

Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

**METODOLOGIA STATISTICA PER LA RICERCA
SCIENTIFICA**

Ciclo XXIV

SETTORE CONCORSUALE DI AFFERENZA: 13/D3- DEMOGRAFIA E STATISTICA SOCIALE

SETTORE SCIENTIFICO-DISCIPLINARE:SECS-S/04 – DEMOGRAFIA

TITOLO TESI

**Parameter estimation in a growth model for a
biological population**

Presentata da: **Elettra Pignotti**

Coordinatore Dottorato

PROF. ANGELA MONTANARI

Relatore

PROF. DANIELA COCCHI

Correlatore

PROF. BARBARA CAFARELLI

Esame finale anno 2013

Introduction	4
1 The growth curves in Biology	7
1.1 Growth Curve Model	7
1.1.1 Bell-shaped curve	8
1.1.2 Unbounded growth	9
1.1.3 Bounded growth	9
2 The use of the von Bertalanffy curve to estimate the solitary coral growth	13
2.1 Coral growth and problems of its modelling	13
2.2 The von Bertalanffy growth function for solitary corals	15
2.3 The methods used in marine biology for estimating the von Bertalanffy growth parameters	18
2.3.1 The Gulland-and-Holt (GH) plot	18
2.3.2 The size-increment method proposed by Fabens	19
2.3.3 The linearization proposed by Basso and Kehr	19
2.4 Gallucci and Quinn parameterization	20
2.5 A new proposal of parameterization based on Gallucci and Quinn model	20
3 An alternative approach to estimate the growth curves: hierarchical models	22
3.1 The hierarchical approach	22
3.2 Basic nonlinear regression model	24
3.3 Newton's method for nonlinear function estimation	26
3.4 The hierarchical model specification proposed by Lindstrom and Bates	27
3.4.1 Intra-site variation	28
3.4.2 Inter-site variation	29
3.5 Hierarchical nonlinear model for the solitary corals	30
3.5.1 The standard parameterization	30
3.5.2 The new proposed parameterization	33
4 Growth curves for corals	35
4.1 The Solitary Corals of our study	35
4.1.1 <i>Balanophyllia europaea</i> and <i>Leptopsammia pruvoti</i>	35
4.2 The Data	37

4.3 Explorative analysis	38
4.3.1 Balanophyllia europaea	38
4.3.2 Leptopsammia pruvoti	40
5 Results with the traditional Methods used in marine biology for the	
Von Bertalanffy Growth Function fitting	42
5.1 The Gulland-and-Holt (GH) plot	42
5.2 The size-increment method proposed by Fabens	47
5.3 The linearization proposed by Basso and Kehr	51
5.4 Newton's method for nonlinear function estimation	57
5.4.1 The standard parameterization	57
5.4.2 The new parameterization	59
5.5 Results with nonlinear hierarchical model	62
5.5.1 The method giving the best fit for the Balanophyllia Europaea	64
5.5.2 The method giving the best fit for Leptopsammia Pruvoti	71
6 Conclusions	77
Bibliography	79

Introduction

Marine organism growth has been the focus of increasing interest, over recent years, among marine biologists (Stolarski et al., 2007), (Goffredo et al., 2000). The goal of this work is therefore the problem of estimating the growth of some very common marine organisms in the Mediterranean sea, belonging to the family of corals, better known as solitary corals. Little is known about their growth mechanism and the influence of environmental factors. Furthermore the coral reproductive ability depends on how fast they reach the minimum size for reproduction.

The individual age of these corals may be estimated through their body size: standard age-based growth and population dynamics models may be subsequently applied. Demographic parameters highlight relationships between organisms and their environment, and contribute to the assessment of habitat stability; in addition, information on population turnover may contribute to techniques for the restoration of damaged or degraded coastal areas.

Due to difficulties and costs of recording the ages of corals, it is necessary to build up reliable growth models to infer coral age from their body size. In fact corals have a truncated cone geometry and their growth is isometric. Since growth depends on annual rings of calcium carbonate in coral bodies, the coral body size is directly related to the coral age and the ring thickness, which is in turn related to genetic and environmental factors. As a consequence, coral growth can be described as a function of several parameters, having a biological meaning, linked to age by a non-linear relationship.

In biology, the estimates of such parameters are generally obtained by using different strategies of linearization (see for instance Gulland and Holt ,1959; Fabenas,1965; Basso and Kehr,1991).

Biologists follow two main strategies for collecting data useful to estimate these parameters and consequently the coral growth.

In the first one the sample consists of corals randomly chosen in the site under study; each coral is measured repeatedly at fixed intervals of times in situ; the results are the mm/months of growth of each coral; then the parameters describing individual growth are estimated by a suitable method. Finally, the parameters for the

population are subsequently estimated starting from the individual ones and a growth curve for the population is finally constructed (Harper, 1977; Grigg et al. ,1984; Ross et al, 1988; Goffredo. et al., 2000).

In the second approach, the sample consists of colony of corals. Corals of different sizes of each colony are random collected, measured and dated. Each coral is measured only once because collection and dating causes the destruction of the coral. The growth curve for each colony is constructed using the dimensions of the body size of corals at different ages. The parameters of the growth curve for the population are then estimated by the parameters of the growth curve for each colony (Goffredo et al., 2000; Epstein et al. 2001; Goffredo and Chadwick-Furman, 2003; Goffredo and Lasker, 2008).

Apart from the method used for collecting data, the estimate of the growth curve parameters is traditionally is traditionally obtained by linear models that simplify non-linear relationships. The influence of environmental factors is accounted for by their linear correlation coefficient with the parameters (McClanahan, 2009; Goffredo and Caroselli, 2010; Harrell Yee and Barron, 2009; Soto et al., 2011). For example, Goffredo and Caroselli (2010) assessed the influence of the sea surface temperature (SST) on coral growth by finding a linear correlation between SST and estimated growth parameters in colonies located in places having different SST.

The above approaches share two main drawbacks: first, linearization is introduced regardless the behaviour of observed data, without considering the presence of variability in the body size of corals of the same age. This could cause an error in the parameter estimates. Furthermore no connection is hypothesised between the parameter estimate error at the coral (or colony) level, the parameter estimate error at the population level and the error of the linear regression between the parameter and the environmental factors.

For these reasons we propose instead a non-linear mixed effect model with the aim to overcome all these limits related to the estimation of coral growth and to obtain more reliable parameter estimates of the population growth curve, in this case avoiding forced linearization.

In addition, a new parameterization of the coral growth curve is proposed in order to identify parameters that are sensitive to the environment from those depending only on the genetics of the coral.

The dissertation starts in chapter 1 with an outline of the basic growth curves chapter 2 is focused on the description of the curve used to describe coral growth, the reasons for choosing it and a summary of the

traditional methods used to estimate the parameters of the curve; in the chapter there is also the description of a new proposed parameterization for the curve. Chapter 3 describes the new approach using a hierarchical nonlinear model is performed, whereas chapter 4 describes the data sets and the models using both parameterizations of the growth curve. Chapter 5 shows the results of the traditional linearization methods used in biology to estimate parameters; furthermore an estimate of the parameters using a nonlinear regression is proposed. Chapter 5 presents the results for the nonlinear hierarchical model applied to both parameterizations with and without considering the influences of environmental covariates.

1 The growth curves in Biology

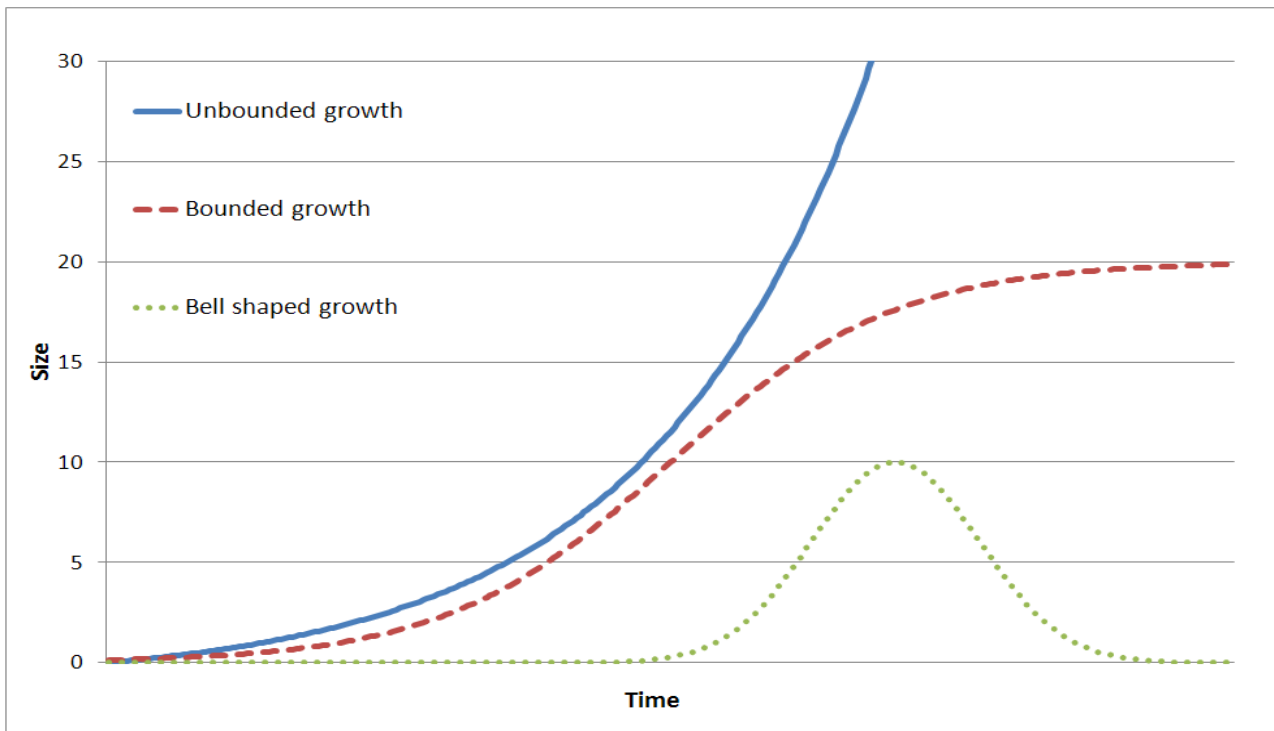
1.1 The Growth Curve Model

Growth curve models are used to describe how a particular quantity increases or evolves over time and consists of a sequence of data points taken at successive moments in time commonly spaced at uniform intervals. These models are also used to identify the type of growth pattern of different populations and to study the related variables in different fields of applications, for examples biology, ecology, demography, population dynamics, finance, econometrics, where a lot of models are developed including: mechanistic models, time series models, stochastic differential equations, etc.

In biology, growth is considered to be both related to the physical dimension and to the population of an organism; it is considered to be a fundamental property of biological systems studied at the colony level (group level), as well as for each organism (individual level). Growth curve models are based on longitudinal data, so measurements are repeatedly taken on a response variable at different time points.

Different kinds of growth patterns have been used in the literature to model the various types of realistic growth mechanisms. The first study pertaining to growth curves was presented by Wishart (1938) and differences between growth curves were discussed by Burnaby (1966). A wide and organic discussion of growth curve models (GCM) was introduced by Potthoff and Roy (1964) and subsequently expanded among others by Rao (1965). Results lead to summarizing growth curve shapes into three different growth mode categories: unbounded, bounded and bell-shaped curves (Figure 1.1). They behave similarly in early times because the energy available is all used for growth, then saturation starts to play a significant role. In fact saturation is due to an increasing part of energy used to maintain the reached dimension and this process can be very different among organisms. In bell-shaped curves, the increase eventually turns into decrease after passing through a peak point: bell-shaped curves can be symmetric or asymmetric. However, the curves displayed in Figure 1.1 have to be considered as examples, and there are a multitude of mathematical functions available for each type of growth mode (Höök et al., 2011).

Figure 1.1 general growth modes



1.1.1 Bell-shaped growth

Bell-shaped growth (Hubbert, 1956) includes a portion of time with a positive slope, an inflection point which becomes also a maximum and finally a portion of time with a negative slope. Bell-shaped curves have many different shapes and may be symmetric or asymmetric. They are often closely related to sigmoid functions and commonly appear as their derivatives (Höök et al., 2011). These curves have frequently been used in a wide array of disciplines, but initially were developed to describe and predict production growth in biological systems (cell production) or in bioenergetics systems (fuel production). Hubbert (1956) was among the first to formulate a basis for the extrapolation of finite resource production curves into the future and bell-shaped growth curves were an important cornerstone in such framework. He assumed that production levels begin at zero, before production has started, and end at zero, when the resource has been fully exhausted. In between, production would pass through one or several maxima. Conceptually, all resources are subjected to a physical limitation, due to the earth's intrinsic finiteness. Such physical resource limitations affect the general growth pattern. The limitation lies primarily in the maximum cumulative production that can be reached. So after an initial increase the production reaches a maximum, then slows down towards zero because of the finiteness of the resource. In biology they were developed to describe and

predict growth in biological systems such as in organs of the human body that show degeneration or involution: a typical example is the human brain which grows fast in children and slowly in adults; it reaches its maximum development in adult age and then shrinks slowly until old age when shrinking occurs rapidly.

1.1.2 Unbounded growth

All forms of unbounded growth, regardless of their mathematical nature, are clearly not suitable in the long-term. Linear growth is obviously slower than exponential growth but in the long run all unbounded growths tend towards infinity (Höök et al., 2011). Growth in economic theory is the most commonly used unbounded model; in fact, growth is believed to continue forever, but these models can be interesting only if they are studied over a “short” time (Bartlett, 1993 1999 and 2004) because models change according to different economic conditions. Unbounded growth is supported by those who believe that the obvious physical limitations of the earth do not necessarily imply an economic limitation (Simon, 1981); in fact conditions for the economic growth are continuously changing so when a source of economic growth is ending it is replaced by another one brought about by new technology or inventions or different behavioural patterns of humans. Others believe that that natural resources are not needed for economic growth (Solow, 1974) or even that human ingenuity can act as a powerful force capable of overcoming all possible physical limitations (Radetzki, 2007).

1.1.3 Bounded growth

This growth model is subject to physical limitations that affect growth rates, making growth slow down over time and asymptotically strive towards a maximum value (Janoschek, 1957; Beverton and Holt, 1957). Bounded growth occurs in many cases, such as human growth, organ growth, population growth and finite resource production. Growth may actually continue indefinitely but the rate of growth approaches zero as time tends to infinity. This is well known in many biological systems where an organism may grow quickly in its juvenile stage but then its growth slows down as it reaches maturity. The limiting factor lies rather in the growth process itself than in the amount of resources available for the growth. The upper limit is reached when all the available resources have been used to maintain the dimension achieved and there are no more

resources left for further growth. In the biological sciences, this condition is often modelled as proportionality between growth rate and actual size.

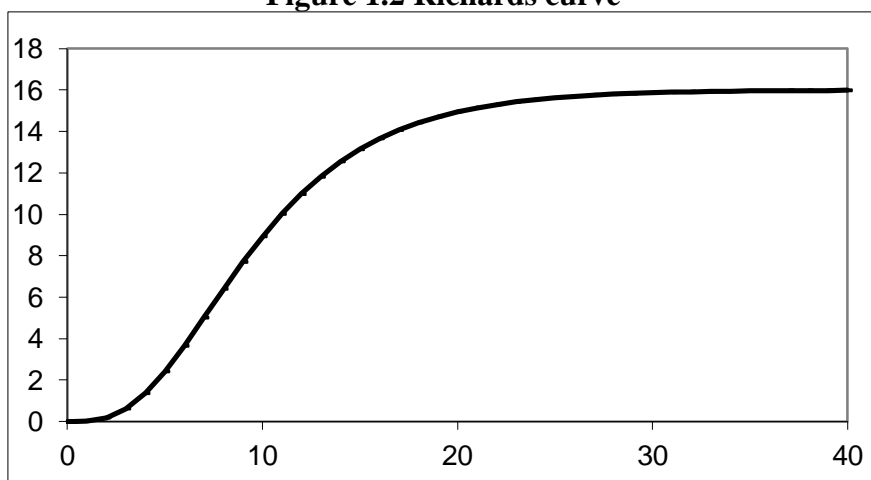
Sometimes these growth curves imply a rapid growth in the beginning, which later slows down; as in the case of exponential growth curves they are commonly used to estimate the length of some organisms, or the size of the skull and the brain.

Some curves have an inflection point (flex): the initial increase is slow therefore the first part of the curve appears flat; then growth increases rapidly so the curve turns upwards sharply until a second slow increase occurs, thus giving the curve a flat appearance again. An example can be found in sigmoid growth curves where growth is slow at the beginning of the time of observation and at the end but faster in the central time of observation.

The weight and volume of the body and of most organs show a sigmoid growth pattern initially, since the rate of growth in mass is low but increasing. The growth rate reaches a maximum, which corresponds to the inflection point in the curve, and then slowly declines to zero when the organism achieve their mature weight.

Richards (1959) proposed a sigmoid function developed as a generalization of classical growth curves and shown in Figure 1.2. This function is characterised by four parameters not very easy to interpret in many applications. Moreover it converges less frequently than the curves that we describe in what follows, especially if data do not already include the inflection point. It is actually used to fit data on forest growth.

Figure 1.2 Richards curve



The function is

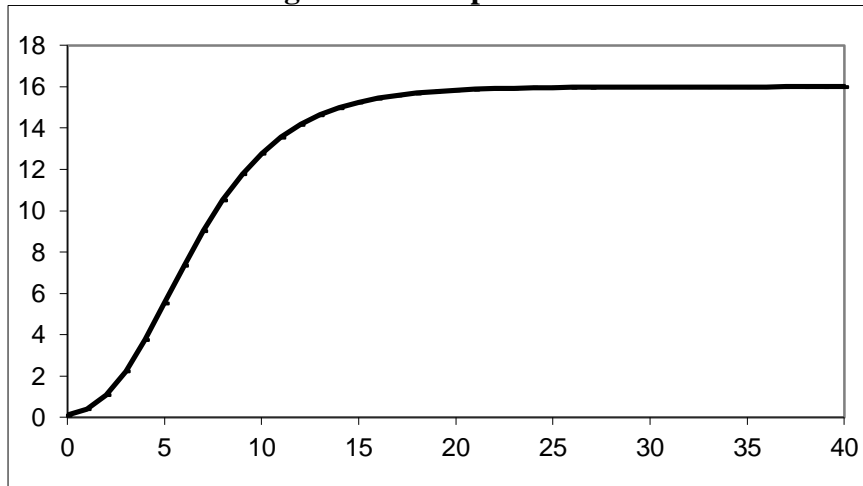
$$y(t) = L_{\infty}(1 - be^{-kt})^M \quad \mathbf{1.1}$$

where $y(t)$ is the dimension, L_∞ is the asymptotic maximum, b depends on the difference between the asymptotic maximum and the initial dimension $y(0)$, k is the growth rate and M shapes the way the asymptotic maximum is reached and the position of the inflection point.

Other well-known sigmoid curves are the Gompertz curve (Gompertz, 1825), the logistic curve (Verhulst, 1838) and the von Bertalanffy curve (von Bertalanffy, 1938).

The Gompertz curve (Gompertz, 1825) is shown in Figure 1.3, and was developed for the calculation of mortality rates. It is one of the most commonly used curves in growth mathematics and it is characterized by slow growth at the beginning and the end.

Figure 1.3 Gompertz curve



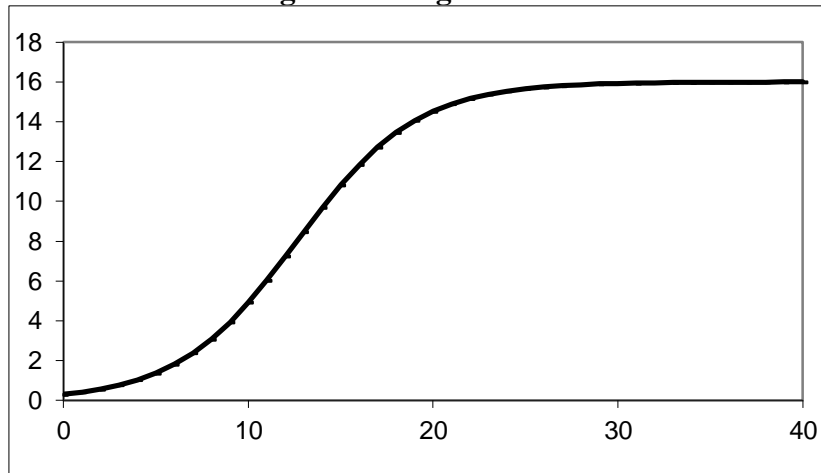
The function is

$$y(t) = L_\infty (e^{-be^{-kt}}) \tag{1.2}$$

where $y(t)$ is the dimension, L_∞ is the asymptotic maximum, b sets the displacement along the abscissa and k is the growth rate.

The Logistic curve, shown in Figure 1,4, was developed by Verhulst (1838) as a model for population growth. It is characterized by an inflection point independent of measurements so it is often used for sigmoid growth where the inflection corresponds approximately to one-half of final size.

Figure 1.4 Logistic curve



The function is

$$y(t) = \frac{L_{\infty}}{(1 + be^{-kt})} \quad \mathbf{1.3}$$

where $y(t)$ is the dimension, L_{∞} is the asymptotic maximum, b sets the displacement of the inflection point and k is the growth rate.

The von Bertalanffy curve is a bounded growth curve, derived from the von Bertalanffy Growth Law (1938) and based on the balance by the energy income provided by “food” and the fraction of this energy used for growth. The most common parameterisation of the von Bertalanffy Growth Law was described by Beverton and Holt (1957) and consists of a Richards model with the M shape parameter equal to 1. In this curve, dimension is a non-linear function of time based on 3 parameters, each with a physical meaning: maximum dimension, growth rate, and time at which the coral size is equal to zero. Their role and importance are discussed widely in the next chapter.

2. The use of the von Bertalanffy curve to estimate the solitary coral growth

2.1 Coral growth and problems of its modelling

Most simple marine organism populations are analysed considering individual body size. This measure is fundamental, since all physiological processes are related to size, including rates of metabolism, foraging and digestion (Peters, 1983; Calder, 1984). Body size is directly linked to life-history traits and individual competitive ability (Fox, 1975; Arendt, 2007). For these reasons, a reliable estimate of the growth parameters for a population of a certain area is important. In particular, the body size of corals is strictly related to the reproductive activity because size has to be big enough to let the planulae get out of the oral disk; where corals reaching the dimension of reproduction in early ages have more reproductive success.

Growth modelling and the analysis of intra-population patterns of body size variability over time are the central topics in animal population biology, since the internal size structure of populations may have a decisive influence on the population dynamics (De Angelis et al., 1993; Imsland et al., 1998; Uchmanski, 2000; Kendall and Fox, 2002; Fujiwara et al., 2004).

De Angelis et al. (1993) observed that model structure and parameter values are based on observations of individuals, so the natural variation among individuals and the stochastic nature of their fates should be incorporated in any model used. In this framework, the numbers of individuals necessary for estimation needs to be high. This model produces a reliable estimate of the population response to environmental variation only if data come from well-documented long-term studies of individuals belonging to the same populations. Imsland et al. (1998), Kendall and Fox (2002) and Fujiwara et al. (2004) underline that a model should incorporate population-specific data on ecological energetics, thermal and size dependence of digestive physiology and metabolic rates, energetics of individual growth, allometric relationships, social structure and mating system, and the dependence of mortality rates on age, size, and social status of individuals. This means that individual-level processes are determined by the organism's physiological state and interactions with its environment, whereas the population state is the distribution of individuals over all possible individual states.

In general, the von Bertalanffy growth function (VBGF, von Bertalanffy, 1938) is the most widely accepted relationship to describe the growth of fish and other marine organisms (Ricker, 1979; Cailliet et al., 2006). The VBGF describes the relationship between age and mean length of a population, whereas the variability among individuals of the same age (e.g. the variance or even the distribution of each cohort) is not included. Each individual is born with a specific genetic constitution that, to a certain extent, controls its growth profile (Sainsbury, 1980), but many physical and biological factors, such as water temperature (Sumpter, 1992), dissolved oxygen (Brett, 1979), photoperiod (Imsland et al., 2002), and the availability and type of food sources (Rilling and Houde, 1999) affect the actual growth rates achieved. In addition, the plasticity of phenotypes has been shown to be adaptive across environmental gradients (Conover and Munch, 2002; Ernande et al., 2004). Therefore, a suitable growth model should take into account individual and environmental variability.

In both ecological (Arino et al., 2004) and evolutionary (Conover and Munch, 2002; Ernande et al., 2004) contexts, one of the challenges of researchers is to model how the body size of an individual changes over time and to understand from the growth model what kind of probability distribution is suitable for the size of an individual as age increases (Lv and Pitchford, 2007; Fujiwara et al., 2004). In corals, as in other animals, the first source of variability is rooted in physiological processes and is the net result of two opposing processes, catabolism and anabolism (von Bertalanffy, 1938). The inter-individual variability in growth is the result of several internal (genetic) and external (environmental) factors, which affect these physiological processes. Pilling et al. (2002) and Clarke et al. (2011) state that the estimate of the growth curve of a population in each collection site is better performed by the growth model proposed for the size-at-collection data than a repeated measure model on a sample of individuals; in fact a single observation per individual better describes the mean growth parameters.

In fact, whilst each individual is born with a personal genetic architecture, which primarily determines its growth profile, a number of physical and biological factors, such as water temperature, solar radiation, the availability of appropriate food sources etc, have been shown to affect growth rates.

