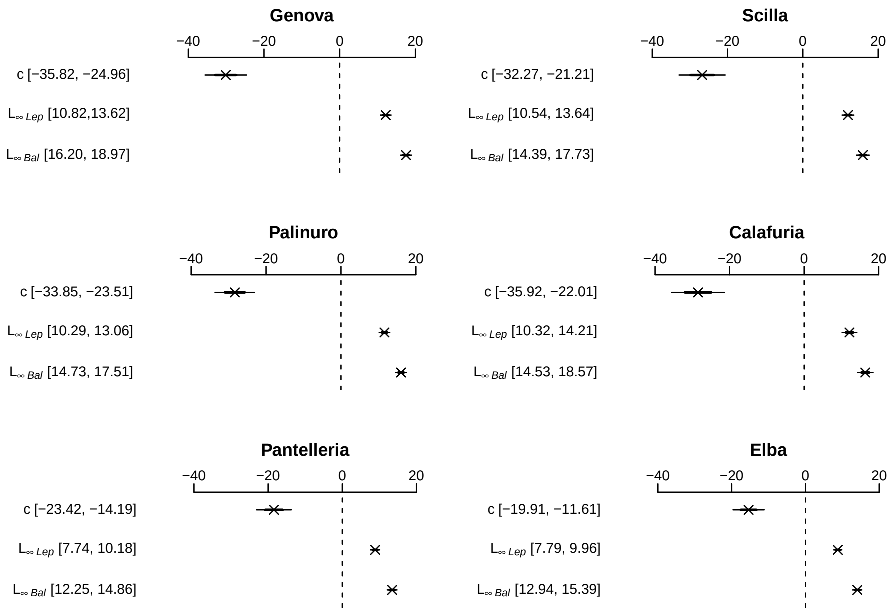


Fig. 2: Syntheses of posterior distributions of VBGF parameters. Crosses represent the means of the marginal posterior distributions; horizontal-bars represent 50% (thicker ones) and 95% (lighter ones) CIs (1 and 2 sd around the mean, respectively). For each parameter, 95% Credibility Intervals extremes are reported in brackets

The combination of species and locations is relevant in affecting L_∞ values. Results suggest a higher overall asymptotic length for *B. europea* than for *L. pruvoti*, confirming the well-known differences between the two species (Figure 1b). Posterior distributions of c parameters allow to distinguish among different sites, summarizing the relevant geographical effects on the growth rates of both species. The negative values of c are due to the fact that growth rates in (2) are less than 1 for this dataset. Sites that exhibit negative posterior c values closer to 0, tend to have lower L_∞ values with more concentrated posterior distributions. This result witnesses that very different environmental conditions influence coral growth at each site, which however translate into a direct relationship between a favorable situation and growth rate.

5 Conclusions

The examined case study provides a flexible framework that allows biologically meaningful estimates of the VBGF parameters. In particular, the use of Bayesian hierarchical models enables a better evaluation of uncertainty typical of length-at-age data and an innovative description of the relations between species growth and environmental conditions which can be represented by priors in a hierarchical model. The proposed approach demonstrates how to address ecological problems where environmental information is often limited, sparsely sampled or does not suit with collected data. By considering individual and population variation, the growth of two coral species has been estimated accurately and precisely, improving the understanding of their biological characteristics. Furthermore, this approach might be conveniently extended to multi-species models allowing tools to analyze entire communities.

Acknowledgements This work was supported by the PRIN2015 project EphaStat - *Environmental processes and human activities: capturing their interactions via statistical methods*, funded by the Italian Ministry of University and Research (MIUR).

References

1. Basso, N.G., Kehr, A.I.: Postmetamorphic growth and population structure ofthe frog Leptodactylus latinasus (anura: Leptodactylidae). Stud. Neotrop. FaunaEnviron **26**, 39–44. (1991)
2. Cadigan, N. G., Campana, Steven E.: Hierarchical model-based estimation of population growth curves for redfish (*Sebastes mentella* and *Sebastes fasciatus*) off the Eastern coast of Canada. ICES Journal of Marine Science **74** (3), 687 – 697 (2017)
3. Cafarelli, B., Calculli, C., Cocchi, D., Pignotti, E.: Hierarchical non-linear mixed-effects models for estimating growth parameters of western Mediterranean solitary coral populations. Ecological Modelling **346**, 1–9 (2017)
4. Caroselli, E., Zaccanti, F., Mattioli, G., Falini, G., Levy, O., Dubinsky, Z., Goffredo, S.: Growth and demography of the solitary scleractinian coral Leptopsammia pruvoti along a sea surface temperature gradient in the Mediterranean Sea. PLoS ONE **7**(6) , e37848 (2012)
5. Fabens, A.J.: Properties and fitting of the von Bertalanffy growth curve. Growth **29**, 265–289 (1965)
6. Gelman, A., Hill, J.: Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, NY (2007)
7. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B. , Vehtari, A., Rubin, D. B.: Bayesian Data Analysis, 3nd edition. Chapman & Hall/CRC, Boca Raton (2013)

8. Goffredo, S., Caroselli, E., Mattioli, G., Pignotti, E., Zaccanti, F.: Relationships between growth, population structure and sea surface temperature in the temperate solitary coral Balanophyllia europaea (scleractinia, dendrophylliidae). Coral Reefs **27**, 623–632 (2008)
9. Goffredo, S., E. Caroselli, E.Pignotti, G. Mattioli, F. Zaccanti: Variation in biometry and demography of solitary corals with environmental factors in the Mediterranean Sea. Mar. Biol. , **152**, 351–361 (2007)
10. Quintero, F. O. L., Contreras-Reyes, J. E., Wiff, R., Arellano-Valle, R. B.: Flexible Bayesian analysis of the von Bertalanffy growth function with the use of a log-skew-*t* distribution. Fish. Bull., **115**(1), 13–26 (2017)
11. Plummer, M.: JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003); Vienna, Austria. 124 (2003)
12. Su, Y., Yajima, M.: R2jags: Using R to Run "JAGS". R package version 0.5-7.(2015) https://CRAN.R-project.org/package=R2jags