Differences in size among individuals that are established early in their life history can persist or be amplified (Ricker, 1958) if growth is positively size dependent and if there are positive correlations in

growth over time among individuals. The latter phenomenon is referred to as ‘growth autocorrelation’ (Pfister and Stevens, 2002) and is the reason for persistent growth differences among individuals. Similarly, differences among individuals in size can dampen over time (Ricker, 1958) if growth is negatively related or unrelated to size and there is no growth autocorrelation among individuals. The pattern of growth autocorrelation may be the result of several mechanisms, including factors that are intrinsic to the organisms, such as genetic or behavioural traits that confer performance differences among individuals (see Fraser et al. 2001). Alternatively, factors extrinsic to organisms, such as environmental heterogeneity, can cause persistent differences among individuals.

In conclusion, the underlying sources of growth variability in a population cannot generally be known. For size-at-collection data the consequences of not accounting for individual growth variability, or assuming the wrong source of variability, are less with respect to tag–recollection data, even when individual variability is high or data coverage is poor. So to reach a reliable size-at-age estimation the best way is to estimate VBGF parameters under a stochastic model using size-at-collection data.

2.2 The von Bertalanffy growth function for solitary corals

The main approach followed by biologists to obtain insights into metabolic phenomena is the study of the metabolism as a balance of energies: the energy entering the organism from food, heat, radiation and the energy necessary for feeding, growth, reproduction, maturation and maintenance. The mechanisms that are responsible for the organization of metabolism are not species specific (Kooijman, 2000). This care for generality is supported both by the universality of physics and evolution and the existence of widespread biological empirical patterns among organisms. In particular the growth of isomorphic organisms with abundant food is well described by the von Bertalanffy growth curve (Putter 1920; von Bertalanffy 1938). The identification of the curve is based on the physical principle that mass and energy are fixed quantities as starting points, so that the maintenance rate coefficient is the ratio between the cost of volume maintenance

E_M and the cost of growth E_G ($k_M = \frac{E_M}{E_G}$) under the biological principle that maintenance has priority over

growth and maturity maintenance has priority over maturation or reproduction. This leads to the biological

proposition that organisms of the same species have a maximum structural length L_{∞} . (Kleiber ,1947) All these considerations point to von Bertalanffy's law: the growth curve of an isomorphic juvenile or adult individual with constant food availability or abundant food is then

$$\frac{dL}{dt} = k(L_{\infty} - L) \quad 2.1$$

The von Bertalanffy growth rate k is given by

$$k = \left(\frac{3}{k_M} + \frac{3(L_h + L_{\infty})}{\nu} \right)^{-1} \quad 2.2$$

where L_h is the reduction in length due to the energy used for heating (L_h is the ratio between the cost of the surface maintenance E_S , associated mainly with heating, and the cost of volume maintenance E_M) and ν is the energy conductance (a measure of energy transfer efficiency). The set of parameter values is individual-specific. Individuals differ in parameter values and selection leads to evolution characterized by a change in the (mean) value of these parameters. However, it is important to underline that, in agreement with **Errore. L'origine riferimento non è stata trovata.**, L_{∞} and the growth parameter k are correlated; the von Bertalanffy growth rate decreases, in fact, with ultimate length: different combinations of k and L_{∞} can give almost the same fit to data, except when a wide range of ages is represented. Again, a high value of k combines with a low value of L_{∞} and vice versa (Sparre and Venema, 1992).

The von Bertalanffy's law (Putter, 1920; von Bertalanffy, 1938) is one of the most universal biological patterns (Fraser et al., 1990; Strum, 1991; Schwartz and Hundertmark, 1993; Ferreira and Russ, 1994; Ross et al., 1995) and can be considered the pillar of the laws describing the growth of organisms. Of course complex organisms have interactions between the different parts of their body and with the environment, so the description of growth only in terms of a physical law is very difficult. For simple organisms like corals this law well describes the body growth and can be used also for organisms of the same species with different food availabilities. In the latter case Kooijman et al. (2007) asserts that unlike **Errore. L'origine riferimento non è stata trovata.**, the logarithm of the von Bertalanffy growth rate decreases with ultimate length,

$$\ln(k) \propto \frac{1}{L_{\infty}} \quad 2.3$$

The most common parameterization of the solution to the differential equation 2.1 is

$$y(t) = L_{\infty}(1 - e^{-k(t-t_0)}) \quad 2.4$$

where $y(t)$ is the dimension, L_{∞} is the asymptotic maximum, k is the growth rate and t_0 is interpreted as the age when an individual would have been of zero length.

Biologists use the von Bertalanffy model with size at birth equal to 0 to describe solitary coral growth; when the size at birth is considered zero, thus $t_0=0$, the curve starts from the origin, and the

Errore. L'origine riferimento non è stata trovata. becomes

$$y(t) = L_{\infty}(1 - be^{-kt}) \quad 2.5$$

From a biological point of view the VBGF has three meaningful parameters:

1. L_0 is the mean length at birth ($t = 0$), which is species-specific and for solitary corals is universally considered to be very close to zero

$$L_0 = 0 \quad \text{or equivalently} \quad t_0 = 0$$

2. L_{∞} is the maximum mean length achievable by the species, with set environmental conditions and food availabilities, when t goes toward infinity.

3. k is the so-called Brody growth rate coefficient, but it is actually the exponential rate of approach to the asymptotic size, his unit is the reciprocal of the unit time (e.g. year⁻¹).

The main problem for biologists is then a reliable estimate of these parameters; the estimate should face two basic matters: the strong correlation between the parameters and the potential influence on them of environmental covariates which shouldn't affect, in agreement with the metabolic theory, the ultimate length L_{∞} .

For these reasons, we propose a new parameterization able to capture the effect of environmental covariates only by one parameter, isolating the role of the ultimate length L_{∞} and, in addition, to propose a method to obtain reliable estimates of the two parameters of the curve.

2.3 The methods used in marine biology for estimating the von Bertalanffy growth parameters

2.3.1 The Gulland-and-Holt (GH) plot

The GH plot (Gulland and Holt, 1959) is one of the most widely used methods in biology: it is based on annualized growth rates plotted vs. mean length at first and second measurements. Indeed the von Bertalanffy growth curve implies that the growth rate ($\frac{dL}{dt}$) declines linearly with length.

This relationship between length and growth rate can be used to estimate the two parameters L_∞ and k . In a standard GH plot, the growth rate $\frac{dL}{dt}$ of an experimental interval is plotted over the mean length in that interval). The differential form is

$$L'(t) = \frac{dL}{dt} = kL_\infty - kL_{mean} = k(L_\infty - L_{mean}) \quad 2.6$$

or in terms of growth increments per interval (length L_1 and L_2);

$$\frac{L_2 - L_1}{t_2 - t_1} = a + b \left(\frac{L_2 + L_1}{2} \right) \quad 2.7$$

$$k = -b \quad L_\infty = \frac{a}{-b} = \frac{a}{k}$$

For corals the mm/year growth is calculated for each individual and then plotted against the individual length: the least squares estimation of the straight line parameters are then used to calculate L_∞ and k .

This method has several limitations:

- it is adequate only if the time interval $\Delta t = t_2 - t_1$ is infinitesimal;
- it should be used only with follow up measurements, it is commonly used instead also with the length-at-capture data;
- it does not take into account the correlation between L_∞ and k ;
- it does not take into account the possible influence of environmental parameters;

- it is a deterministic method which does not take into account any possible statistical fluctuation and does not provide any confidence measure of the estimate.

2.3.2 The size-increment method proposed by Fabens

The size-increment method proposed by Fabens (1965), fits the von Bertalanffy model by the least squares method to observed lengths, using data on known growth increments in known time intervals but making no assumption about absolute age according to:

$$L_{t+\Delta t} = L_t e^{-k\Delta t} + L_\infty (1 - e^{-k\Delta t}) \quad 2.8$$

where Δt is the time increment between the two measured points L_t and $L_{t+\Delta t}$.

Indeed for corals the growth is measured for each individual at fixed intervals of time, so Δt is constant and then a linear regression of $L_{t+\Delta t}$ against L_t will generate a slope of $e^{-k\Delta t}$ and an intercept of $L_\infty(1 - e^{-k\Delta t})$.

This method has several limitations:

- it overestimates k and underestimates L_∞ . The bias appears to be associated with a failure to account for the redistribution of the error term when the basic growth equation is transformed to eliminate the necessity of estimating age;
- it does not take into account the correlation between L_∞ and k ;
- it does not take into account the possible influence of environmental parameters;
- it does not take into account the level and the distribution of the error at the individual and population of a colony level.

2.3.3 The linearization proposed by Basso and Kehr

The linearization proposed by Basso and Kehr (1991) fits the von Bertalanffy model imposing as L_∞ the largest size found in individuals of each species and then considering the linear regression between

$\ln\left(1 - \frac{L_t}{L_\infty}\right)$ as dependent variable and t as independent variable. So the estimate of k and $kt_0 = L_0$ is the

solution via the least squares method of the equation

$$\ln \left(1 - \frac{L_t}{L_\infty} \right) = -kt + kt_0 \quad 2.9$$

The above models are stated as deterministic, while linearization is valid only considering a small neighbourhood of the studied times, so nonlinear statistical methods can definitely improve the results.

2.4 Gallucci and Quinn parameterization

Gallucci and Quinn (1979) noted that comparisons of "growth" between two or several groups should involve both k and L_∞ . However, because of the generally high correlation between these two parameters, simultaneous hypothesis tests of these two parameters are non-standard and difficult to interpret. They introduced a new parameter, $\omega = kL_\infty$ asserting that it is a sensible index to compare two or more growth curves because it captures both the essential features of body-size growth. The new parameterization can be then achieved, by solving for L_∞ and substituting it In **Errore. L'origine riferimento non è stata trovata.**,

$$y(t) = \frac{\omega}{k} (1 - e^{-kt}) \quad 2.10$$

In the same article Galucci and Quinn (1979) state that ω can be thought of as a growth rate because the units are in length-per-time, which, in fact, corresponds to the instantaneous growth rate near $t=0$ in the case of corals. Furthermore, they claim that ω is the appropriate parameter to use to compare populations because of its statistical robustness owing to its normality (with small variance) for different populations of the same species (Moreau et al. 1987).

2.5 A new proposal of parameterization based on Galucci and Quinn model.

Kooijman (2000) underlines that for organisms of the same species with different food availabilities the logarithm of the von Bertalanffy growth rate decreases linearly with ultimate length as stated in **Errore. L'origine riferimento non è stata trovata.**

$$\ln(k) \propto \frac{1}{L_\infty}$$

So, different combinations of k and L_∞ can give almost the same fit to data, except when a wide range of ages is represented. Again, a high value of k combines with a low value of L_∞ and vice versa (Sparre and Venema, 1992)

Our proposal is then to use a different parameterization of

Errore. L'origine riferimento non è stata trovata.

$$y(t) = L_\infty \left(1 - e^{-te^{\frac{c}{L_\infty}}} \right) \quad 2.11$$

where

$$c = \ln(k) L_\infty \quad 2.12$$

This new approach might lead to more reliable results when including the effect of environmental covariates, in fact, as Kooijman (2008) pointed out, for simple isomorphic organisms with different food availabilities L_∞ could be considered independent of external factors; so the influence of covariates x_1, \dots, x_v representing environmental influences could be attributed to the new parameter

$$c = f(x_1, \dots, x_v) \quad 2.13$$

The new parameter c could be seen then as the part of the length growth accountable to site-specific conditions such as environmental factors.

This parameterization, compared to the traditional one, has the advantage of isolating the parameter sensible to environmental influences, so that it is possible to obtain a more meaningful and parsimonious statistical model when covariates are involved. The deterministic methods used by biologists do not suit with this parameterization because they were designed to find k and L_∞ , while the new parameterization does not contain k .

3 An alternative approach to estimate the growth curves: hierarchical models

3.1 The hierarchical approach

The models used by biologists are deterministic in the sense that they do not take into account differences among individuals of the same population and differences among populations related to different environmental conditions. They furthermore introduce a forced linearization which produces an unreliable estimation of the VBGF parameters; in fact the VBGF is a nonlinear relationship. A first step to improve the parameters estimation is the use of a nonlinear regression model so that it is possible to consider directly in the model the functional form linking growth and age. The nonlinear models are based therefore on assumptions often violated in the corals growth; in addition the studied data have two sources of variation: the coral population and the sites of collection. It comes natural then to get to nonlinear hierarchical models. Therefore a better approach could be a non-deterministic model in which each coral is characterized by different parameter values of the VBGF: in this way, each single coral has its own couple of parameters, (L_{∞}, k) , in the case of the classic VBGF parameterization, or (L_{∞}, c) , in the case of the new proposed VBGF parameterization, which is retained throughout its life (Sainsbury, 1980). According to biologists, the parameters values of corals collected in the same site, leaving in an environment characterised by the same sea surface temperature, sea current and solar radiation, should be more alike than those of corals collected in different sites. The parameters describing the growth of a coral can then be seen as a sum of different contributions: the species contribution common to all corals, the site contribution common to all corals exposed to the same environmental characteristic and a random contribution typical of the single coral attributable to immeasurable aspects. This approach points directly to a hierarchical model. Another critical aspect of the deterministic approach is the forced linearization; the statistical approach provides techniques to estimate the parameters while maintaining the nonlinear relationship. Combining the two aspects we come to the hierarchical nonlinear models: the coral growth curve can be then estimated via the nonlinear least squares method; where the species contribution, common to all the corals, is considered as a fixed effect; the site contribution related to all corals exposed to the same environmental characteristics is

considered as a random effect, which might depend on environmental covariates; the random contribution typical of the single coral can be seen as the residual error.

In fact, hierarchical nonlinear models for data in the form of continuous, repeated measurements on each of a number of individuals, are a popular platform for analysis when interest focuses on individual-specific characteristics and has gained broad acceptance as a suitable framework for such problems. The central concept of hierarchical models is that certain model parameters are themselves modelled; in other words, not all the parameters are directly estimated from the data, rather some of them are calculated from estimates of the model's hyperparameters which are in turn estimated from the data. The latter parameters are sometimes referred to as "random effects". They are to be distinguished from "fixed effects," which are not modelled, but are instead estimated directly from the data. A hierarchical model can have both, so it is often described as a "Mixed effects model".

Hierarchical nonlinear models may be regarded as both an extension of the nonlinear regression models and the hierarchical linear models. A natural framework is the two-stage model that takes into consideration intra- and inter-individual variations, as in mixed effect models. Hierarchical nonlinear models can be considered mixed effect models where some, or all, of the fixed and random effects occur non-linearly in the model function. From a non-linear point of view they can be seen as nonlinear regression models for independent data (Bates and Watts, 1988) where random effects are incorporated in the coefficients to allow them to vary by group, thus inducing correlation within the groups. From a mixed effect point of view they can be seen as linear mixed effect models where the conditional expectation of the response, given the random effects, is allowed to be a nonlinear function of the coefficients.

Pinheiro and Bates (2000) discuss three important advantages of nonlinear hierarchical models:

- interpretability. The modelling approach requires that one explicitly model. Judgment as well as background empirical or theoretical knowledge can be used to guide the choice of nonlinear functional form;
- parsimony. A well-chosen nonlinear function can model a non-linear process with fewer parameters than a linear model with multiple polynomial terms. In addition, the hierarchical modelling approach

allows one to replace a potentially large number of subject-specific indicator variables and interaction terms with a small number of hyperparameters;

- validity beyond the observed range of the data. Of course it is always dangerous to use a model to extrapolate beyond the data. However, this approach at least offers a framework within which one can harness one's background knowledge when specifying a model. Such an approach is less likely to lead one astray than a less parsimonious or more atheoretical "curve-fitting" approach.

3.2 Basic nonlinear regression model

The nonlinear regression models are the starting point to improve the estimation of the parameter values of the VBGF. For longitudinal data the used method are mostly based on least-square, maximum-likelihood and Bayesian estimation procedures as can be seen in Gallant (1987), Seber and Wild (1989), Genning et.al (1989), Davidian and Giltinan (1995), Vonesh and Chinchilli (1997).

Let y_j be the generic observation of the i th of M sites, where $j=1, \dots, n_i$, the model can be written then as

$$y_j = f(x_j, \boldsymbol{\phi}) + \varepsilon_j \quad 3.1$$

where $f(\cdot)$ is a nonlinear function, $\boldsymbol{\phi}$ is the $(p \times 1)$ vector considering the p parameters of the functions and ε_j is the random error. The function $f(\cdot)$ should satisfy $\mu_j = f(x_j, \boldsymbol{\phi})$ for one value of $\boldsymbol{\phi}$ in Φ and for all values of x . This condition, if verified, implies that $E(\varepsilon_j) = 0$ and consequently $E(y_j) = \mu_j$. The error should in addition satisfy the classical assumptions so that ε_j should be independent and identically distributed with zero mean and common variance σ^2 .

In the biological area, for growth curve or repeated measures this assumption could be often unrealistic.

The model formulation proposed by Davidian and Giltinan (1995) allows departures from the assumptions to be accommodate through some generalizations.

A general way to consider the intra-individual variance heterogeneity consists of specifying the variance function $g(\cdot)$ which may depend on the mean response, on constants z_j which may include the influence of environmental site-specific covariates and on an additional q -dimensional parameter vector $\boldsymbol{\theta}$ fully specifying the variance functional form.

$$\begin{aligned} \text{Var}(y_j) &= \sigma^2 g^2(\mu_j, z_j, \boldsymbol{\vartheta}) \\ \mu_j &= E(y_j) = f(x_j, \boldsymbol{\phi}) \end{aligned} \quad 3.2$$

Furthermore because of the repeated measures, errors might be correlated, that could be accommodate delineating the correlation of the error by the correlation matrix $\Gamma(\boldsymbol{\alpha})$ where $\boldsymbol{\alpha}$ is a s -dimensional vector of correlation parameters

Furthermore, for repeated measures, errors might be correlated. This intra-site correlation could be considered by using the correlation matrix $\Gamma(\boldsymbol{\alpha})$ where $\boldsymbol{\alpha}$ is a s -dimensional vector of correlation parameters (Davidian and Giltinan, 1995).

Moreover in growth curves both correlation among measurements and intra-site heterogeneity may be evident; if this happens, variance function $g(\cdot)$ could be used to define the diagonal variance matrix

$$\mathbf{G}(\boldsymbol{\phi}, \boldsymbol{\vartheta}) = \text{diag} \left[g^2(\mu_1, z_1, \boldsymbol{\vartheta}), \dots, g^2(\mu_n, z_n, \boldsymbol{\vartheta}) \right] \quad 3.3$$

with $\mathbf{G}^{1/2}(\boldsymbol{\phi}, \boldsymbol{\vartheta})$ the diagonal matrix with elements the square root of those of $\mathbf{G}(\boldsymbol{\phi}, \boldsymbol{\vartheta})$ (Davidian and Giltinan, 1995). Here $\boldsymbol{\phi}$ appears as an explicit argument to emphasize the possible dependence of intra-individual variance on the regression parameters through the mean response $\mu_j = f(x_j, \boldsymbol{\phi})$.

Considering a correlation pattern described by the matrix $\Gamma(\boldsymbol{\alpha})$ then the specification

$$\text{Cov}(\boldsymbol{\varepsilon}) = \sigma^2 \mathbf{G}^{1/2}(\boldsymbol{\phi}, \boldsymbol{\vartheta}) \Gamma(\boldsymbol{\alpha}) \mathbf{G}^{1/2}(\boldsymbol{\phi}, \boldsymbol{\vartheta}) = R(\boldsymbol{\phi}, \boldsymbol{\xi}) \quad 3.4$$

where $\boldsymbol{\xi} = [\sigma, \boldsymbol{\vartheta}', \boldsymbol{\alpha}']$ is the $(q+s+1)$ -dimensional combined vector of all intra-site covariance parameters

The **Errore. L'origine riferimento non è stata trovata.** implies that

$$\text{Var}(y_j) = \sigma^2 g^2(\mu_j, z_j, \boldsymbol{\vartheta}), \quad \text{Corr}(y_{j_1}, y_{j_2}) = \Gamma_{(j_1, j_2)}(\boldsymbol{\alpha}) \quad 3.5$$

These latter considerations improve the model but still doesn't fulfil the exigency to have a model which takes into account the different sites of collection and maybe the influence of environmental covariates points: this can be done in a more complete model like hierarchical nonlinear one.

3.3 Newton's method for nonlinear function estimation

Let y_j denote the response obtained at the j th covariate value t_j where $j=1,\dots,n$. The response vector $\mathbf{y} = [y_1, \dots, y_n]'$ contains the information at values $\mathbf{x} = [x_1, \dots, x_n]'$ so that

$$\begin{aligned}y_j &= f(x_j, \boldsymbol{\phi}) + \varepsilon_j \\ \boldsymbol{\varepsilon} &= [\varepsilon_1, \dots, \varepsilon_n]'\end{aligned}$$

where $f(\cdot)$ is a nonlinear function, $\boldsymbol{\phi}$ is the $(p \times 1)$ vector considering the p parameters of the functions. By the Least squares method, the parameter estimates provide the best fit of the mean function $y_j = f(x_j, \boldsymbol{\phi})$ to the observations obtained by minimisation of the residual sums of squares (RSS) with respect to $\boldsymbol{\phi}$ as follows:

$$\text{RSS}(\boldsymbol{\phi}) = \sum_{j=1}^n (y_j - f(x_j, \boldsymbol{\phi}) + \varepsilon_j)^2 \quad \mathbf{3.6}$$

The minimisation of the RSS is known as least-squares estimation, and the solution is the least-squares parameter estimates, denoted by $\hat{\boldsymbol{\phi}}$. The minimisation of the RSS is a nonlinear problem due to the nonlinearity of $f(x_j, \boldsymbol{\phi})$, and therefore numerical optimisation methods are needed. These methods start from some initial values and then repeatedly calculate next available value according to some optimization rules so that the iterative procedures will ideally approach the optimal parameter values in a stepwise manner. At each step, the proposed algorithm computes the new parameter values based on the data, the model, and the current parameter values. By far the most popular algorithm for estimation in nonlinear regression is the Gauss-Newton method, which relies on linear approximations to the nonlinear mean function at each step; unfortunately, two main complications arise when using it: how to choose the initial/starting parameter value and how to ensure that the procedure reached the global minimum rather than a local minimum. These two issues are interrelated. If the initial parameter values are sufficiently close to the optimal parameter values, then the procedure will usually reach the optimal parameter value (the algorithm is said to converge) within a few steps. Therefore, it is very important to provide sensible starting

parameter values. Poorly chosen starting values on the other hand will not reach convergence. If lack of convergence persists regardless the choice of the starting values, the conclusion is that the model is not appropriate for the data at hand. As the solutions to nonlinear regression problems are numeric, they may differ as a consequence of different algorithms, different implementations of the same algorithm (for example, different criteria for declaring convergence or computing first derivatives numerically or analytically), different parameterisations, or different starting values. However, the final parameter estimates ought not differ much. If there are large discrepancies, they might indicate that a simpler model should be preferred. Once the parameter estimates $\hat{\phi}$ are found, the estimate of the residual variance σ^2 is obtained as the minimum value of RSS (attained when parameter estimates are inserted divided by the degrees of

freedom $(n-p)$, as $s^2 = \frac{RSS(\hat{\phi})}{n-p}$ (Fox et al. 2002). The residual standard error is then s .

3.4 The hierarchical model specification proposed by Lindstrom and Bates

According to Lindstrom and Bates (1990) a general nonlinear mixed effects model for repeated measures can be defined at two levels. At the first step the j th observation on the i th site is modelled as

$$y_{ij} = f(\mathbf{x}_{ij}, \phi_i) + \varepsilon_{ij} \quad i = 1, \dots, M \text{ and } j = 1, \dots, n_i \quad 3.7$$

where y_{ij} is the j th response on the i th individual, \mathbf{x}_{ij} is the covariate vector for the j th response on the i th site and ϕ_i is the M -dimensional parameter vector, f is a nonlinear function and ε_{ij} is a normally distributed error term (Lindstrom and Bates, 1990). In the second step, the parameter vector ϕ_i , is modelled as

$$\phi_i = \mathbf{A}_i \boldsymbol{\beta} + \mathbf{B}_i \mathbf{b}_i \text{ and } \mathbf{b}_i \sim \mathbf{N}(0, \sigma^2 \mathbf{D}) \quad 3.8$$

where $\boldsymbol{\beta}$ is a p -dimensional vector of fixed effects, \mathbf{b}_i is a q -dimensional random effects vector associated with the i th individual, the matrices \mathbf{A}_i and \mathbf{B}_i are, respectively, the design matrices for the fixed and random effects and $\sigma^2 \mathbf{D}$ is a general variance–covariance matrix.

This formulation assumes the observations, corresponding to different groups, as independent and the within-group errors ε_{ij} as *i.i.d.* $N(0, \sigma^2)$ and independent on the \mathbf{b}_i .

The assumption of independence and homoschedasticity for the within-group errors can be combined therefore in a more general model.

3.4.1 Intra-site variation

The model **Errore. L'origine riferimento non è stata trovata.** describes the systematic and random variation associated with measurements on the i th site. In particular the systematic variation is taken into account by the regression function f , random variation is taken into account by a distributional assumption for the random errors e_{ij} and the specification of a model for its distribution. As already seen, for a given site variability in the y_{ij} may be a systematic function of the mean response for that site, other known constants and additional, possibly unknown parameters; correlation among measurements on a given site may also arise. In many contests, and growth curve should be one of those, it is reasonable to expect a comparable pattern of intra-site variation across sites (Davidian and Giltinan,1995). The pattern of correlation of measurement taken in a given site would also be likely to remain constant across sites.

According to **Errore. L'origine riferimento non è stata trovata.** and collecting the errors for the 1th site into the vector $\boldsymbol{\varepsilon}_i = [\varepsilon_{i1}, \dots, \varepsilon_{in_i}]'$ it is possible to write a general specification of the common intra-site variance structure as

$$Cov(\boldsymbol{\varepsilon}_i / \boldsymbol{\phi}_i) = R_i(\boldsymbol{\phi}_i, \boldsymbol{\xi}), \quad \boldsymbol{\xi} = [\sigma, \boldsymbol{\vartheta}', \boldsymbol{\alpha}']' \quad 3.9$$

allowing for variance heterogeneity and correlation within sites. The most common assumption about the conditional distribution of the error for a given $\boldsymbol{\phi}_i$ is that of intra-site normality of the response which comes from the error specification (Davidian and Giltinan, 1995)

$$\boldsymbol{\varepsilon}_i / \boldsymbol{\phi}_i \sim N(\mathbf{0}, R_i(\boldsymbol{\phi}_i, \boldsymbol{\xi})) \quad 3.10$$

3.4.2 Inter-site variation

Variation among different sites is taken into account by the site specific regression parameters ϕ_i . The standard approach as already seen is to specify a model for the ϕ_i which could consider part of the parameters variation due to a systematic dependence on individual characteristic (possibly covariates) and part due to unexplained (random) reasons. To account for these possibilities the parameters ϕ_i are considered as depending on systematic and random components respectively β and \mathbf{b}_i .

Let ϕ_i be a p -dimensional vector of regression parameters specific to the i th individual and \mathbf{a}_i be an a -dimensional covariate vector corresponding to the attribute of the i th individual. If \mathbf{b}_i is a k -dimensional vector of random effects associated with the i th individual and β is a $(r \times 1)$ vector of fixed effects then a general model for ϕ_i could be given by

$$\phi_i = d(\mathbf{a}_i, \beta, \mathbf{b}_i) \tag{3.11}$$

where d is a p -dimensional vector-valued function. Each element of d is associated with the corresponding element of ϕ_i , so that the functional relationship may be of a different form for each element. A complete characterization of the inter-site variation requires an assumption about the distribution of the random effects \mathbf{b}_i . The most common distributional assumption is

$$\mathbf{b}_i : N(\mathbf{0}, \sigma^2 \mathbf{D}) \tag{3.12}$$

where $\sigma^2 \mathbf{D}$ is a $(k \times k)$ covariance matrix (Davidian and Giltinan, 1995)

As an alternative to normality it is possible to assume a multivariate t (Wakefield 1995) or a mixture of normal distributions (Beal and Sheiner 1992). The t distribution with its heavier tails may provide a robust alternative to handle outlying individuals; the mixture of normals accommodates the possibility of multimodality of the distribution of the random effects. It is also possible to consider situations in which the assumption of zero mean is not made.

3.5 Hierarchical Nonlinear model for the solitary corals

In the following paragraphs the hierarchical model to design the solitary corals growth curve based on VBGF will be specified.

3.5.1 The standard parameterization

Let y_{ij} be the length of j th coral in the i th site; let t_{ij} be the age of corals and let ϕ_i be a bi -dimensional vector

$\phi_i = [L_{\infty_i}, k_i]^T$ and considering then M sites each one measured n_i times, the nonlinear model can be

integrated and the **Errore. L'origine riferimento non è stata trovata.** in accordance with the (3.6) becomes

$$y_{ij} = L_{\infty_i} (1 - e^{-k_i t_{ij}}) + \varepsilon_{ij} \quad j = 1, \dots, n_i \quad i = 1, \dots, M \quad \mathbf{3.13}$$

In designing the distribution of the error ε_{ij} we have to consider if the classical assumption are reliable for corals growth.

The assumption of zero mean of the error is not called into question as the relationship between the response and the covariate based on the VBGF has a physical meaning and is derived, as already seen, by biochemical considerations. The assumption that the error has common variance σ^2 and is identically distributed for all t_{ij} risks to be easily violated for two reasons: the first one is that young corals are less variable than old one as environmental factors have less time to influence them, the second one depends on the way they are measured. The age of solitary corals are determined by counting the growth rings, so in small corals that could be less precise than in adult corals, on the other hand in very old corals, slowing their growth after a certain dimension is very difficult to count ultimate rings because they are very close, sometimes collapsed. The error may be also correlated: the growth in a certain site depends on the yearly fluctuation of environmental parameters, such as temperature, solar radiation and current. So corals closer in age should have had the same fluctuation of environmental parameters.

A more flexible model for the variance of the error can be then build considering the function described in

Errore. L'origine riferimento non è stata trovata. and

Errore. L'origine riferimento non è stata trovata. according to which the error variance matrix becomes

$$R_i(\phi_i, \xi), \text{ where } \xi = [\sigma, \mathfrak{D}', \alpha']'$$

So that

$$\varepsilon_{ij} \sim N(\mathbf{0}, R_i(\phi_i, \xi))$$

At the inter-site level it is possible to attribute the variations of parameters $L_{\infty i}$ and k_i among the M sites to systematic and random sources.

Parameter estimates can be obtained by the method of Least Squares; however minimization of residual sum of squares yields equations nonlinear in the parameters. Since it is not possible to solve nonlinear equations in closed forms, the alternative is to obtain approximate analytic solutions by employing iterative procedures. The main methods being (Draper and Smith 1998): the Taylor Series Method, the Steepest Descend Method and the Levenberg-Marquardt's Gauss-Newton based Method.

It is then possible to see each parameter comprising a fixed effect (L_{∞} and k) due to known site-specific characteristics and a random effect respectively b_{1i} and b_{2i} due to unexplained variation among the sites. The vector ϕ_i considering both the effect can be then defined

$$\phi_i = \begin{pmatrix} L_{\infty i} \\ k_i \end{pmatrix} = \begin{pmatrix} L_{\infty} + b_{1i} \\ k + b_{2i} \end{pmatrix} \quad \mathbf{3.14}$$

Unlike the traditional marine biology approach, the **Errore. L'origine riferimento non è stata trovata.** approach to parameter estimations has the desirable feature of contemporarily taking into account the site model and the “population” model introducing a hierarchy between them.

Analysis of data of repeated measurement over time is a recurrent challenge to statisticians engaged in biological applications: growth curves are among this kind of data. The inter-individual variability in growth is the result of several internal (genetic) and external (environmental) factors which affect these physiological processes. The choice of applying the growth model to the length-at-collection data rather than following the growth process over time of several individuals allows the description through mean growth

parameters which oversee individual variability in each site of capture. The measures are then repeated in a wide sense. Anyway Pilling et al., (2002) and Schaalje et al., (2001) focused on the statistical nuances of fitting back-calculated lengths at age data in a repeated measures context and obtaining better estimates of individual growth variability than taking repeated measurements on a sample of individuals.

The main strategy followed was to incorporate these features in an inferential setting by building a hierarchical model. Inter-site variation is then considered as consisting of a model for variation in the regression parameters ϕ_i . Thus variation can be modelled using a distributional assumption for ϕ_i at various levels of complexity. For instance a possible specification is

$$\phi_i = \mathbf{A}_i \boldsymbol{\beta} + \mathbf{b}_i \quad 3.15$$

where ϕ_i is assumed to depend linearly on a two-dimensional vector of parameter $\boldsymbol{\beta}$ and on site-specific information such as temperature and radiation summarized in a design Matrix \mathbf{A}_i . as shown further on.

Error \mathbf{b}_i corresponds to the random component of inter-site variation, which is supposed not depending on environmental covariates and is taken to have mean zero and covariance matrix \mathbf{D} .

Adding the restriction that the distribution of ϕ_i belongs to a particular parametric family, the bivariate normal-lognormal distribution is often chosen (Helser and Lai, 2004)

$$\phi_i = \begin{pmatrix} L_{\infty i} \\ \ln k_i \end{pmatrix} \sim N(\mathbf{A}_i \boldsymbol{\beta}, \mathbf{D}) \quad 3.16$$

So, considering a linear influence of radiation gradient R and temperature gradient T on L_{∞} , and exponential influence of radiation gradient R and temperature gradient T on k in the parameterization of the VBGF, expression (3.14) is composed by the following elements

$$\mathbf{A}_i = \begin{pmatrix} 1 & R_i & T_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & R_i & T_i \end{pmatrix} \quad 3.17$$

$$\boldsymbol{\beta} = \begin{pmatrix} L_{\infty} \\ a_1 \\ a_2 \\ \ln k \\ a_3 \\ a_4 \end{pmatrix} \quad 3.18$$

$$\mathbf{b}_i = \begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} \quad 3.19$$

The elements of the bi-dimensional vector-valued function described in

Errore. L'origine riferimento non è stata trovata. are then expressed as

$$\begin{aligned} d_1(\boldsymbol{\alpha}_i, \boldsymbol{\beta}, \mathbf{b}_i) &= L_\infty + a_1 R_i + a_2 T_i + b_{1i} \\ d_2(\boldsymbol{\alpha}_i, \boldsymbol{\beta}, \mathbf{b}_i) &= \ln(k) + a_3 R_i + a_4 T_i + b_{2i} \end{aligned} \quad 3.20$$

where $\boldsymbol{\alpha}_i = \begin{pmatrix} R_i \\ T_i \end{pmatrix}$

3.5.2 The new proposed parameterization

Similarly to the previous chapter the above VBGF **Errore. L'origine riferimento non è stata trovata.** in accordance with the **Errore. L'origine riferimento non è stata trovata.** becomes

$$y_{ij}(t) = L_{\infty_i} \left(1 - e^{-t_{ij} \frac{c_i}{L_{\infty_i}}} \right) + \varepsilon_{ij} \quad j = 1, \dots, n_i \quad i = 1, \dots, M \quad 3.21$$

where the error distribution is $\varepsilon_{ij} \sim N(\mathbf{0}, R_i(\boldsymbol{\phi}_i, \boldsymbol{\xi}))$

The $\boldsymbol{\phi}_i$ is a bi-dimensional vector $\boldsymbol{\phi}_i = [L_{\infty_i}, c_i]'$ and also in those case it is possible to see each parameter comprising a fixed effect L_∞ and c and a random effect b_{1i} and b_{2i} . The vector $\boldsymbol{\phi}$ considering both the effect can be then defined

$$\boldsymbol{\phi}_i = \begin{pmatrix} L_{\infty_i} \\ c_i \end{pmatrix} = \begin{pmatrix} L_\infty + b_{1i} \\ c + b_{2i} \end{pmatrix} \quad 3.22$$

This new parameterization is designed to be more parsimonious: in fact the two parameters are divided in the one sensitive to environmental and, in general, external, influences that is c and the one derived by genetic and moreover, site not-dependending, characteristics, that is L_∞ .

We consider for the distribution of $\boldsymbol{\phi}_i$ the bivariate normal distribution

$$\boldsymbol{\phi}_i = \begin{pmatrix} L_{\infty_i} \\ c_i \end{pmatrix} \sim N(\mathbf{A}_i \boldsymbol{\beta}, \mathbf{D}) \quad 3.23$$

Considering a linear influence of radiation (R) and temperature (T) on c which is designed to be the only parameter sensitive to environmental covariates, the parameterization of the VBGF the **Errore. L'origine riferimento non è stata trovata.** is composed by the following elements

$$\mathbf{A}_i = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & R_i & T_i \end{pmatrix} \quad 3.24$$

$$\boldsymbol{\beta} = \begin{pmatrix} L_\infty \\ a_1 \\ a_2 \\ c \\ a_3 \\ a_4 \end{pmatrix} \quad 3.25$$

$$\mathbf{b}_i = \begin{pmatrix} b_{1i} \\ b_{2i} \end{pmatrix} \quad 3.26$$

The elements of the two-dimensional vector-valued function described in the **Errore. L'origine riferimento non è stata trovata.** become then

$$\begin{aligned} d_1(\boldsymbol{\alpha}_i, \boldsymbol{\beta}, \mathbf{b}_i) &= L_\infty + b_{1i} \\ d_2(\boldsymbol{\alpha}_i, \boldsymbol{\beta}, \mathbf{b}_i) &= c + a_3 R_i + a_4 T_i + b_{2i} \end{aligned} \quad 3.27$$

where $\boldsymbol{\alpha}_i = \begin{pmatrix} R_i \\ T_i \end{pmatrix}$

4 Growth curves for corals

4.1 The Solitary Corals of our study

For this study we considered two species of solitary corals: *Balanophyllia europaea* and *Leptopsammia pruvoti*, because these corals are very important in determining the health of other colonial corals of the area with regards to climatic changes and the presence of pollution. Moreover for these corals age-based models can be easily applied. In fact, age can be determined by counting the growth bands of the skeleton using the Computerized Tomography scans (CT) and the size can be easily measured by a calliper.

4.1.1 *Balanophyllia europaea* and *Leptopsammia pruvoti*

Balanophyllia europaea is a solitary, ahermatypic, zooxanthellate scleractinian coral that lives on a rocky substratum and is endemic to the Mediterranean Sea. Owing to its symbiosis with zooxanthellae, depth distribution appears to be restricted in this species; it is found between 0 m and a maximum of 50 m depth (Zibrowius 1980), though congeneric azooxanthellate corals have been reported at depths of up to 1,100 m (Cairns 1977). The reproductive biology of this species is characterized by simultaneous hermaphroditism and brooding. During the annual cycle of sexual reproduction, fertilization takes place from March to June and planulation in August and September.

Figure 4.1 *Balanophyllia europaea*



This coral has been chosen because Goffredo et al. (2007) observed that unlike other tropical and temperate corals, in which calcification is positively correlated with solar radiation and surface sea temperature (SST),

balanophyllia calcification is not correlated with solar radiation, whereas it is negatively correlated with SST. The conjecture is that photosynthesis of the symbiotic algae of *Balanophyllia europaea* is inhibited at high temperatures, consequently causing an inhibition of calcification so that the calcification of *Balanophyllia europaea* would be depressed at 20.5–21 °C mean annual SST.

Leptopsammia pruvoti is an ahermatypic, nonzooxanthellate and solitary scleractinian coral, which is distributed in the Mediterranean basin and along the European Atlantic coast from Portugal to Southern England and Ireland. It is one of the most common organisms in semi-enclosed rocky habitats, under overhangs, in caverns and small crevices at 0–70 m depth, with mean abundances of [104 individuals m⁻², i.e., [2 kg m⁻² of CaCO₃ biomass (Goffredo et al. 2007). When released, the planulae (695–1,595 µm in length) are ready to settle and swim by ciliary movement for 1–20 days.

Figure 4.2 *Leptopsammia pruvoti*



This coral was chosen because Goffredo et al. (2007) observed that unlike other tropical and temperate corals, in which calcification was positively correlated with solar radiation and SST, the biometry of individuals, thus skeletal density, corallite length, width and height, and their abundance are not affected by SST and solar radiation along an 850-km latitudinal gradient in western Italian coasts.

Solitary coral size growth and reproduction is usually related to the length of maximum diameter of the oral disk, so the study was based on the growth of the radial length according to the VBGF.

As observed by Goffredo (2008), the population structures of these two species of temperate solitary corals become less stable and have deviated from the steady state in recent years as a result of a progressive

deficiency of young individuals. The study of the variation in calcification rate, linear extension rate, and skeletal density in populations arranged along a temperature and solar radiation gradient can provide therefore an explanation for this observation. The results can be extended to the other temperate scleractinian corals and considered in light of the most recent scenarios on climate changes for the near future.

4.2 The data

The data were collected by S. Goffredo et al.(2007) e S. Goffredo et al. (2008). Specimens of different ages (from 1 to 14 years) of two species of solitary corals: *Balanophyllia Europaea* and *Leptopsammia Pruvoti* were collected at maximum biomass density depth in 6 different sites having different mean annual radiation, sea surface temperature and sea current; they were then dated and measured. Two sites were islands and four were on the coast. Table 4.1 shows the values of annual mean and standard error of sea surface temperature, solar radiation and sea current for the sites under study. The other variables considered in this analysis were:

- the length in *mm* of the corals measured by callipers
- the age in *years* measured by the mean over 3 repeated counts of the growth rings visualized by CT scan
- the mean annual radiation in (W/m^2) of the 6 sites (Italian Air Force Weather Service)
- the mean annual temperature in ($^{\circ}C$) of the 6 sites (Italian Air Force Weather Service)
- the mean annual surface current in (m/s) of the 6 sites (Atlas of the surface currents of the Italian Navy)

Table 4.1 Environmental Characteristics of each site

	Genova #1	Calafuria #2	Elba #3	Palinuro #4	Scilla #5	Pantelleria #6
Mean annual radiation (SE) (W/m^2)	166.95 (1.02)	170.07 (1.02)	172.74 (1.02)	181.48 (1.01)	187.31 (1.02)	190.95 (1.02)
Mean annual surface temperature (SE) ($^{\circ}C$)	19.56 (0.04)	18.02 (0.04)	18.74 (0.04)	19.14 (0.03)	19.54 (0.02)	19.88 (0.04)
Mean annual Current (m/s)	0.08	0.08	0.2	0.15	0.6	0.15
Island/Coast	Coast	Coast	Island	Coast	Coast	Island

4.3 Explorative analysis

4.3.1 *Balanophyllia europaea*

238 corals of *Balanophyllia europaea* were collected in the six sites considered. Table 4.2 shows the number of corals per site and the relative descriptive analysis for the age and the oral disk length. It can be argued that the different sites show different mean oral disk lengths; this difference is not always ascribable to a different mean age; in fact, corals from Calafuria are the smallest but not the youngest. Corals from Elba and Pantelleria are at least as young as those of Calafuria but they are definitely wider. Corals collected in Genova, on the other hand, are older and much wider. Figure 4.3 shows the oral disk length plotted against the relative age of each coral: The graphic shows the different pattern of the sites: corals from Elba, for example, are wider than the others only until the age of five years, corals from Genova are wider from the age of seven years, and corals from Palinuro are smaller at all ages. This may suggest different growth patterns in the different sites. Figure 4.4 shows the oral disk length plotted against the relative age of each coral in each site and confirms the idea of a different growth pattern according to the sites. Figure 4.5 shows the oral disk length distribution at different ages: the distributions of several ages are clearly not Gaussian, maybe because of the influence of the sites; the shape of the distribution of the oral disk length against the age is confirmed by a Von Bertalanffy-like growth curve.

Table 4.2 *Balanophyllia europaea*: descriptive analysis

	Sites	Number of corals	Mean	Std. Deviation	Median	95% Confidence Interval for Mean		Minimum	Maximum
Age	Genova	42	7.4	3.3	7.0	6.4	8.5	1.0	14.0
	Calafuria	34	5.5	1.9	5.0	4.9	6.2	2.0	9.0
	Elba	34	4.6	2.2	4.0	3.9	5.4	2.0	12.0
	Palinuro	54	6.9	3.3	6.5	6.0	7.8	1.0	14.0
	Scilla	32	6.2	2.6	6.5	5.3	7.2	1.0	13.0
	Pantelleria	42	5.2	2.3	5.0	4.5	6.0	1.0	9.0
	Total	238	6.1	2.8	6.0	5.7	6.5	1.0	14.0
Length	Genova	42	11.7	3.9	11.6	10.4	12.9	2.4	19.0
	Calafuria	34	8.3	2.7	10.0	7.4	9.3	4.0	14.5
	Elba	34	9.0	2.9	10.5	7.9	10.0	3.0	16.3
	Palinuro	54	9.9	3.2	10.6	9.0	10.8	2.5	16.9
	Scilla	32	9.9	2.9	11.0	8.8	10.9	3.7	16.0
	Pantelleria	42	8.8	2.8	9.7	8.0	9.7	2.6	13.3
	Total	238	9.7	3.3	10.5	9.2	10.1	2.4	19.0

Figure 4.3 The length of *Balanophyllia europaea* vs age

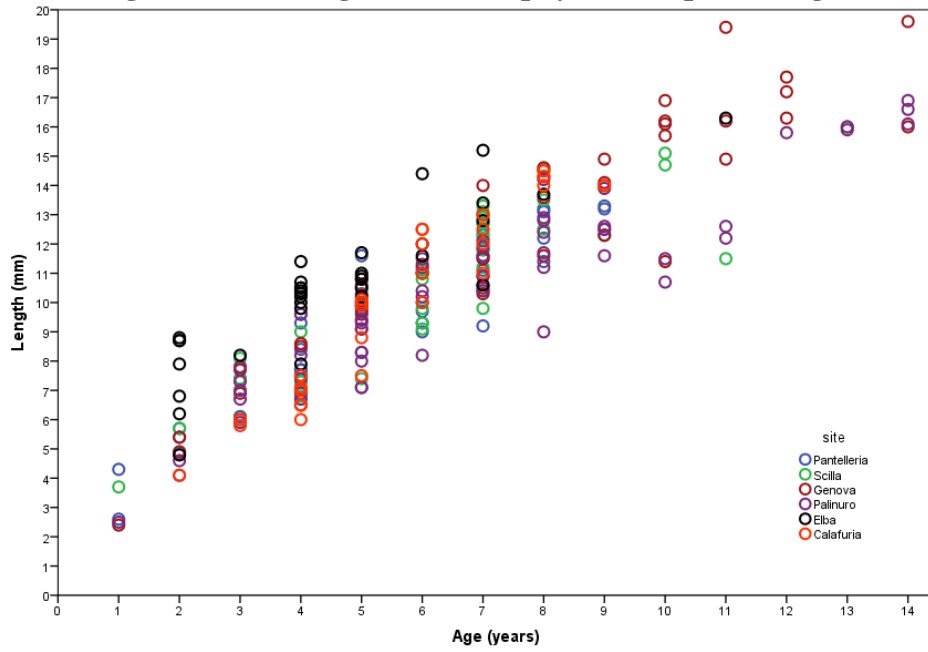


Figure 4.4 The length of *Balanophyllia europaea* vs age for each site under study

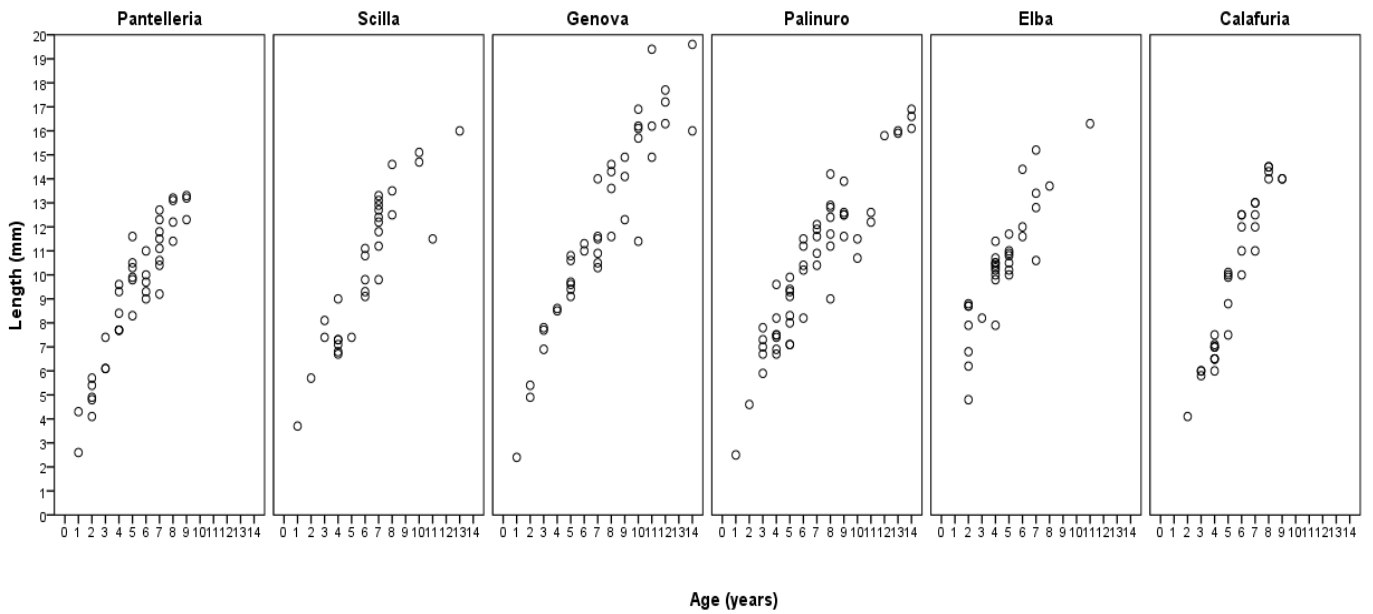
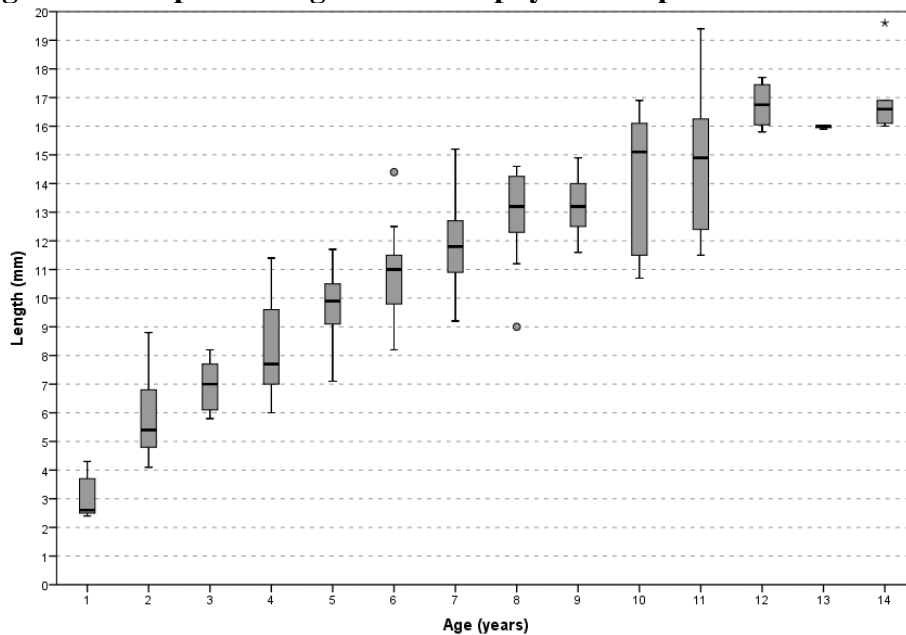


Figure 4.5 Boxplot of length of *Balanophyllia europaea* at different ages



4.3.2 *Leptopsammia pruvoti*

179 corals of *Leptopsammia pruvoti* were collected in the six sites considered. Table 4.3 shows the number of corals per site and the relative descriptive analysis for the age and the oral disk. As for *Balanophyllia europaea*, the different sites show different mean oral disk lengths not always ascribable to a different mean age: corals from Genova and Pantelleria have similar dimensions, but corals from Genova are younger, The same can be said of corals from Calafuria and Palinuro or corals from Elba and Scilla where corals from northern sites are younger than those of southern sites with similar dimensions. Figure 4.6 shows the oral disk length plotted against the relative age of each coral. The graphic shows a wide spread of the dimensions at older ages and furthermore, unlike *Balanophyllia europaea*, the corals from Elba are wider than the others only from the age of six years, where corals from Genova are smaller from the same ages; corals from Pantelleria are smaller at all ages. Therefore, for *Leptopsammia pruvoti* a different growth pattern can also be hypothesized in the different sites. Figure 4.7 shows the oral disk length plotted against the relative age of each coral in each site and confirms the idea of a different growth pattern according to site. In particular it can be noticed that corals from Calafuria are very young and fast growing, but due to the lack of older corals it will be difficult to understand if this fast growth will be maintained. Anyway, the difference in age distribution of the corals from Calafuria might be a problem when carrying out separate estimations according to site. . Figure 4.8 shows the oral disk length distribution at the different ages: as for the

Balanophyllia europaea the distributions of several ages are clearly not Gaussian. Furthermore, the oral disk length distribution vs age confirms a Von Bertalanffy-like growth curve only for ages starting from three years; Ages two and moreover one have bigger dimensions than those expected by a Von Bertalanffy-like growth curve. It can be noticed therefore that this coral is smaller than *Balanophyllia europaea* and may be subjected to a bigger error of measurement for very young or small corals; on the other hand the influence of the environment of the six studied sites might produce a different growth rate at early ages thus explaining the gap from the theoretical behaviour.

Table 4.3 *Leptopsammia pruvoti*: descriptive analysis

	Sites	N	Mean	Std. Deviation	Median	95% Confidence Interval for Mean		Minimum	Maximum
Age	Genova	30	6.2	3.9	6.0	4.7	7.6	1.0	14.0
	Calafuria	29	4.1	1.9	4.0	3.4	4.8	1.0	7.0
	Elba	30	6.3	2.0	6.2	5.6	7.0	3.0	10.0
	Palinuro	30	5.6	3.0	6.0	4.5	6.8	1.0	13.0
	Scilla	30	7.8	4.5	8.0	6.1	9.5	1.0	14.0
	Pantelleria	30	7.0	4.2	6.0	5.5	8.6	1.0	14.0
	Total	179	6.5	3.8	6.0	5.9	7.0	1.0	14.0
Length	Genova	30	4.8	2.2	4.9	4.0	5.7	2.0	9.0
	Calafuria	29	4.0	1.5	3.6	3.4	4.5	2.0	6.0
	Elba	30	6.3	2.0	6.2	5.6	7.0	3.0	10.0
	Palinuro	30	4.4	1.8	4.3	3.8	5.1	2.0	8.0
	Scilla	30	6.3	3.1	6.2	5.1	7.4	2.0	12.0
	Pantelleria	30	4.8	2.2	4.4	4.0	5.6	2.0	9.0
	Total	179	5.1	2.3	5.0	4.7	5.4	2.0	12.0

Figure 4.6 The length of *Leptopsammia pruvoti* vs age

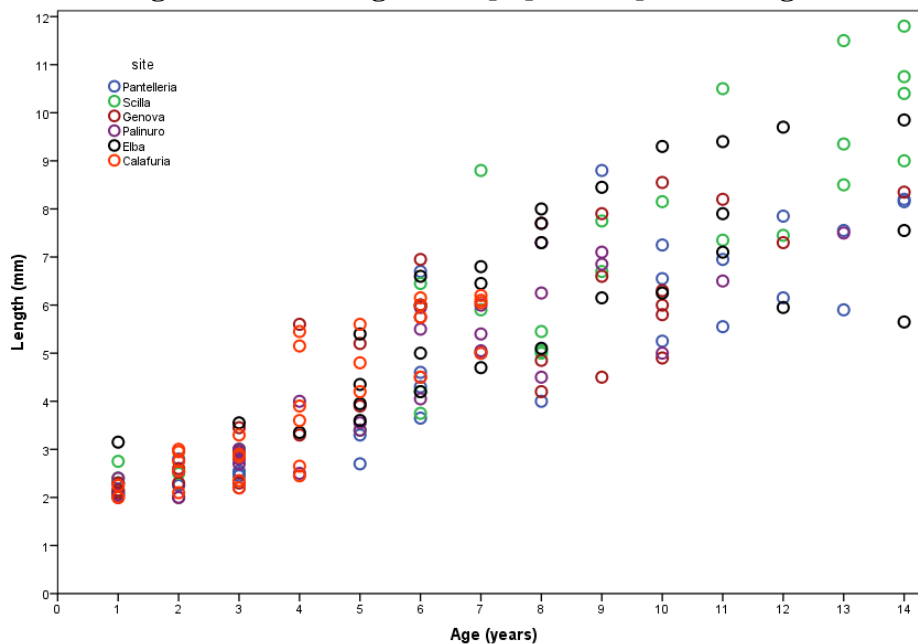


Figure 4.7 The length of *Leptosammia pruvoti* vs age for each site under study

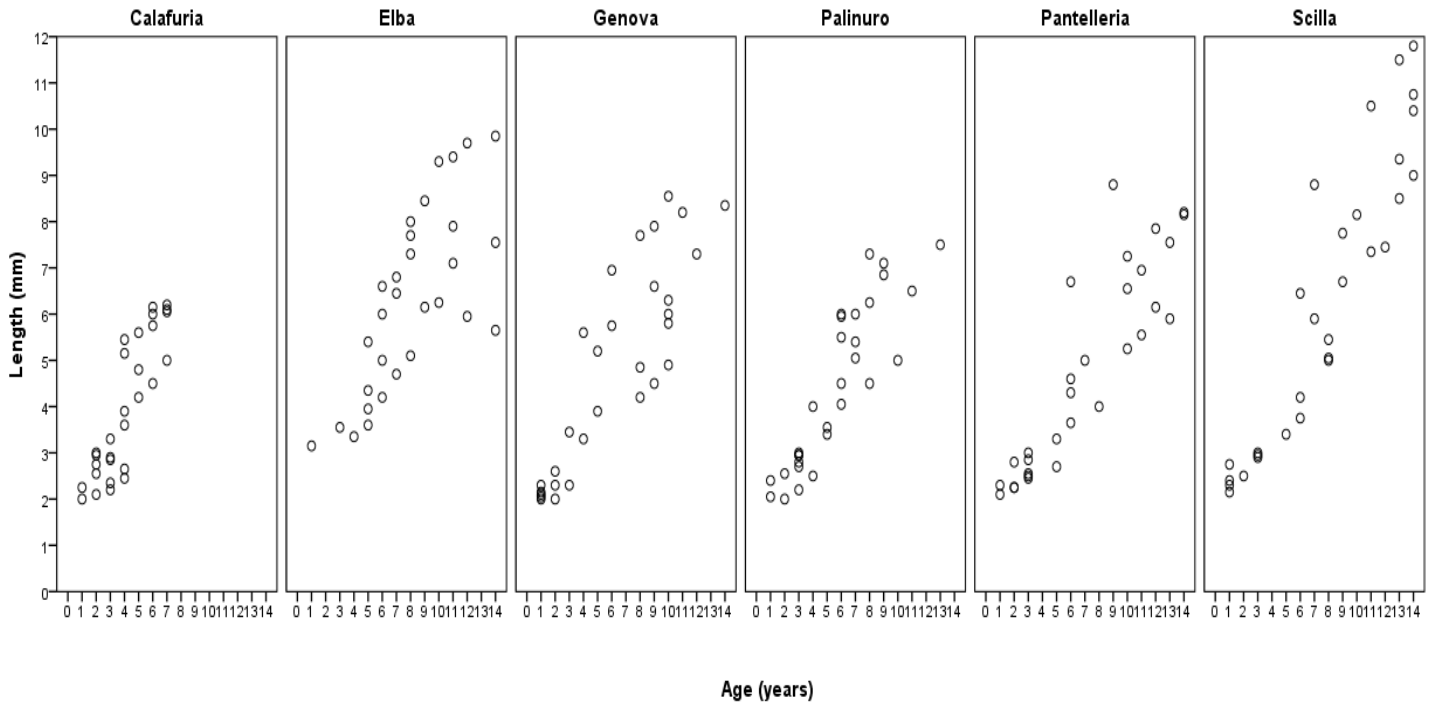
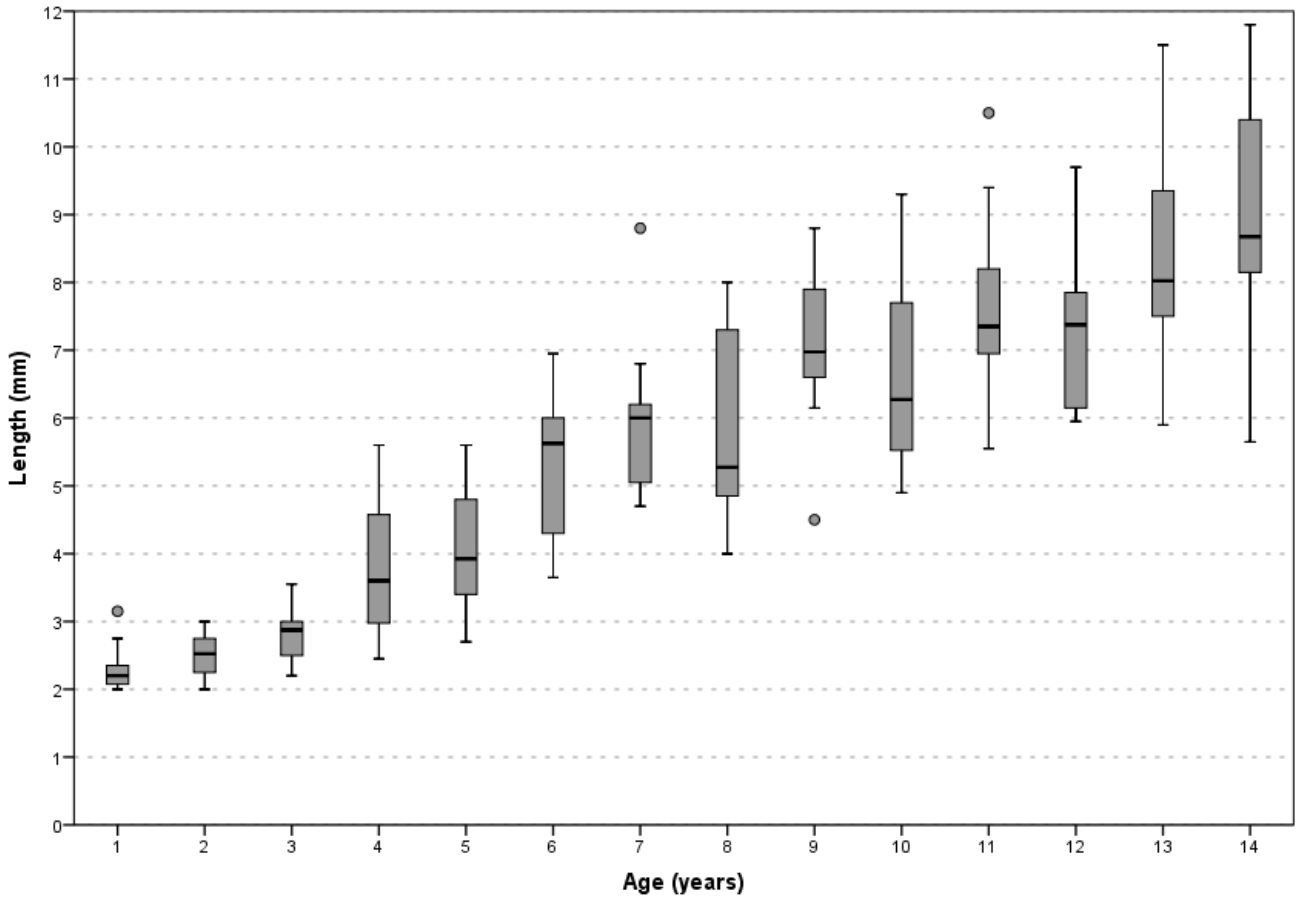


Figure 4.8 Boxplot of length of *Leptosammia pruvoti* at different ages



5 Results with the traditional Methods used in marine biology for the Von Bertalanffy Growth Function fitting

5.1 The Gulland-and-Holt (GH) plot

As pointed out in chapter (2.3.1) this method is used for the estimation of the two parameters L_∞ and k . and is based on the growth rates plotted vs. the mean length at first and second measurements using

Errore. L'origine riferimento non è stata trovata. and

Errore. L'origine riferimento non è stata trovata.. The method consisted of first calculating the elements

of the growth rates (the dependent variable) of the six sites, thus $y_j = \frac{L_j - L_{j-1}}{t_j - t_{j-1}}$ $j = 1, \dots, 14$, and the elements

of the mean length at first and second measurements (the independent variable), thus

$x_j = \frac{L_j + L_{j-1}}{2}$ $j = 1, \dots, 14$. After this a linear regression was fitted. At the end the fitted regression

coefficient was used to estimate parameter k as $k = -b$ and the estimated intercept was used to estimate

parameter L_∞ as $L_\infty = \frac{a}{k}$

The graphic in Figure 5.1 is obtained by applying this method to the *Balanophyllia europaea* data

The estimated regression coefficients were $\hat{a} = 2.7437$; $\hat{b} = -0.1493$ so that the estimation of parameter L_∞ ,

representing the ultimate length was 18.38 mm and the estimated growth rate k , was 0.149. The curve

presented in Figure 5.2 was built according to (2.5) using. $\hat{L}_\infty = 18.38$; $\hat{k} = 0.1493$.

Then standardized residuals e^* were calculated and plotted vs fitted values as shown in Figure 5.3. The

standardized residuals were calculated according to

$$e_j^* = \frac{e_j}{\hat{\sigma} \sqrt{1 - h_{jj}}} \quad 5.1$$

Where $\hat{\sigma}$ is the standard deviation of the residuals and h_{jj} is the Leverage of the j th estimated point calculated according to

$$h_{jj} = \frac{1}{N} + \frac{(e_j - E(e))^2}{\sum_j (e_j - E(e))^2} \quad 5.2$$

Figure 5.2 shows that the dimensions of young and old corals are underestimated; Figure 5.3 suggests that the dimensions of small corals are definitely underestimated. Furthermore, the residuals are asymmetrical, in particular the residuals of large corals are divided into two groups: one exceeding zero and the other one below zero. The overall impression is that the method underestimates parameter k producing a growth curve that is too slow at the young ages and too remote reaching the ultimate length. In addition the variance of the dimension seems to vary according to the dimension itself, but there is no way to consider this behaviour in the model.

Figure 5.1 Gulland and Holt plot for the *Balanophyllia europaea*

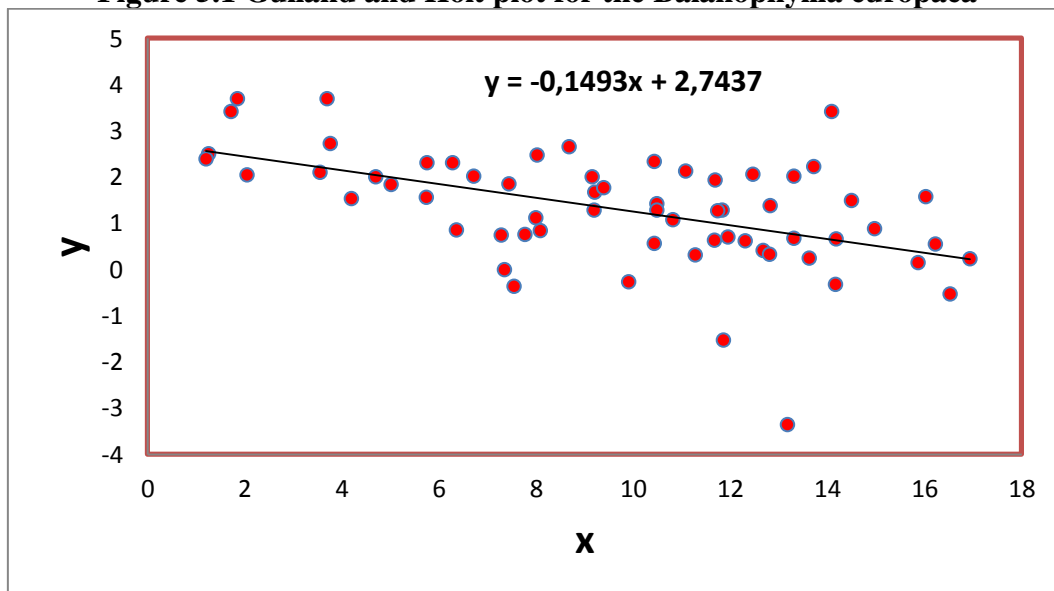


Figure 5.2 Curve fitting for the *Balanophyllia europaea* data using the Gulland and Holt plot method

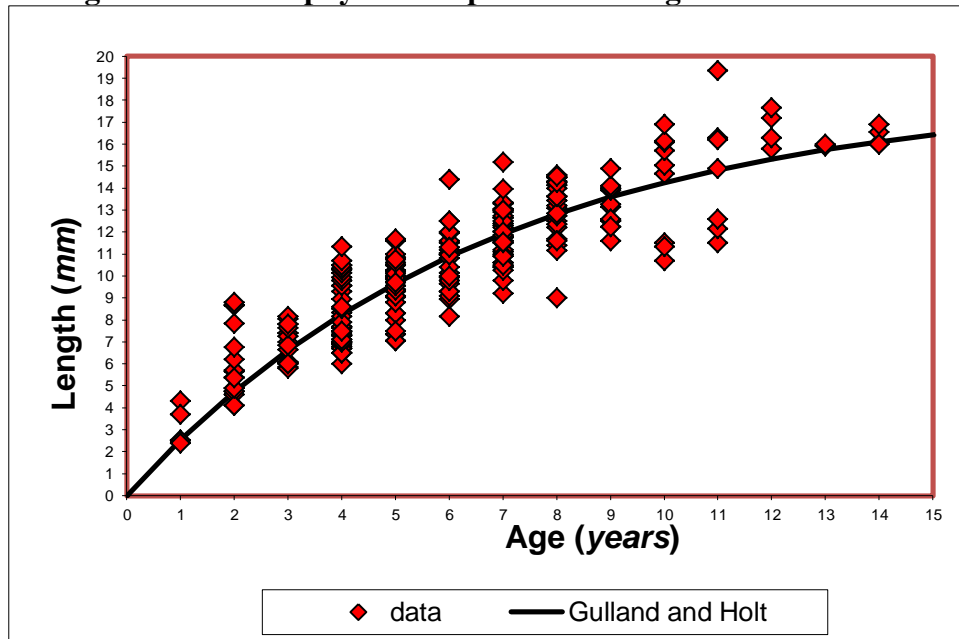
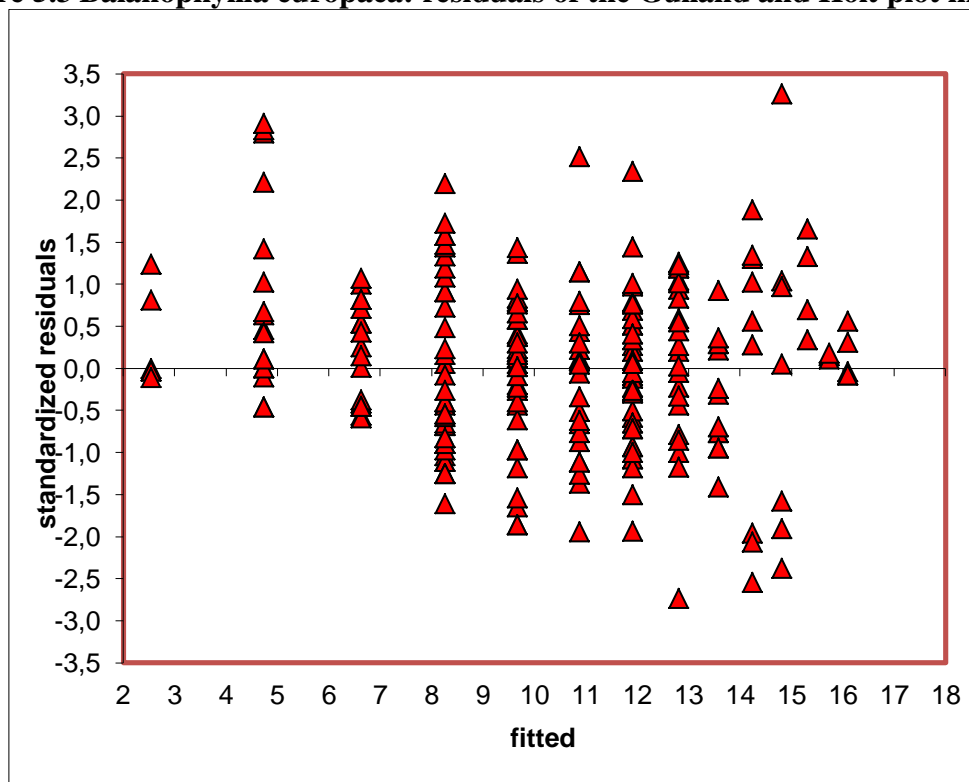


Figure 5.3 *Balanophyllia europaea*: residuals of the Gulland and Holt plot method



The graphic in Figure 5.4 is obtained by applying this method to the *Leptopsammia pruvoti* data

As for the *Balanophyllia* the linear regression of the data shown in Figure 5.4 allows the estimation of

$\hat{a} = 1.2906; \hat{b} = -0.1224$ so that $\hat{L}_\infty = 10.54; \hat{k} = 0.122$. As for *Balanophyllia* data, these two parameters are

used for the construction of the curve shown in Figure 5.5; this curve, in turn, gives origin to the residuals

plotted in Figure 5.6. Here the dimensions of very young corals are also underestimated; furthermore the residuals are very asymmetrical showing a “U” shape; in particular the residuals of big corals are all above zero, instead those of medium dimensions are below zero; this might mean the presence of a nonlinear relationship not correctly considered in the model. Also in this case, the overall impression is that the forced linearization introduced in the estimate of the parameters leads to the underestimation of parameter k thus causing the curve to grow too slowly.

Figure 5.4 Gulland and Holt plot for the *Leptosammia pruvoti*

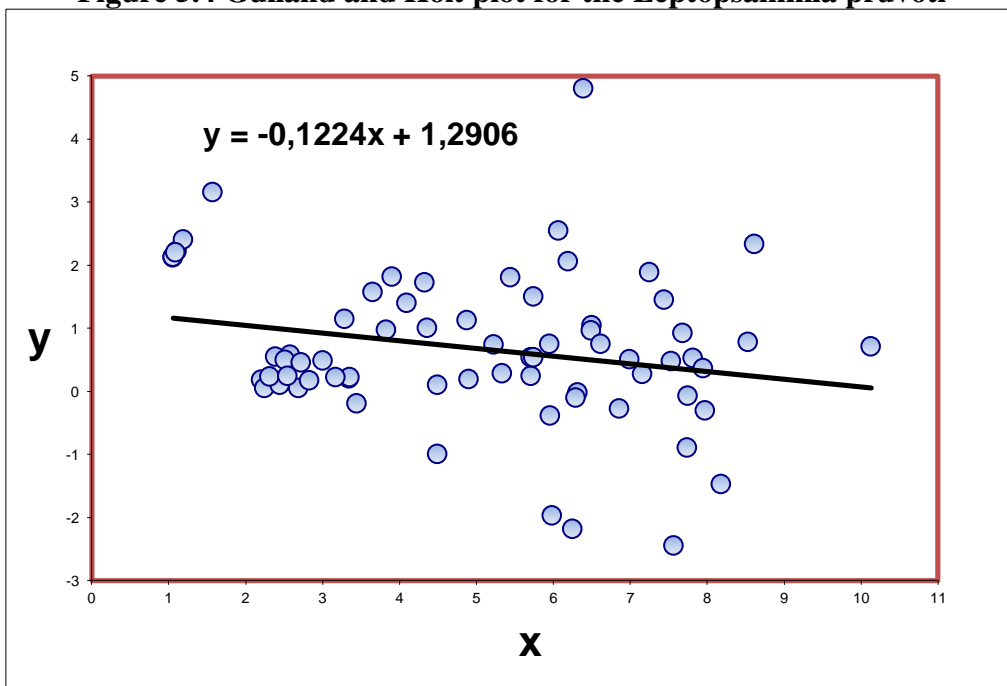


Figure 5.5 Curve fitting for the *Leptosammia pruvoti* data using the Gulland and Holt plot method

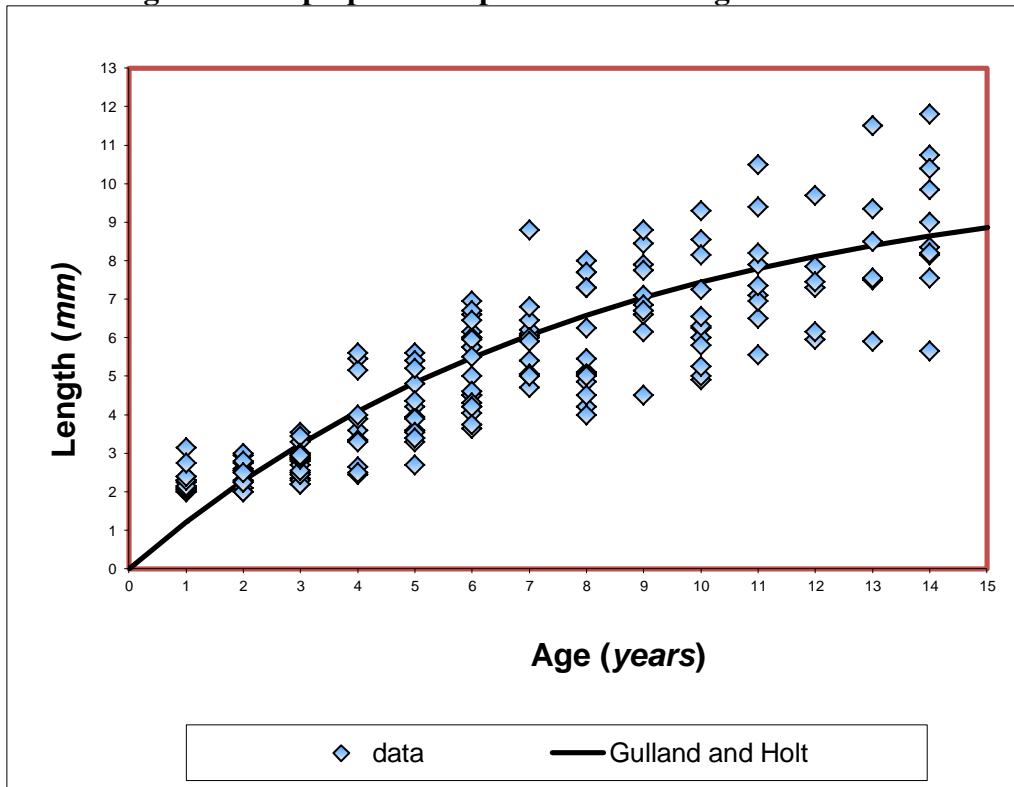
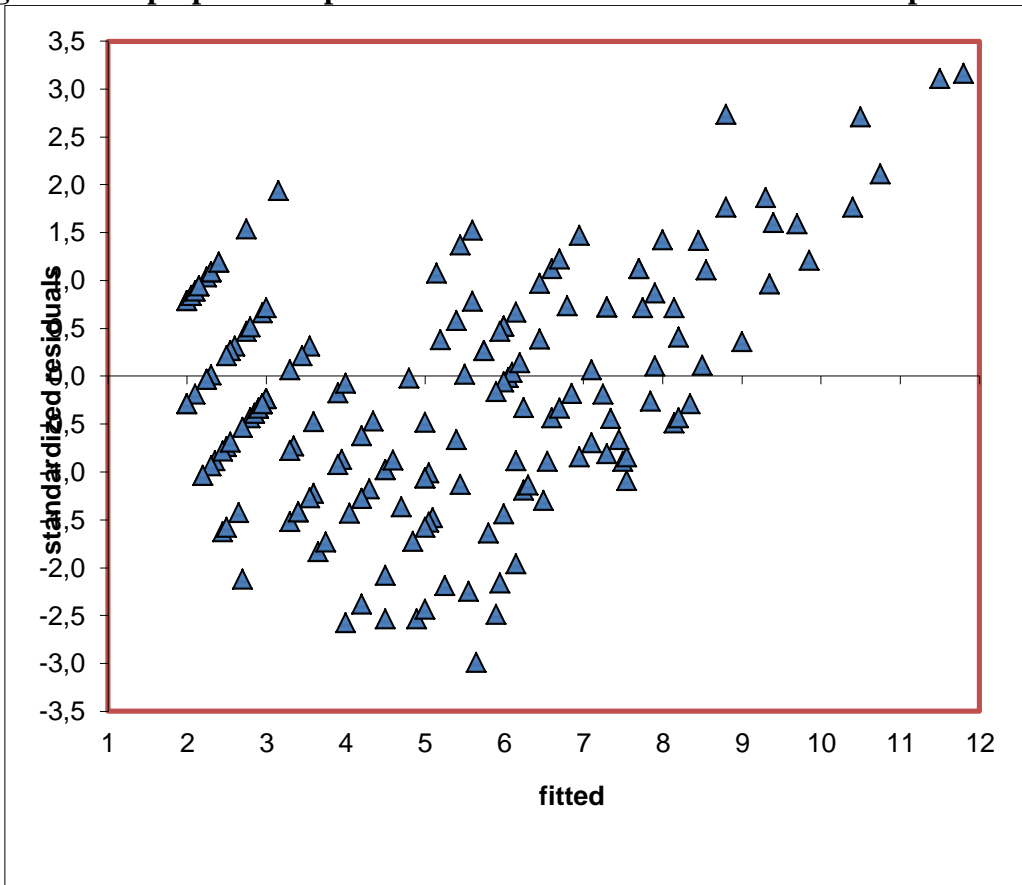


Figure 5.6 *Leptosammia pruvoti*: residuals of the Gulland and Holt plot method



5.2 The size-increment method proposed by Fabens

As already seen in chapter (2.3.2) the size-increment method proposed by Fabens (1965), fits the von Bertalanffy curve by applying the least squares method to the observed lengths, using data of known growth increments in known time intervals but making no assumption about absolute age according to

Errore. L'origine riferimento non è stata trovata. For this data Δt corresponds to one year. The linear regression of $L_{t+\Delta t}$ against L_t will generate a slope of $b = e^{-k}$ and an intercept of $a = L_\infty(1 - e^{-k})$. The

parameters can then be calculated according to $k = -\ln(b)$ and $L_\infty = \frac{a}{1 - e^{-k}}$

The graphic in Figure 5.7 is obtained by applying this method to the *Balanophyllia europaea* data and

considering $y = L_{t+\Delta t}$ and $x = L_t$. The linear regression results are then $\hat{a} = 2.7897; \hat{b} = 0.8455$ so that

$\hat{L}_\infty = 18.06; \hat{k} = 0.168$. The curve in Figure 5.8 is built using these two parameters and originates the

residuals plotted in Figure 5.9. From Figure 5.8 it is possible to argue that the dimensions of old corals are

underestimated. Furthermore, Figure 5.9 shows that the residuals are asymmetrical; in particular the

residuals of big corals are divided into two groups: one mostly above zero and the other one strongly below

zero. The overall impression is that this method is more suitable than the previous one but is still not reliable

for old corals because of the presence of at least two groups of different dimensions for the same age. This is

a problem that is difficult to solve without considering the source of this difference. Here too the variance of

the dimensions seems to vary according to the dimensions themselves, but there is no way to consider this

behaviour in the model.

Figure 5.7 Fabens plot for the *Balanophyllia europaea* data

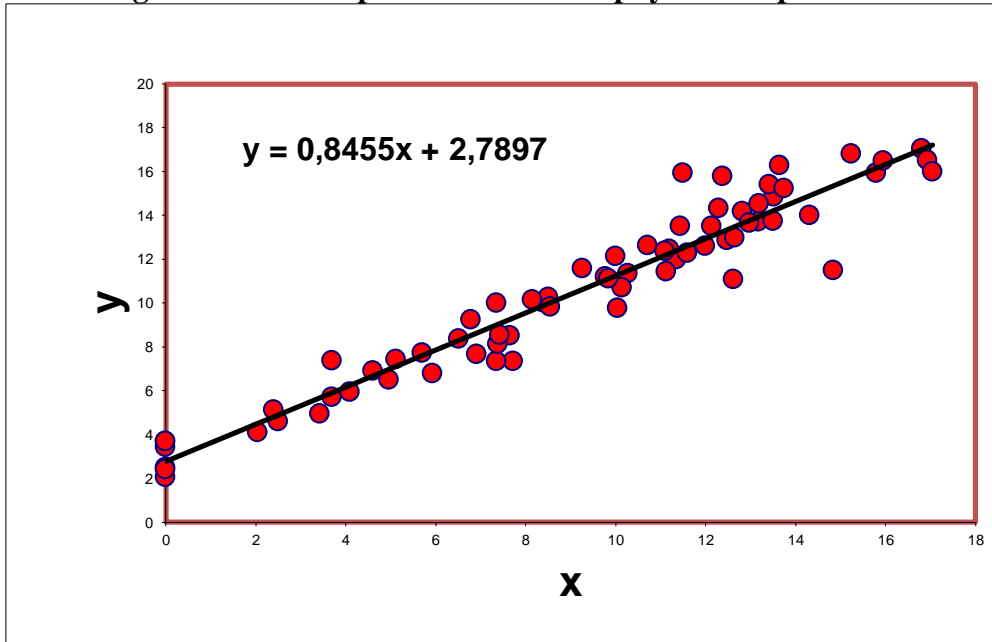


Figure 5.8 Curve fitting for the *Balanophyllia europaea* data using the Fabens method

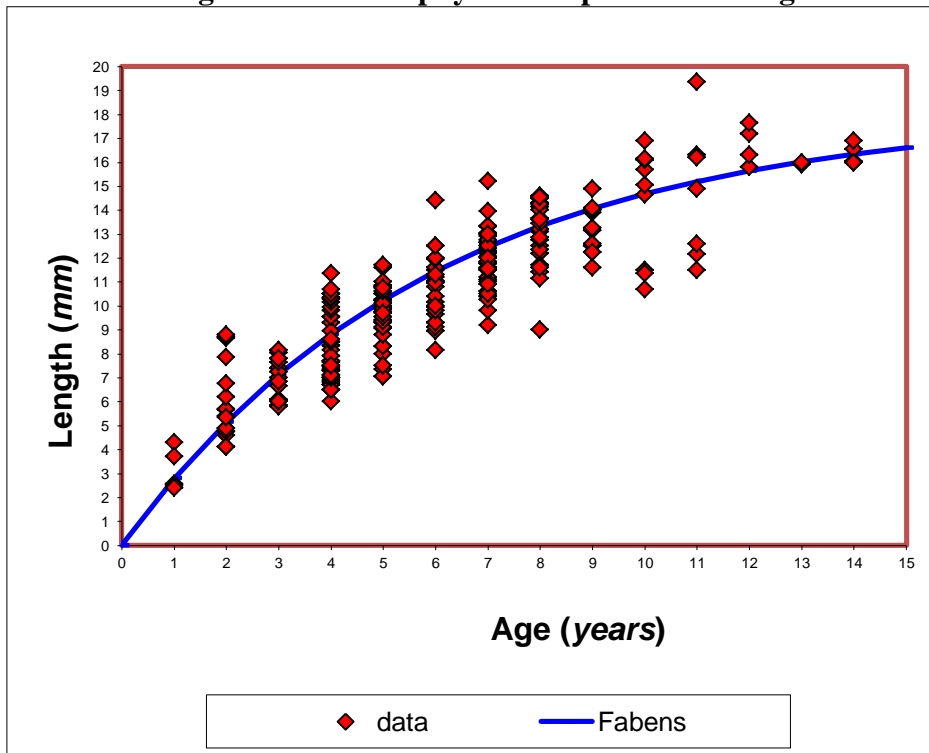
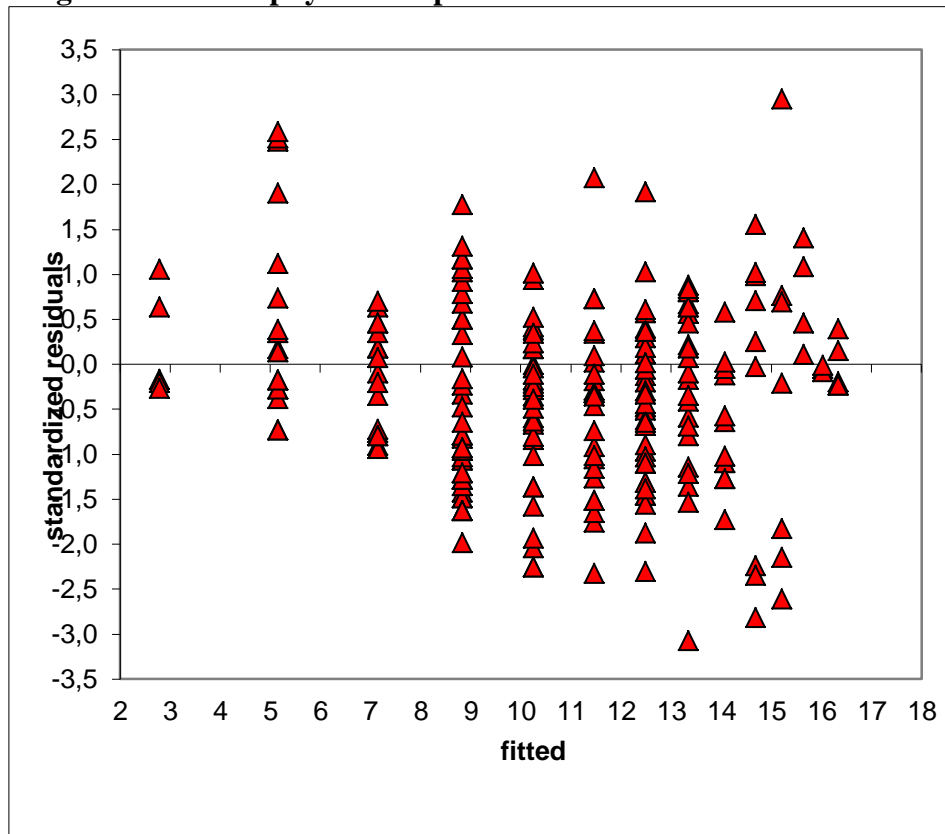


Figure 5.9 Balanophyllia europaea: residuals of the Fabens method



The graphic in Figure 5.10 is obtained by applying the Fabens method to the *Leptopsammia pruvoti* data.

The linear regression estimations are $\hat{a} = 1.6532$; $\hat{b} = 0.7917$ the parameter estimation is then $\hat{L}_\infty = 7.94$; $\hat{k} = 0.233$. The curve shown in Figure 5.11 is constructed using these two parameters and originates the residuals plotted in Figure 5.12. From Figure 5.11 it is possible to argue that the dimensions of very young and old corals are underestimated. Corals 2 to 4 years old are overestimated; furthermore Figure 5.12 shows that the residuals are very asymmetrical with an asymmetrical “U” shape also in this case there is a suspicion of the presence of nonlinearity not well explained by the model and in particular the residuals of big corals are strongly above zero. The overall impression is that the method because of its forced linearization tends to underestimate parameter L_∞ so that all the older corals are underestimated. Here too the variance of the dimensions seems to vary according to the dimensions themselves,

Figure 5.10 Fabens plot for the *Leptopsammia pruvoti* data

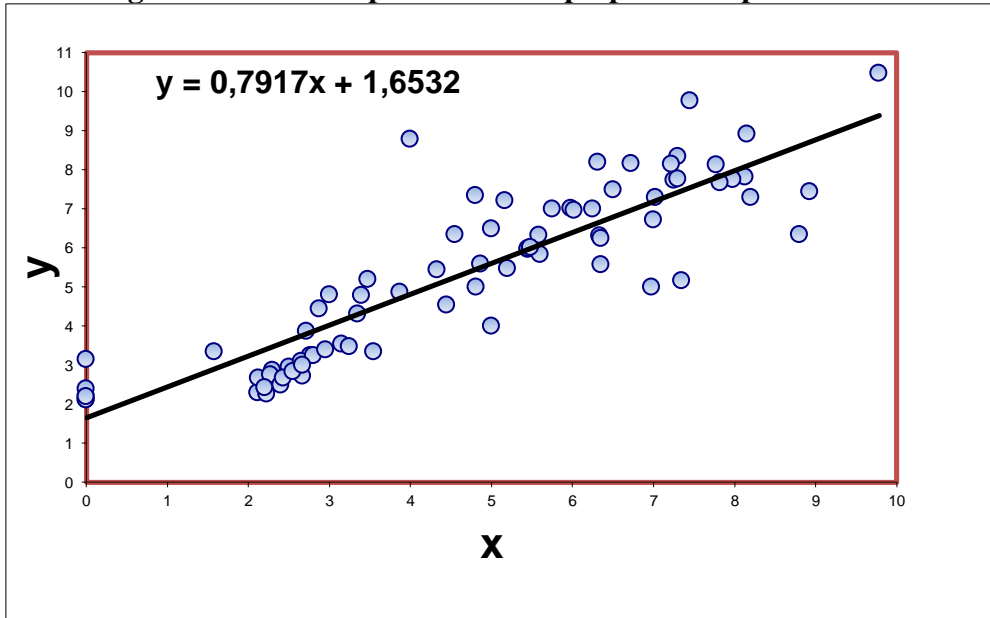


Figure 5.11 Curve fitting for the *Leptopsammia pruvoti* data using the Fabens method

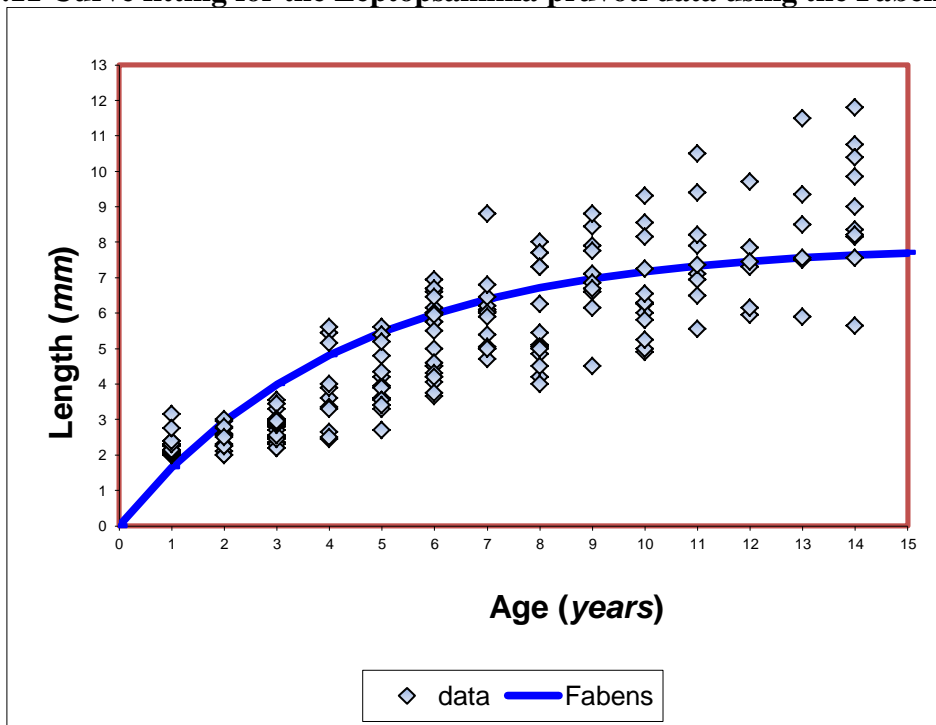
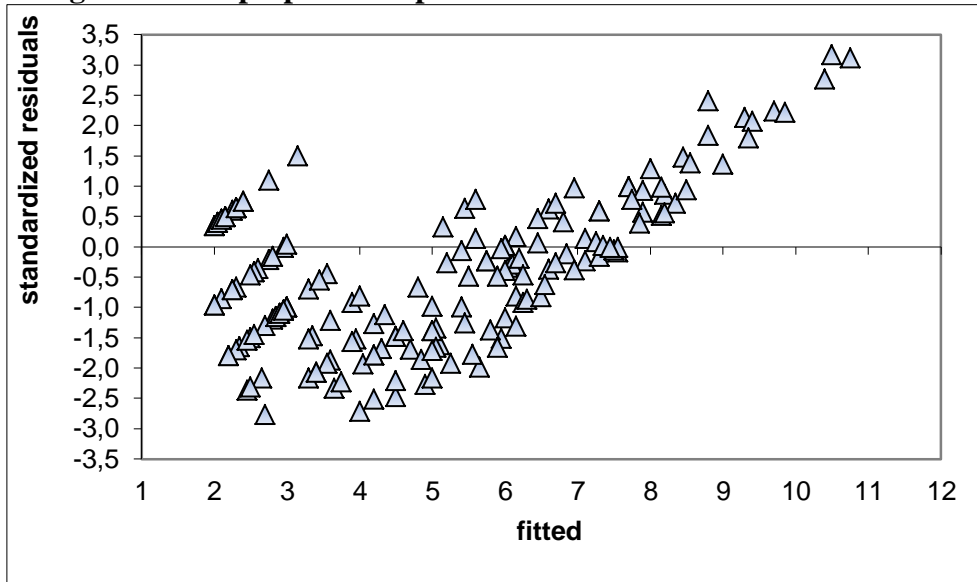


Figure 5.12 *Leptosammia pruvoti*: residuals of the Fabens method



5.3 The linearization proposed by Basso and Kehr

This method fits the von Bertalanffy model by imposing as L_∞ the largest size found in individuals of each species and then considering the linear regression between $\ln\left(1 - \frac{L_t}{L_\infty}\right)$ as the dependent variable and t as the independent variable. So the estimate of k , considering $kt_0 = L_0 = 0$ for corals, will be obtained by the solution via the least square method of the equation **Errore. L'origine riferimento non è stata trovata.** passing through the origin

$$\ln\left(1 - \frac{L_t}{L_\infty}\right) = -kt \quad 5.3$$

so the biggest length measured in the *Balanophyllia europaea* data is 19.40 mm then $\hat{L}_\infty = 19.40\text{mm}$; the graphic of Figure 5.13 calculating parameter k according to (5.2) is based on the linear regression of y over x where $y = \ln\left(1 - \frac{L_t}{L_\infty}\right)$ and $x = t$. According to (5.2) the slope will then be the additive inverse value of the desired parameter k .

The result is then $\hat{L}_\infty = 19.40$; $\hat{k} = 0.142$. The curve in Figure 5.14 is built using these two parameters; the residuals plotted in Figure 5.15 are then calculated. Figure 5.13 shows a wide outlier which corresponds to

the coral used to estimate the ultimate length L_{∞} in so far as it has the greatest measured length. This is a very weak point in the method because it is impossible to eliminate the outlier and, in the meantime, the estimate of k would appreciably change without considering the outlier. From Figure 5.14 it is possible to argue that the dimensions of young corals are underestimated. Figure 5.15 shows that the residuals might depend linearly on fitted values so that they are symmetrical with respect to a line different from the abscissa. This method depends too much on the sample because parameter L_{∞} is directly chosen from the data; so in this case there might be an underestimation of L_{∞} which causes a possible overestimation of k .

Figure 5.13 Basso and Kehr plot for the *Balanophyllia europaea*

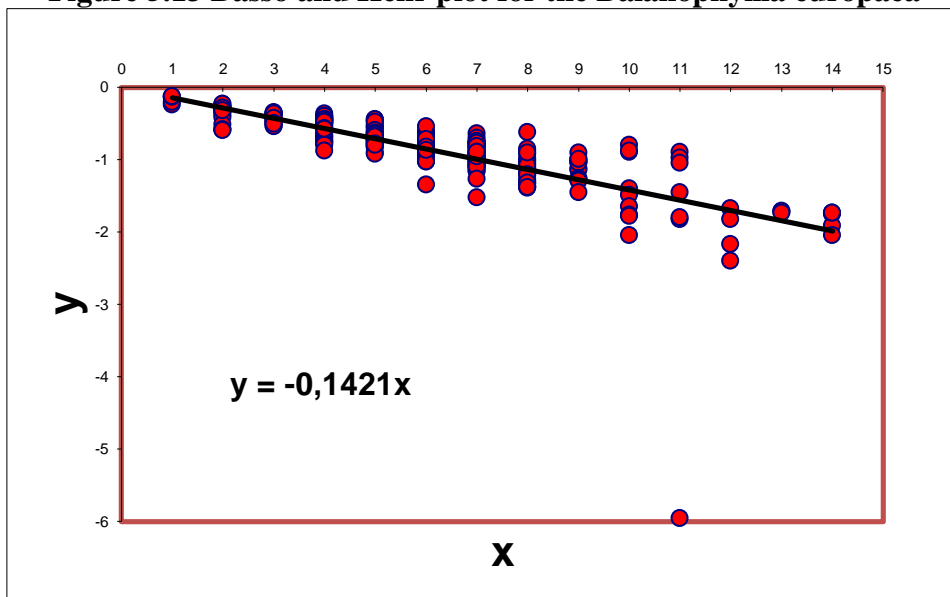


Figure 5.14 Curve fitting for the *Balanophyllia europaea* data using the Basso and Kehr method

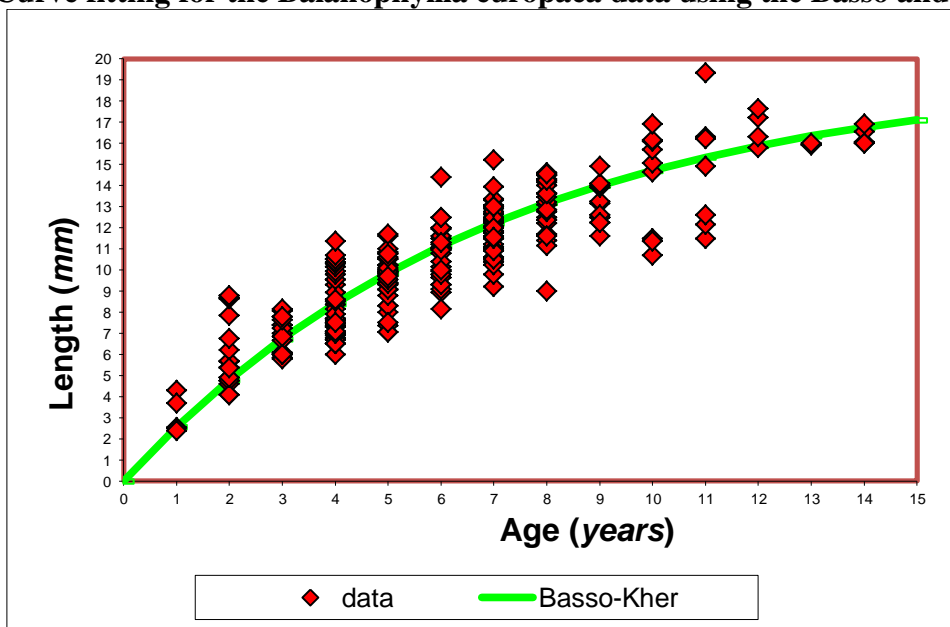
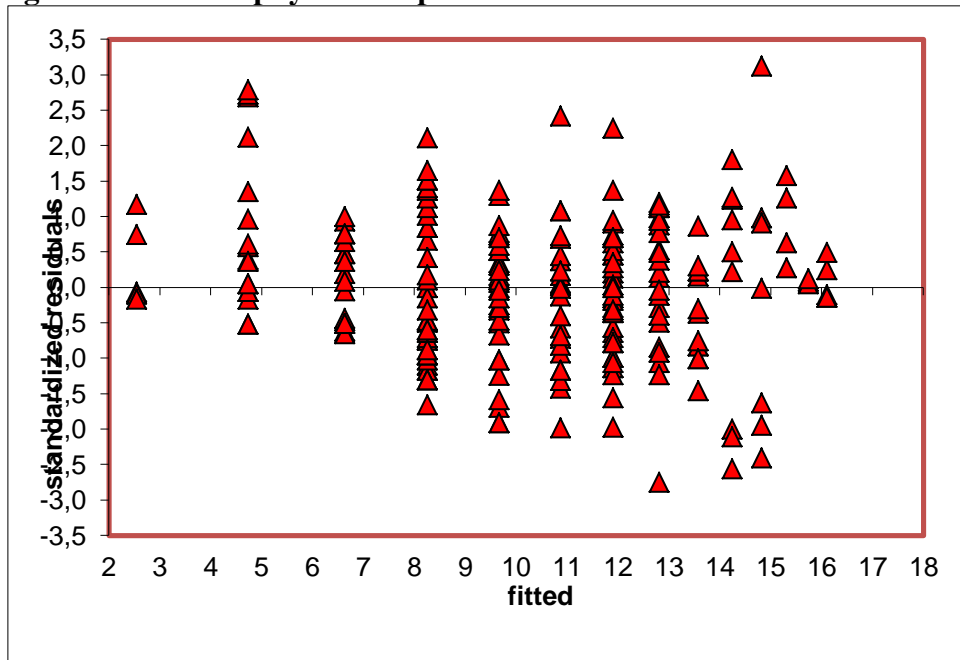


Figure 5.15 Balanophyllia europaea: residuals of the Basso-Kher method



Applying this method to the *Leptosammia pruvoti* data it is possible to determine from the data $\hat{L}_\infty = 11.85mm$; the graphic in Figure 5.16 enables the estimation of parameter k . The result is then $\hat{L}_\infty = 11.85$; $\hat{k} = 0.104$. The curve shown in Figure 5.17 is built using these two parameters and produces the residuals plotted in Figure 5.18. By observing Figure 5.16 the presence of a wide outlier corresponding to the data used for the estimation of the ultimate length can also be noticed here. From Figure 5.17 it is possible to argue that the dimensions of young corals are underestimated, Figure 5.18 shows that the residuals are very asymmetrical and with a “U” shape; in particular the residuals of big corals are strongly above zero and those of medium size are below zero and of course the “U” shape could be a symptom of not considered nonlinearity. Also in this case there might be an underestimation of L_∞ which explains the lack of fit for old-big corals.

Figure 5.16 Curve fitting for the *Leptosammia pruvoti* data using the Basso and Kehr method

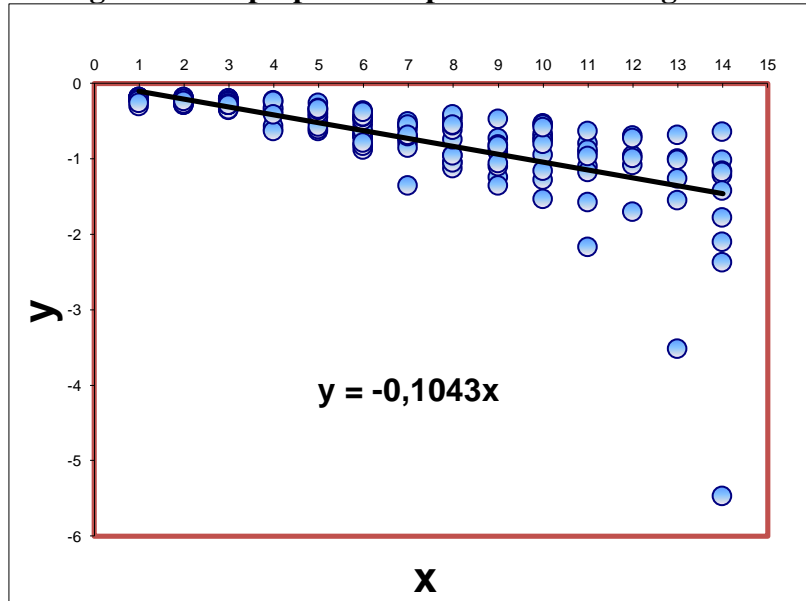


Figure 5.17 Curve fitting for the *Leptosammia pruvoti* data using the Basso and Kehr method

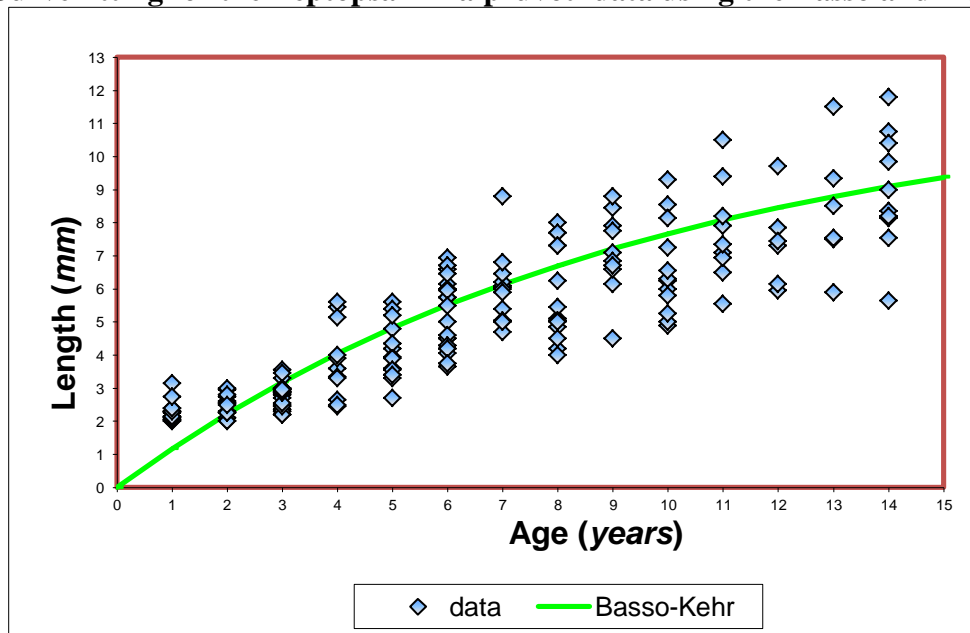
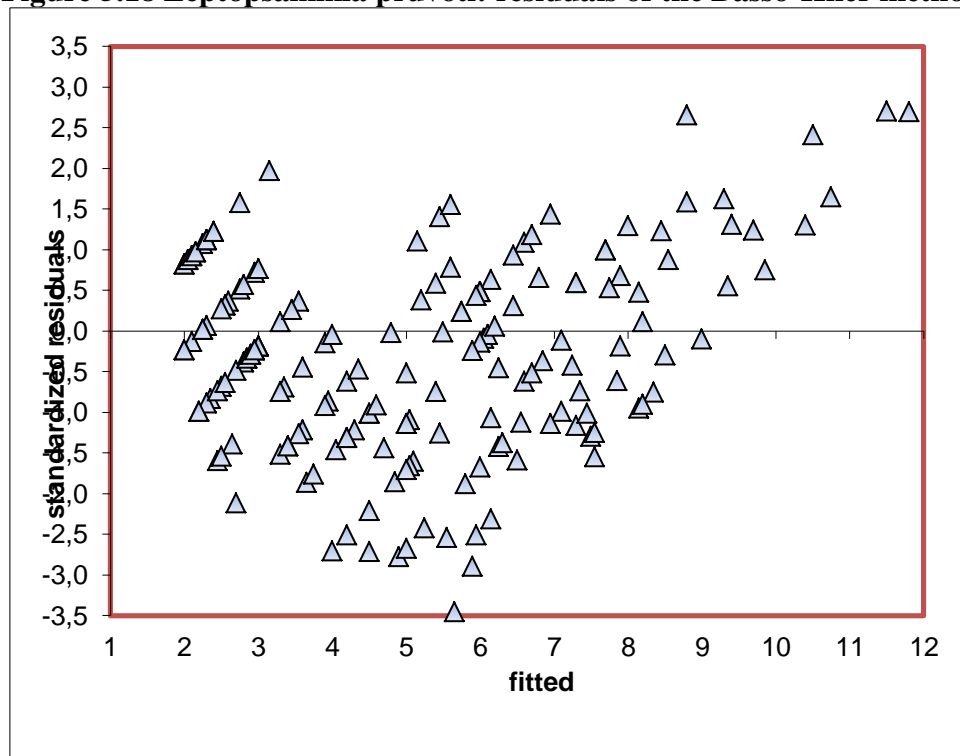


Figure 5.18 *Leptopsammia pruvoti*: residuals of the Basso-Kher method



5.4 Newton's method for nonlinear function estimation

This method fits the von Bertalanffy model following the method described in chapter 3.3, i.e. the two parameters of the VBGF are estimated by the least square method, by minimizing the **Errore. L'origine riferimento non è stata trovata.** via a numeric method based on the Gauss-Newton iterative algorithm. This method does not take into account the collection sites or the influence of covariates. The estimation procedure was carried out using the R function `nls()`, (Bates and Chambers, 1992). The results of the Gulland and Holt plot method were inserted as starting values. In the following paragraphs the method is applied to both parameterizations.

5.4.1 The standard parameterization

Applying the Least square nonlinear regression to the VBGF standard parameterization, the function to minimize **Errore. L'origine riferimento non è stata trovata.** becomes

$$\text{RSS}(\phi) = \sum_{j=1}^n \left(y_j - L_{\infty} (1 - e^{-kt_j}) + \varepsilon_j \right)^2$$

the use of Newton's method to estimate the parameters gives the following results $\hat{L}_{\infty} = 19.16$ with a SE equal to 0.65, $\hat{k} = 0.154$ and SE=0.010 for *Balanophyllia europaea* and $\hat{L}_{\infty} = 10.92$ with a SE equal to 0.87, $\hat{k} = 0.107$ and SE =0.014 for *Leptopsammia pruvoti*.

The above estimated parameters are used to build the growth curve presented in figures 5.19 and 5.20, both graphics show the underestimation of values for very young and old corals, more evident in the *Balanophyllia* data. The fit for corals aged 2 to 9 years is nevertheless improved.

Figure 5.19 Curve fitting for the *Balanophyllia europaea* data using the nls method

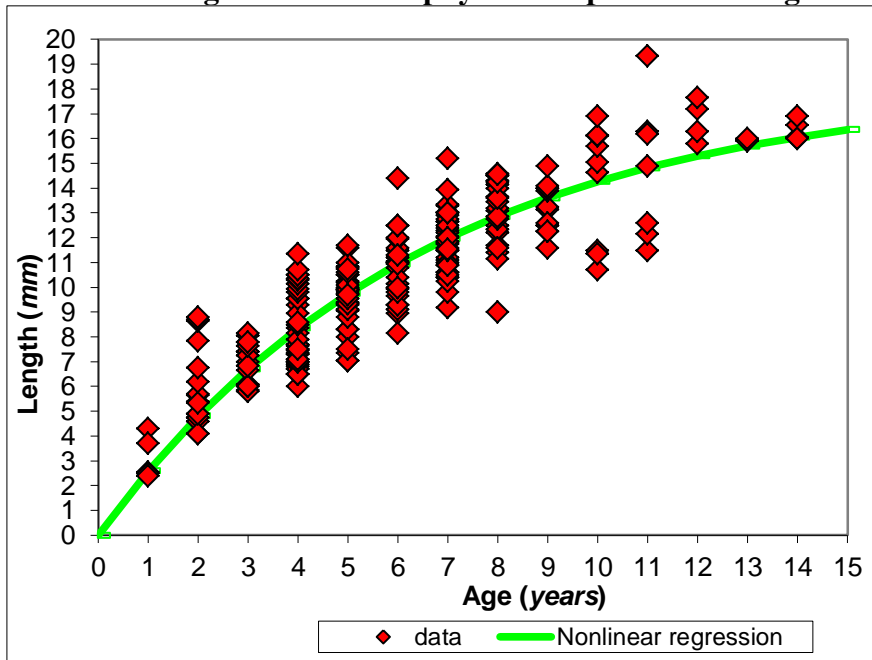
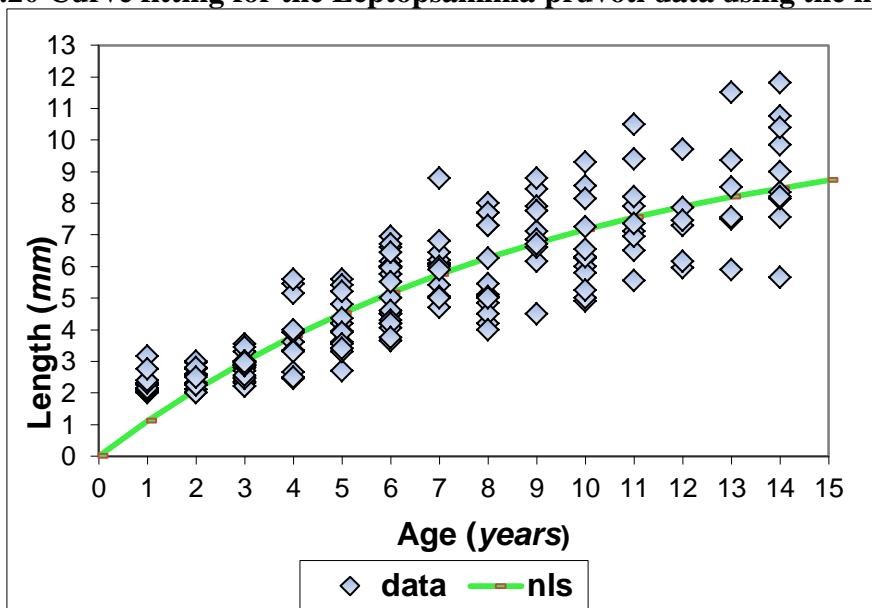


Figure 5.20 Curve fitting for the *Leptopsammia pruvoti* data using the nls method



From the residuals of the model values plotted vs the fitted values in Figure 5.21 A, for *Balanophyllia europaea*, and 5.22 A, for *Leptopsammia pruvoti*, it is possible to extrapolate that residuals are mostly above zero in small corals and that their dispersion around zero increases with the fitted size. The assumption of a normal distribution of residuals was investigated by analysing the residuals plotted vs the quantiles of a standard normal distribution presented in Figure 5.21 B, for *Balanophyllia europaea*, and 5.22 B, for *Leptopsammia pruvoti*, both plots show that residuals are normally distributed except for the extremes. This

method introduces an improvement in comparison to the methods used by biologists but there are still margins for improvement.

Figure 5.21 Residuals with nls method for *Balanopyllia europaea* data

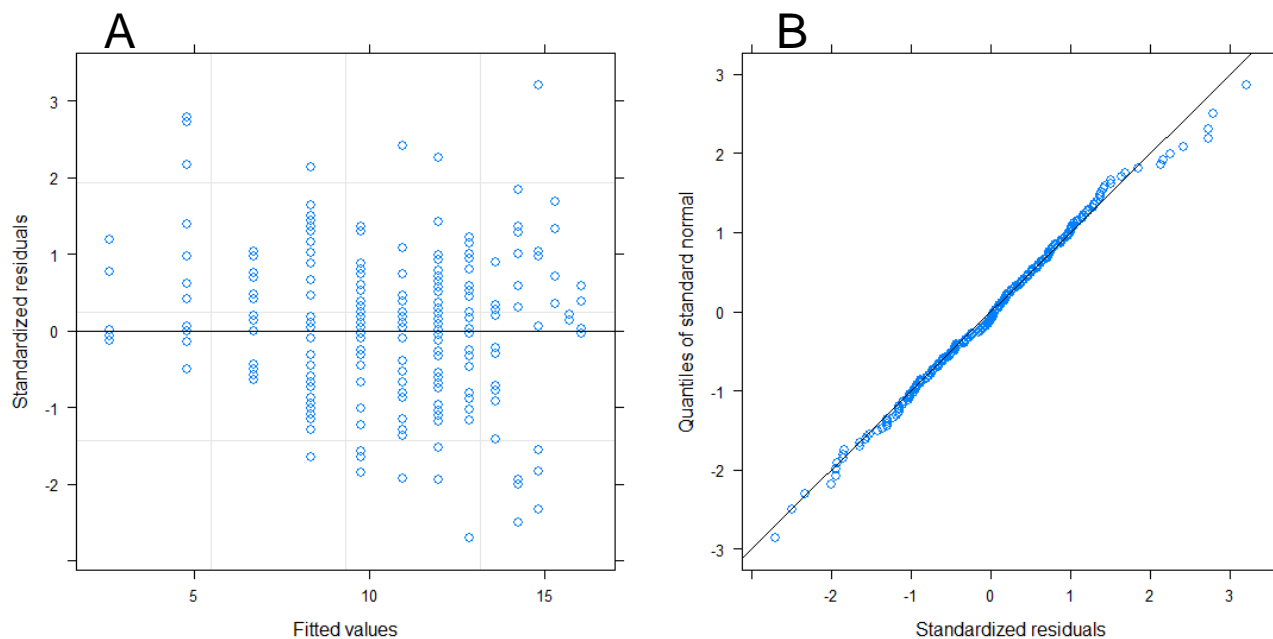
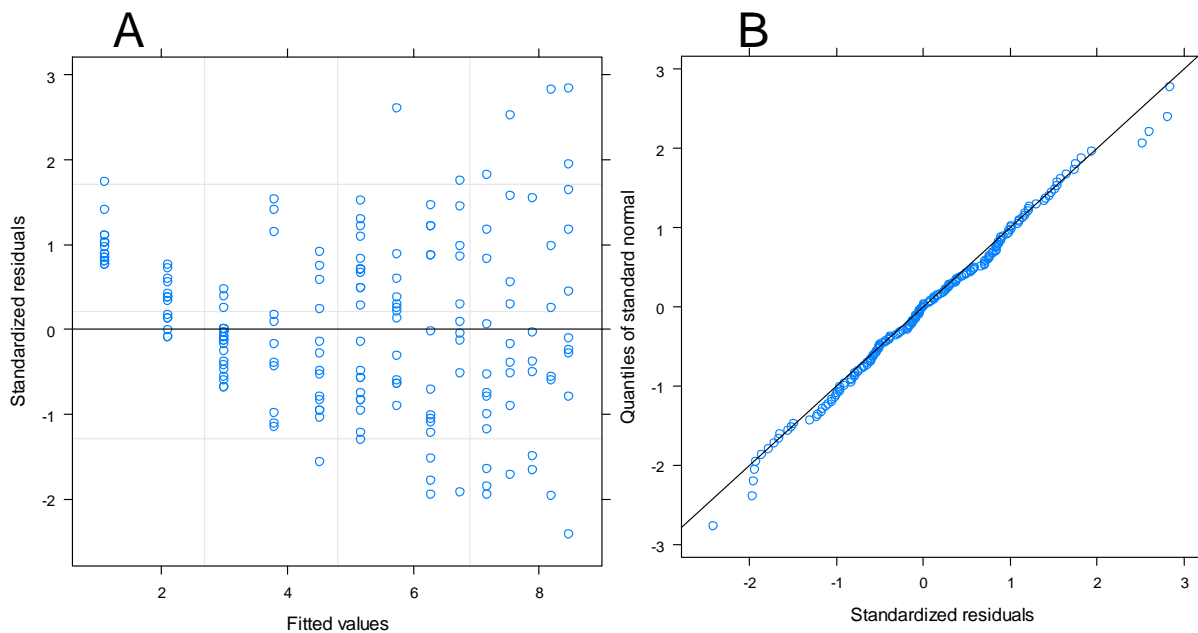


Figure 5.22 Residuals with nls method for *Leptopsammia pruvoti* data



Newton's method is very sensitive to the presence of outliers; this is one explanation for the poor performance around the extreme values; in addition the increase of residuals with size can be avoided by introducing a suitable model for the variance. The poor performance with small and wide corals can be overcome by enriching the model with different patterns of growth according to corals origin sites: this leads to Hierarchical nonlinear models.

5.4.2 The new parameterization

By applying the Least square nonlinear regression to the VBGF new parameterization the function to minimize **Errore. L'origine riferimento non è stata trovata.** becomes

$$RSS(\phi) = \sum_{j=1}^n \left(y_j - L_{\infty} \left(1 - e^{-t_j e^{\frac{c}{L_{\infty}}}} \right) + \varepsilon_j \right)^2$$

the use of Newton's method gives the following estimations of the parameters of the growth curve: $\hat{L}_{\infty} = 18.16$ with a SE equal to 0.65, $\hat{c} = -33.98$ and SE = 2.377 for the *Balanophyllia europaea* and $\hat{L}_{\infty} = 10.91$ with a SE equal to 0.86, $\hat{c} = -24.40$ and SE = 3.345 for *Leptopsammia pruvoti*.

The above estimated parameters are used to build the growth curve presented in Figure 5.23 and Figure 5.24 which show a better performance than the standard parameterization in matching the old corals, even if there is still an underestimation of very young corals.

Figure 5.23 Curve fitting for the *Balanophyllia europaea* data using the nls method

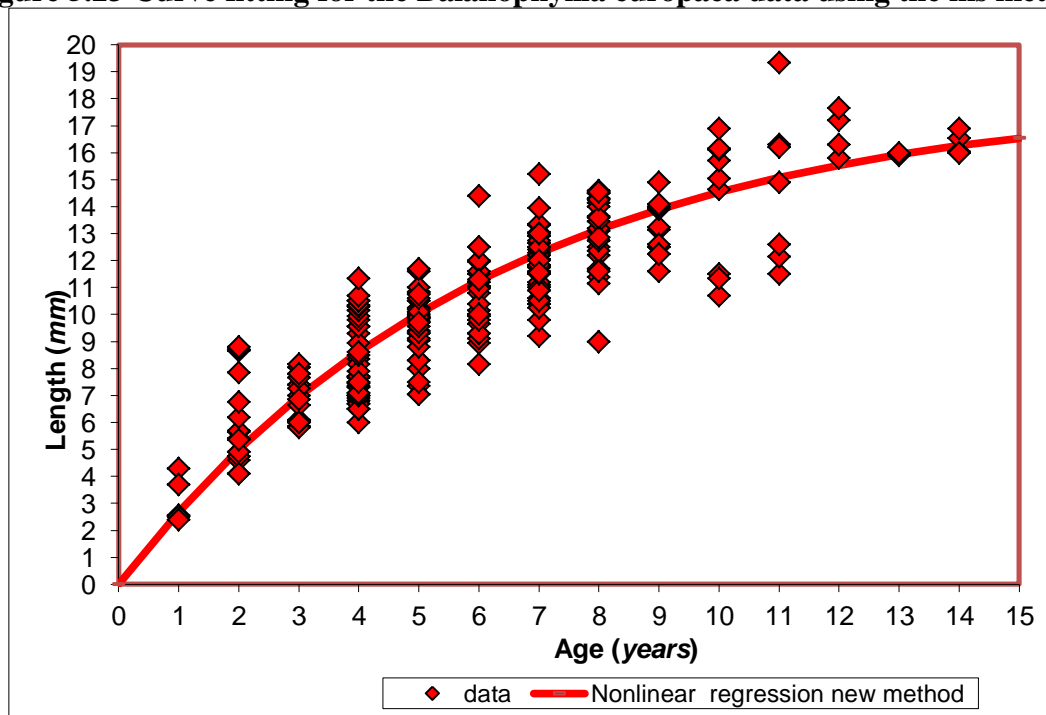
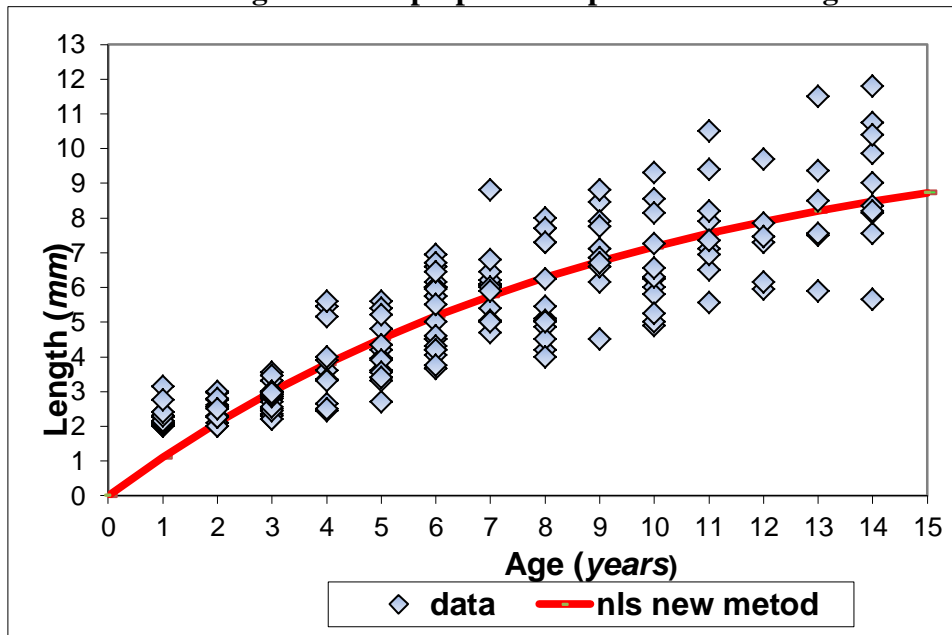


Figure 5.24 Curve fitting for the *Leptopsammia pruvoti* data using the nls method



Diagnostic plots were then made to assess the hypotheses about the residuals. From Figure 5.25 A, showing the residuals plotted vs the fitted value for the *Balanophyllia europaea* data, it is possible to deduct that residuals of small corals are mostly above zero, bigger corals have widely spread residuals. Figure 5.26 A, representing the residuals plotted vs the fitted value for the *Leptopsammia pruvoti* data, shows a linear decreasing trend of small corals residuals; the wide dispersion of residuals of medium and large dimensions are confirmed. Figure 5.25 B (*Balanophyllia europaea*) and 5.26 B (*Leptopsammia pruvoti*) showing the residuals plotted vs the quantiles of a standard normal distribution confirm the normal distribution of residuals but for the extremes for both the studied corals.

Figure 5.25 Residuals with nls method for *Balanopyillia europaea* data

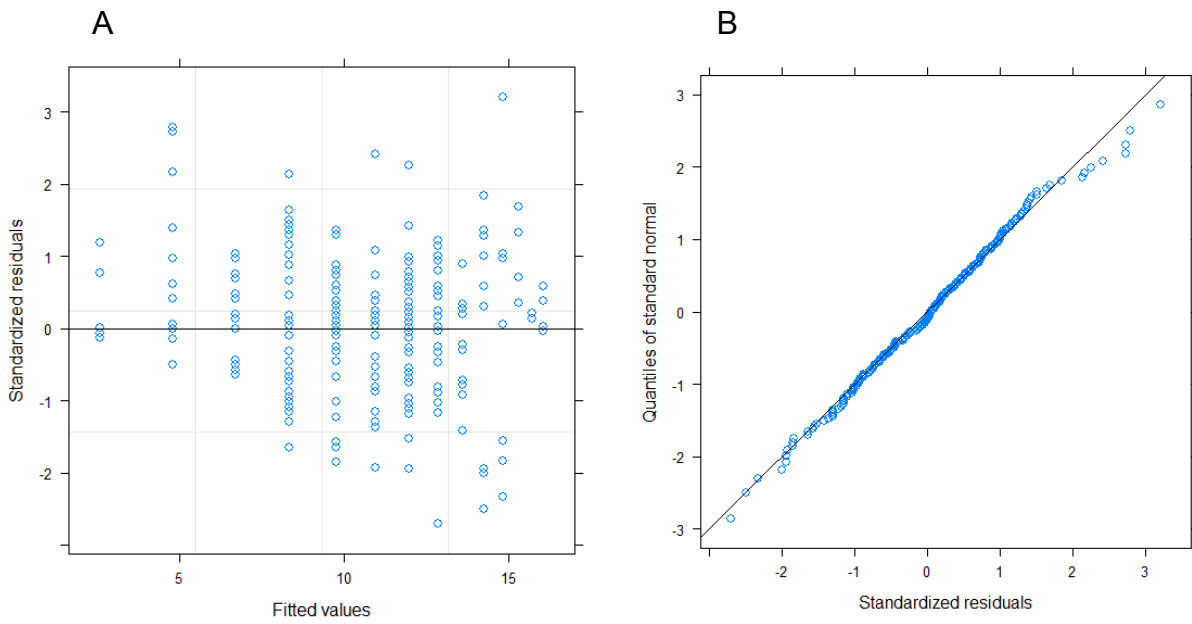
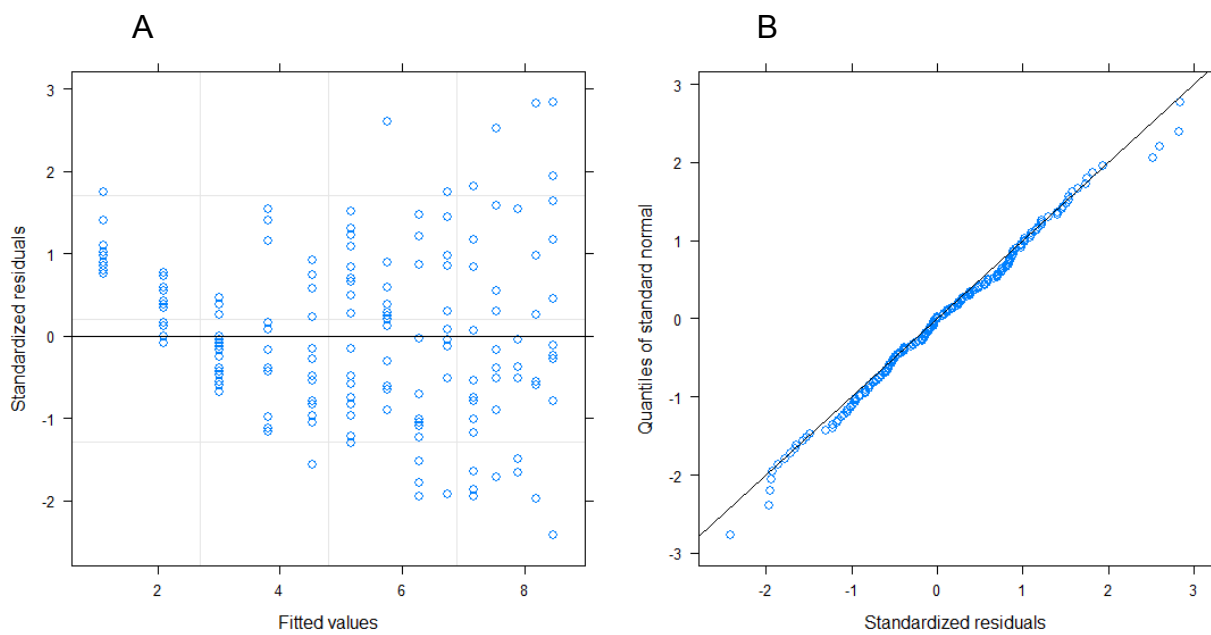


Figure 5.26 Residuals with nls method for *Leptopsammia pruvoti* data



5.5 Results with nonlinear hierarchical model

The implementation of hierarchical models was based on the method described in paragraphs 3.4 and 3.5 according to which two levels can be defined: the site level and the coral level. So each site will have his own growth curve designed according to a couple of site-specific parameters: the parameters are composed

of a value common to all corals, i.e. the fixed effect, and a value specific for each site that is the random effect. This formulation assumes the within-site errors as *i.i.d.* and normally distributed.

The general approach used for the analysis was similar for both corals and both parameterizations; we started by performing the most simple model i.e. considering both parameters having fixed and random effects according to the **Errore. L'origine riferimento non è stata trovata.** for the standard parameterization and the **Errore. L'origine riferimento non è stata trovata.** for the new parameterization; any particular correction for the possible intra-site correlation and heteroscedasticity was taken into account. We then tried with one parameter having a fixed and random effect and the other one having the fixed effect only. The models were compared by choosing the one improving the model comparison criteria such as the Bayesian Information Criterion (BIC) (Schwarz, 1978) and the Akaike Information Criterion (AIC) (Sakamoto et al., 1986). According to the model having the smallest, and then, the best BIC and AIC the **Errore. L'origine riferimento non è stata trovata.** becomes

$$\phi_i = \begin{pmatrix} L_{\infty i} \\ k_i \end{pmatrix} = \begin{pmatrix} L_{\infty} \\ k + b_{2i} \end{pmatrix} \quad 5.4$$

for the *Balanophyllia europaea* and

$$\phi_i = \begin{pmatrix} L_{\infty i} \\ k_i \end{pmatrix} = \begin{pmatrix} L_{\infty} + b_{1i} \\ k \end{pmatrix} \quad 5.5$$

for the *Leptopsammia pruvoti*

The **Errore. L'origine riferimento non è stata trovata.** does not change in either coral; in fact, the best model has a random effect for both parameters.

The assumption that the within-site errors have a zero expected value was then assessed by verifying that the residuals mean was close to zero. Because of the presence of within-site heteroscedasticity observed by plotting the residuals against the fitted values the model for both parameterizations and both corals had to be improved by considering a variance function modelling the variance structure of the within-site errors. The improvement of this new model was assessed by the decreasing of residuals and their normal distribution. Having found a model with any violation of the assumption about the within-site error, the presence of intra-site correlation was investigated by calculating the residuals empirical autocorrelation function (Box et al., 1994) at every lag, i.e. every year. The largest coral, i.e. the *Balanophyllia europaea*, showed an

autoregressive correlation of order 1 (AR(1)); conversely, in the smaller coral, i.e. the *Leptopsammia pruvoti*, no autocorrelation was observed.

The influence of covariates was then tested to see if the model could be further improved. We considered only the mean sea surface temperature gradient from 18°C and the mean solar radiation gradient from 190W/m². The mean sea current measurement was excluded because it is not very reliable for all the collection sites except Scilla where the current is constantly monitored because of his strength. The rationale for considering the gradient of the environmental covariates instead of the measured values is the small variations of the covariate values among the studied sites and the need to test only a linear influence on the parameters of the VBGF.

We considered a different approach for each of the two parameterizations.

In the traditional parameterization the effect of the covariates was attributed to both the VBGF parameters k and L_{∞} , but because of their correlation previously described in **Errore. L'origine riferimento non è stata trovata.**, we considered only one random effect, which is the one that produces a better model.

In the new parameterization the effect of the covariates was attributed only to parameter c , but we considered two random effects for both parameters L_{∞} and c . The results of the models considering the covariates are shown in the following chapters.

All the described analyses were performed using the R function `nlfme()`, (Bates and Chamber, 1992). The previous results obtained by the nonlinear method were inserted as starting points.

5.5.1 The method giving the best fit for the *Balanophyllia Europaea*

The best fit for *Balanophyllia europaea* involves the environmental covariates, a variance function modelling the variance structure and a correlation structure used to model dependence among observations. The estimated parameters of the model providing the best fit are shown in Table 5.1 and the goodness of fit indicators in following Table 5.2

Table 5.1 Best fit estimated parameters for the *Balanophyllia europaea* data in both parameterizations

	<i>Standard Parameterization</i>		<i>New Parameterization</i>		
<i>ith site VBGF parameters</i>	$L_{\infty i} = L_{\infty} + \alpha_1 R_i + \alpha_2 T_i + b_{1i}$		$L_{\infty i} = L_{\infty} + b_{1i}$		
	$k_i = k + \alpha_3 R_i + \alpha_4 T_i + b_{2i}$		$c_i = c + a_3 R_i + a_4 T_i + b_{2i}$		
<i>Fixed effects</i>	$L_{\infty} + \alpha_1 R_i + \alpha_2 T_i$		L_{∞}		
	$L_{\infty} = 18.1$ SE=1.2 p<0.0005		$L_{\infty} = 17.0$ SE=1.3 p<0.0005		
	$\alpha_1 = -0.08$ SE=0.09 p=0.05				
	$\alpha_2 = -2.74$ SE=1.90 p=0.09				
	$k + \alpha_3 R_i + \alpha_4 T_i$		$c + a_3 R_i + a_4 T_i$		
	$k = 0.161$ SE=0.018 p<0.0005		$c = -30.1$ SE=5.0 p<0.0005		
	$\alpha_3 = 0.00002$ SE=0.0020 p=0.79		$a_3 = -0.19$ SE=0.08 p=0.01		
$\alpha_4 = 0.04$ SE=0.03 p=0.14		$a_4 = 0.02$ SE=1.2 p=0.7			
<i>Random effects</i>		b_{1i}		b_{1i}	b_{2i}
	genova	0.006	genova	-5.4	1.3
	calafuria	-0.008	calafuria	-14.7	3.5
	elba	0.027	elba	14.0	-3.3
	palinuro	-0.020	palinuro	-5.5	1.3
	scilla	-0.001	scilla	3.4	-0.8
	pantelleria	0.007	pantelleria	8.2	-1.9
<i>Variance</i>	$\sigma^2(\mu_j, \Theta) = e^{2\mu_j^2} \mu_j^{2\alpha}$		$\sigma^2(\mu_j, \Theta) = (\sigma_j + \mu_j^2)^2$		
	$\vartheta_1 = 0.153$ $\vartheta_2 = .856$		$\vartheta_1 = 13864$ $\vartheta_2 = 3.5$		
<i>Correlation</i>	$\text{Corr}(e_{ij}, e_{ij'}) = \alpha^{1/2 \cdot j-j' }$ $\Gamma(\alpha) = \begin{bmatrix} 1 & \alpha^{1/2} & \alpha^{1/4} & \alpha^{1/8} \\ \alpha^{1/2} & 1 & \alpha^{1/4} & \alpha^{1/8} \\ \alpha^{1/4} & \alpha^{1/4} & 1 & \alpha^{1/8} \\ \alpha^{1/8} & \alpha^{1/8} & \alpha^{1/8} & 1 \end{bmatrix}$		$\text{Corr}(e_{ij}, e_{ij'}) = \alpha^{1/2 \cdot j-j' }$ $\Gamma(\alpha) = \begin{bmatrix} 1 & \alpha^{1/2} & \alpha^{1/4} & \alpha^{1/8} \\ \alpha^{1/2} & 1 & \alpha^{1/4} & \alpha^{1/8} \\ \alpha^{1/4} & \alpha^{1/4} & 1 & \alpha^{1/8} \\ \alpha^{1/8} & \alpha^{1/8} & \alpha^{1/8} & 1 \end{bmatrix}$		
	$\alpha = 0.41$		$\alpha = 0.38$		

Table 5.2 Model comparison criteria and residuals distributions for *Balanophyllia europaea* data

	<i>AIC</i>	<i>BIC</i>	<i>LogLik</i>	<i>GdL</i>	<i>Standardized Error</i>				
					<i>Minimum</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Maximum</i>
<i>Standard Parameterization</i>	758	796	-368	11	-2.36	-0.65	0.13	0.70	3.16
<i>New Parameterization</i>	753	791	-365	11	-2.50	-0.63	0.10	0.71	2.30

By the inspection of Table 5.1 it can be gathered that in the standard parameterization the ultimate length L_{∞} is considered to decrease linearly with solar radiation and sea surface temperature; conversely the growth rate k is considered to increase linearly with solar radiation and sea surface temperature, this might mean that in southern and warmer sites corals grow faster but reach a smaller final dimension.

For both parameters the slope of solar radiation is smaller than the slope of sea surface temperature but with a narrower confidence interval. The only random effect concerns growth rate k . The variance heterogeneity is corrected by a function which is a compound of an exponential increasing function and a power with negative exponent function. This function in fact decreases the influence of the extreme fitted values and increases that of central ones when contemplating the fact that the measures of young and very old corals are generally less reliable because of the difficulty in counting the growth rings. The correlation structure is autoregressive AR(1). The goodness of fit of this model is confirmed by the very close to zero residuals quantiles presented in Table 5.2.

In the new parameterization the ultimate length L_{∞} is independent of solar radiation and sea surface temperature as observed by Kleiber (1947) and confirmed by Kooijman (2008); conversely growth coefficient c is considered to decrease with solar radiation and increase with the sea surface temperature, so southern but colder sites should have more slowly growing corals than those of northern but warmer sites. The slope of solar radiation is larger and has a narrower confidence interval than the slope of sea surface temperature. The random effects are quite large: the similar behaviour of the two islands, i.e. Elba and Pantelleria and the site with certainly a much stronger current, i.e. Scilla can be noticed. This might suggest that by adding further covariates, such as precisely measured sea current, the model might be further improved. The two random effects are correlated, suggesting a more parsimonious model with only one, but, when we estimated it, the results were definitely worse. The variance heterogeneity is corrected by decreasing the influence of only the larger fitted values; in fact, the bad estimation of young corals seems to be less emphasized than in the traditional parameterization. The function having produced the better model is a mixture of a constant function and a power with positive exponent function. Also in this case the correlation function is autoregressive AR(1).

Table 5.2, showing the model comparison criteria and the residuals distributions of both parameterizations, confirms a general good performance of both models; it might be underlined however that the new parameterization has slightly smaller AIC and BIC and residuals slightly closer to zero.

The residuals plotted vs the fitted values in Figure 5.27 show that some extreme estimated values typical of the more common parameterization are absent in the new one; the residuals of both parameterizations are symmetrical and centred on zero.

Figure 5.27 Residuals using the covariates with the two parameterizations for *Balanophyllia europaea* data

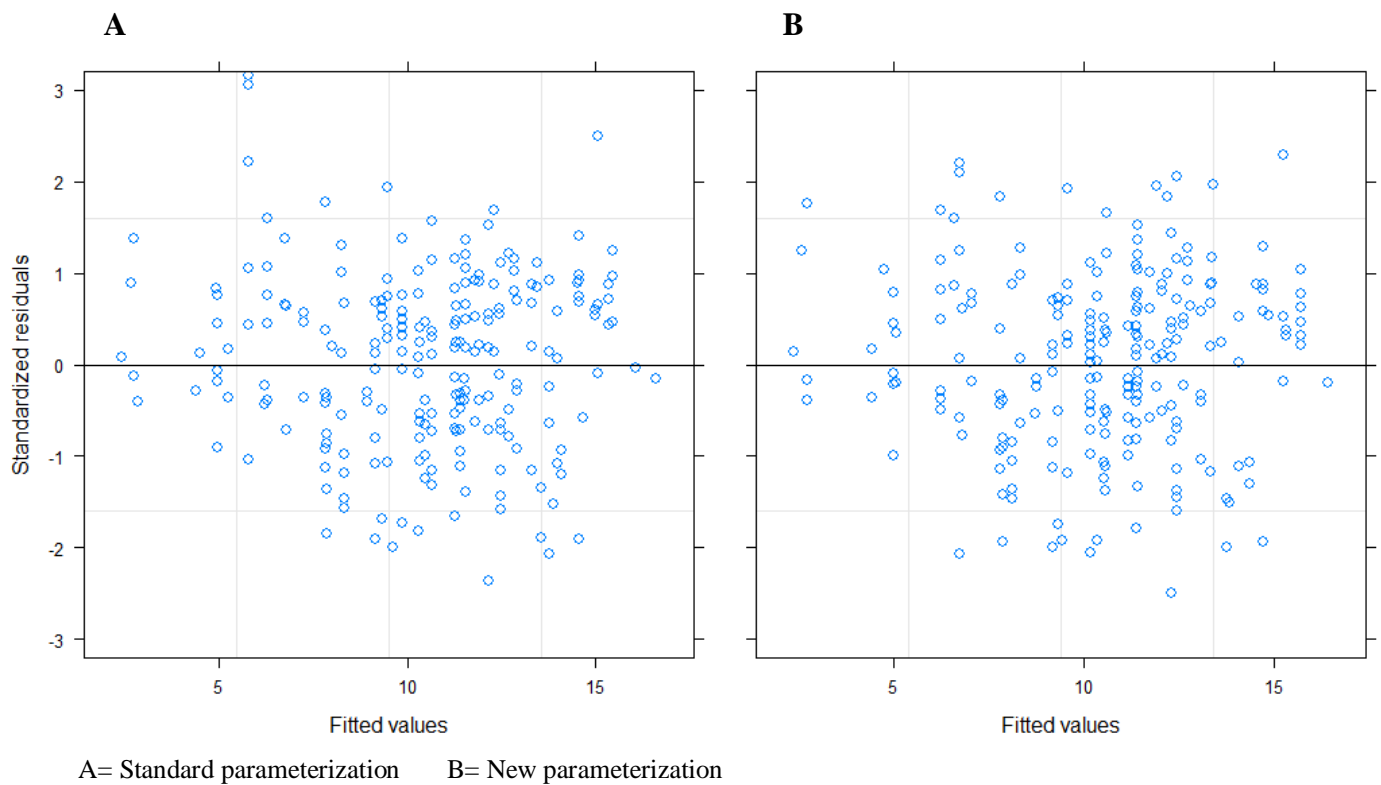
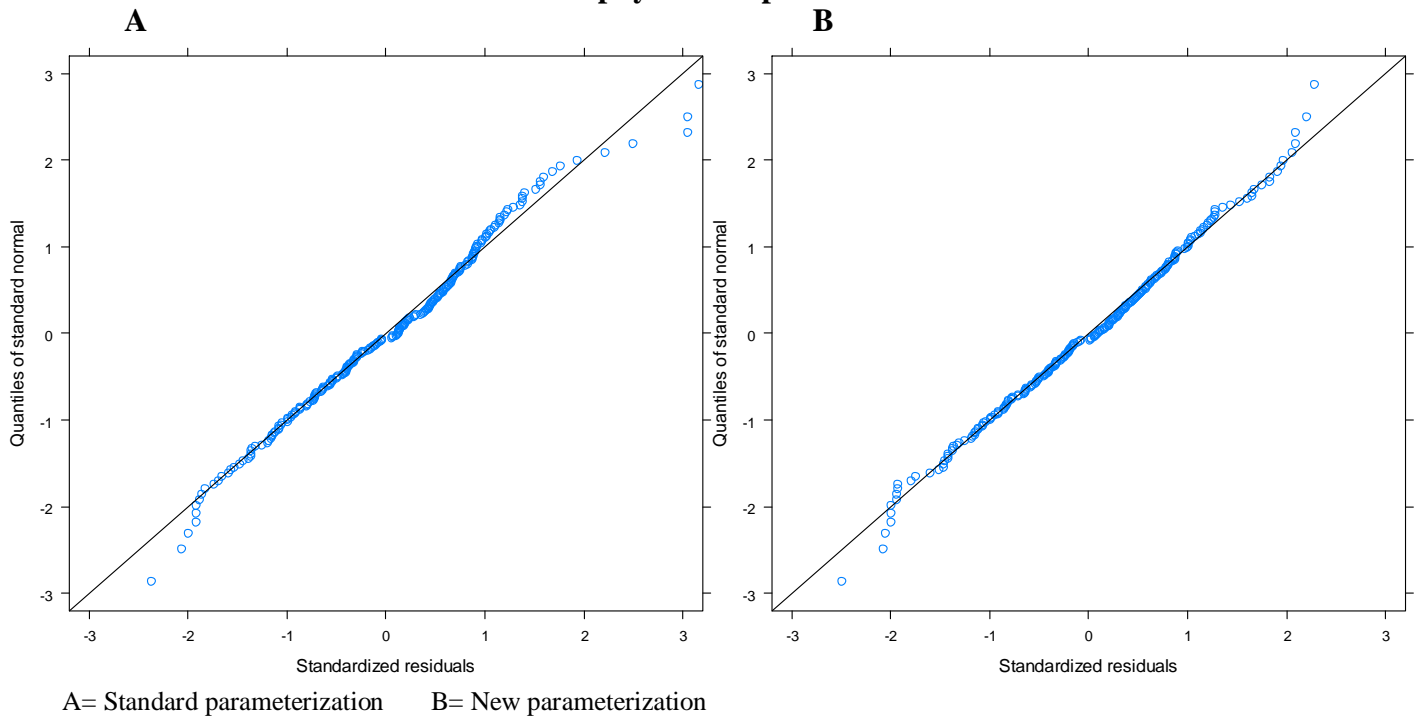


Figure 5.28 shows that the distribution of the residuals of both parameterizations has deviations from normality for extreme values; less strong for the new parameterization.

Figure 5.28 Normal plot of residuals using the covariates with the two parameterizations for *Balanophyllia europaea* data



The observed vs fitted data plot shown in Figure 5.29 highlights, for both parameterizations, the general underestimation of the larger corals, likely due to the lack of consideration of the sea current which can be an important factor of growth.

Figure 5.29 Observed vs fitted plot using the covariates with the two parameterizations for *Balanophyllia* data

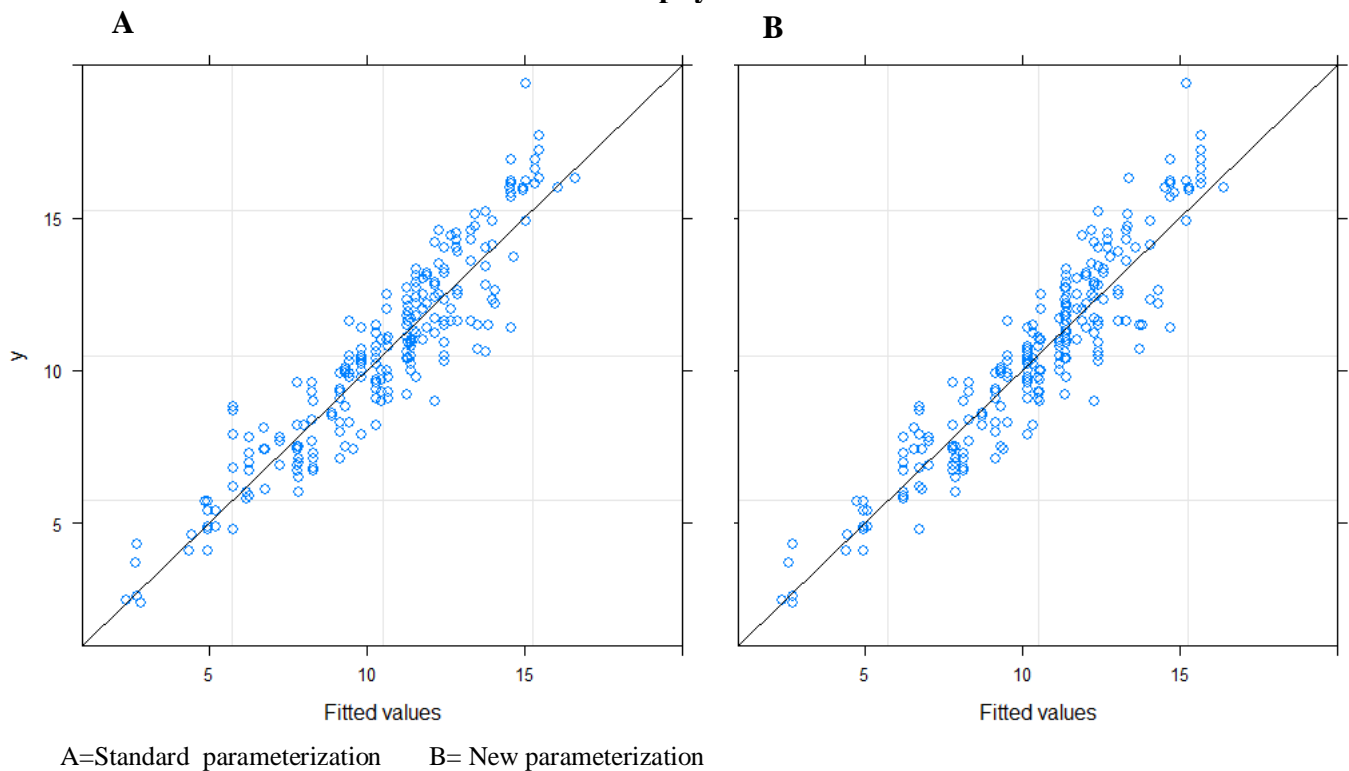
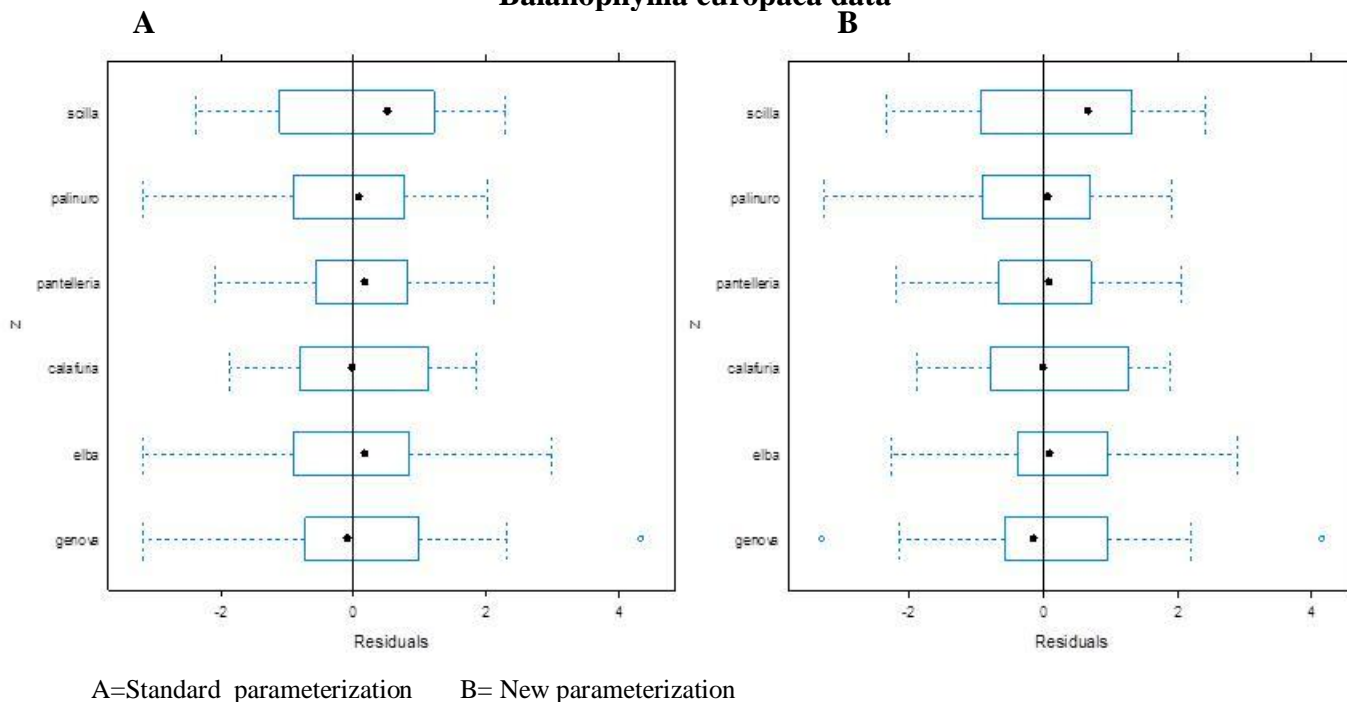


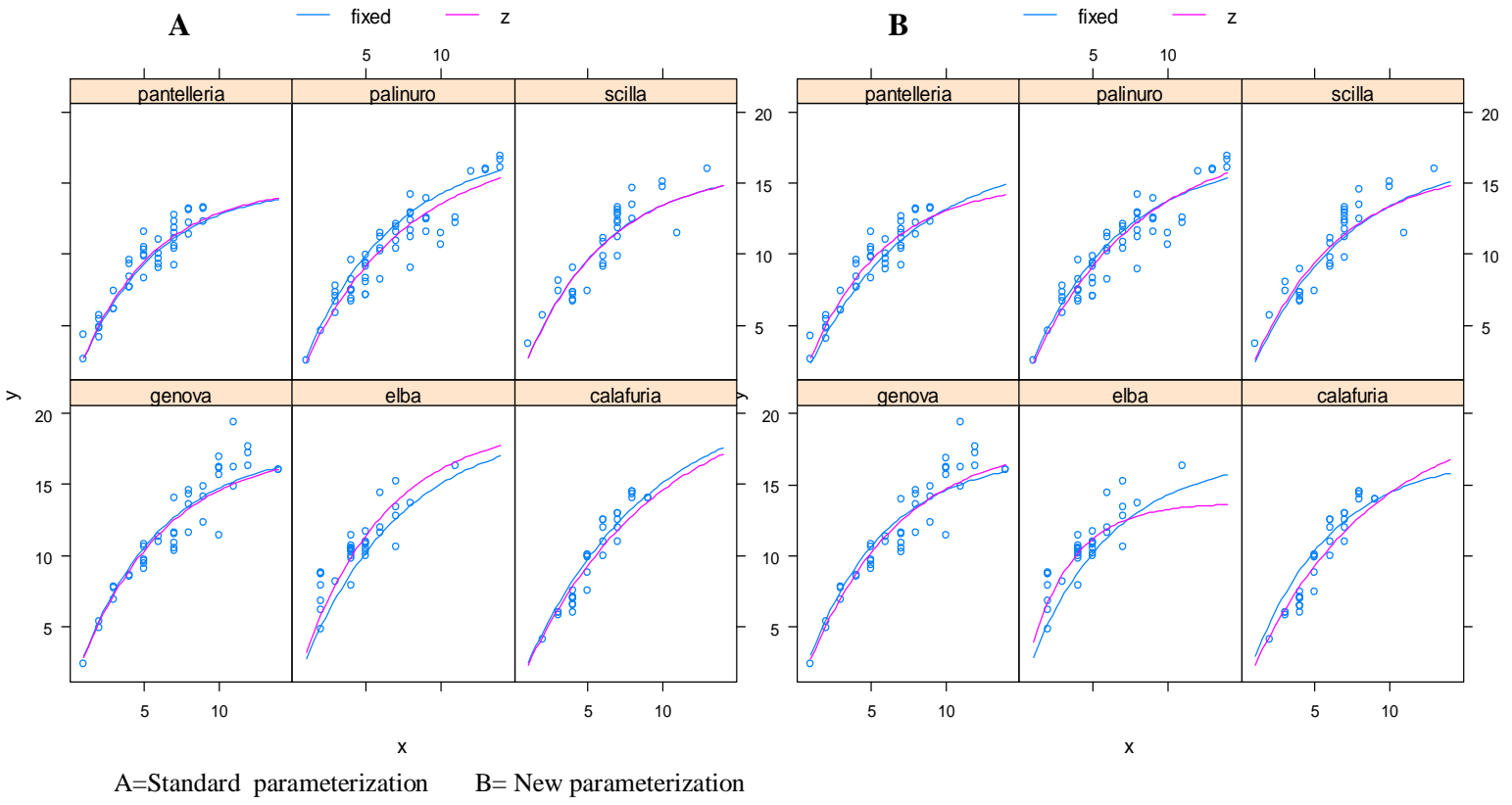
Figure 5.30 shows the distribution of the residuals along sites and confirms that both parameterizations perform poorly only at Scilla which is the site with the largest sea current.

Figure 5.30 Boxplot of residuals by sites using the covariates with the two parameterizations for the *Balanophyllia europaea* data



Finally, Figure 5.31 shows the predicted curve with the intra-site random effect accommodations: the biggest difference between the two models is in the prediction of the coral growth in Elba; there are remarkable differences also for Calafuria and Scilla. The reason for these differences might be once more the lack of information about the sea current which might be the reason for the wide random effects of the new parameterization. Parameter c , introduced to capture the effect of covariates, is in fact more sensitive to poor environmental information. The standard parameterization, distributing the covariate influences between the two parameters, can manage this lack of information.

Figure 5.31 Observed data, population prediction (fixed) and within-site predictions (z) with the two parameterizations for the *Balanophyllia europaea* data



5.5.2 The method giving the best fit for *Leptopsammia Pruvoti*

The model giving the best fit takes into account the environmental covariates and a function modelling the variance structure and is shown in Table 5.3; its goodness of fit is presented in Table 5.4.

Table 5.3 Best fit estimated parameters for the *Leptopsammia pruvoti* data in both parameterizations

	<i>Standard Parameterization</i>		<i>New Parameterization</i>		
<i>ith site VBGF parameters</i>	$L_{\infty i} = L_{\infty} + \alpha_1 R_i + \alpha_2 T_i + b_{1i}$		$L_{\infty i} = L_{\infty} + b_{1i}$		
	$k_i = k + \alpha_3 R_i + \alpha_4 T_i + b_{2i}$		$c_i = c + a_3 R_i + a_4 T_i + b_{2i}$		
<i>Fixed effects</i>	$L_{\infty} + \alpha_1 R_i + \alpha_2 T_i$		L_{∞}		
	$L_{\infty} = 10.5$ SE=0.9 p<0.0005		$L_{\infty} = 10.1$ SE=1.1 p<0.0005		
	$\alpha_1 = 0.15$ SE=0.08 p=0.015				
	$\alpha_2 = -1.03$ SE=1.30 p=0.7				
	$k + \alpha_3 R_i + \alpha_4 T_i$		$c + a_3 R_i + a_4 T_i$		
	$k = 0.115$ SE=0.01 p<0.0005		$c = -21.5$ SE=3.6 p<0.0005		
	$\alpha_3 = 0.025$ SE=0.027 p=0.3		$a_3 = -0.08$ SE=0.03 p=0.08		
$\alpha_4 = -0.004$ SE=0.0018 p=0.03		$a_4 = -0.04$ SE=0.4 p=0.6			
<i>Random effects</i>		b_{1i}		b_{1i}	b_{2i}
	genova	0.009	genova	-1.7	5.7
	calafuria	0.010	calafuria	-0.5	1.8
	elba	0.048	elba	0.5	-1.7
	palinuro	-0.313	palinuro	-0.7	2.5
	scilla	0.454	scilla	3.2	-10.8
	pantelleria	-0.209	pantelleria	-0.7	2.4
<i>Variance</i>	$g^2(\mu_j, \theta) = (\theta_1 + \mu_j^{\theta_2})^2 \mu_j^{2\theta_3}$		$g^2(\mu_j, \theta) = (\theta_1 + \mu_j^{\theta_2})^2 \mu_j^{2\theta_3}$		
	$\theta_1 = 10.4$ $\theta_2 = 4.6$ $\theta_3 = -3.4$		$\theta_1 = 10.1$ $\theta_2 = 5.1$ $\theta_3 = -4.0$		

Table 5.4 Model comparison criteria and residuals distributions for *Leptopsammia pruvoti* data

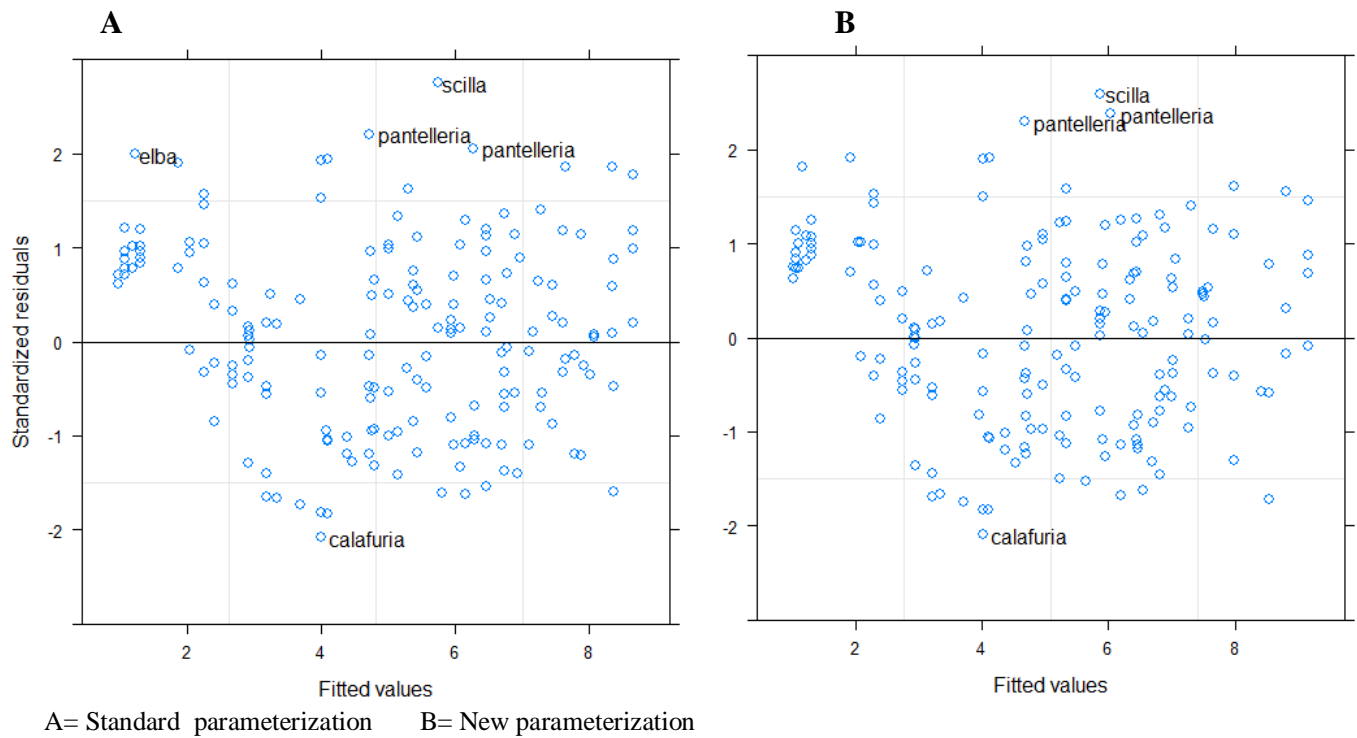
	<i>AIC</i>	<i>BIC</i>	<i>LogLik</i>	<i>GdL</i>	<i>Standardized Error</i>				
					<i>Minimum</i>	<i>Q1</i>	<i>Median</i>	<i>Q3</i>	<i>Maximum</i>
<i>Standard Parameterization</i>	531	566	-254	11	-2.01	-0.64	0.10	0.86	2.75
<i>New Parameterization</i>	527	562	-252	11	-2.01	-0.62	0.13	0.81	2.60

In accordance with the results presented in Table 5.3, in the standard parameterization the ultimate length L_{∞} decreases linearly with solar radiation and increases linearly with the sea surface temperature; conversely

growth rate k decreases linearly with both solar radiation and sea surface temperature, so that southern and warmer sites have corals growing slower but reaching a larger final dimension if the site is very warm. For both parameters the slope of solar radiation is smaller but with a narrower confidence interval than the slope of sea surface temperature. The only random effect is due to the ultimate length L_{∞} . For this coral the variance heterogeneity also needs to be corrected by decreasing the influence of the extreme fitted values and increasing that of central ones: in this case because of the smaller dimension of the coral, the age of only very young corals, mainly one year-olds, is strongly underestimated and those of old corals is generally poorly estimated because of the difficulty in counting the growth rings. The most suitable function for modelling the variance structure is thus a compound of an exponential increasing function, a constant function and a power with negative exponent function. In fact, no autocorrelation was found, because of the very small dimensions; the yearly fluctuation conditions should not have any influence on the coral growth rings. In the new parameterization the ultimate length L_{∞} is, obviously, independent of solar radiation and sea surface temperature; growth coefficient c is considered to decrease with solar radiation and sea surface temperature, so southern and warmer sites should have slower coral growth than that of corals in northern and cold sites. As for the *Balanophyllia europaea*, the slope of solar radiation gradient is larger and has narrower confidence intervals than the slope of the sea surface temperature gradient. The random effects confirm the common behaviour only for Elba and Scilla. Variance heterogeneity is corrected by a function similar to that used for normal parameterization. Table 5.4 shows the model comparison criteria and the residuals distributions of both parameterizations: the performance is similar even if the new parameterization has a smaller AIC and BIC and also the residuals range is smaller.

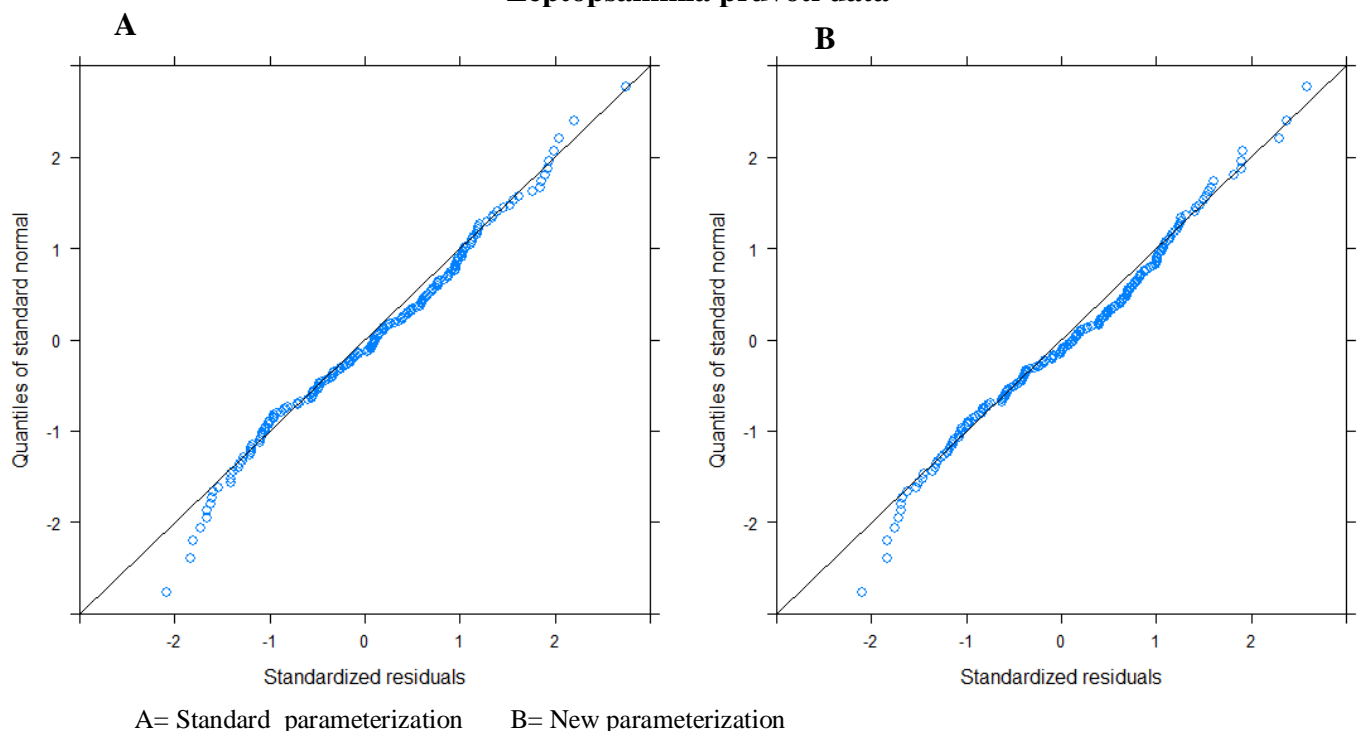
In addition, Figure 5.32 shows that some outliers of the normal parameterization are reduced in the new one. The residuals of both parameterizations are centred on zero, but despite the correction they are still not symmetrical for very young corals.

Figure 5.32 Residuals using the covariates with the two parameterizations for *Leptopsammia pruvoti* data



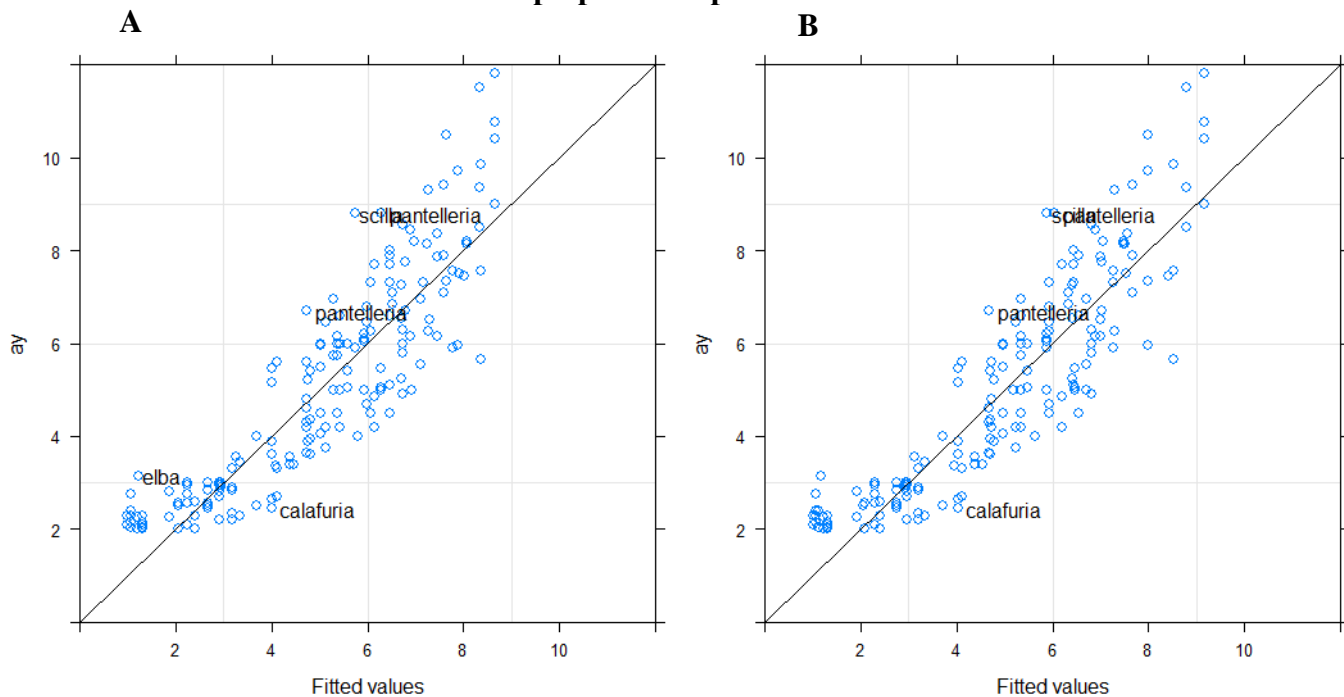
The normality of residuals distributions was investigated by plotting the quantiles of a standard normal distribution against those of residuals shown in Figure 5.3; both parameterizations have evident deviations from normality for small values and the distribution of the new parameterization seems to be slightly closer to normal than the standard parameterization.

Figure 5.33 Normal plot of residuals using the covariates with the two parameterizations for *Leptopsammia pruvoti* data



The fitted vs measured plot shown in Figure 5.34 highlights for both parameterizations the general underestimation of very small and very big corals; the new parameterization seems to slightly limit this underestimation.

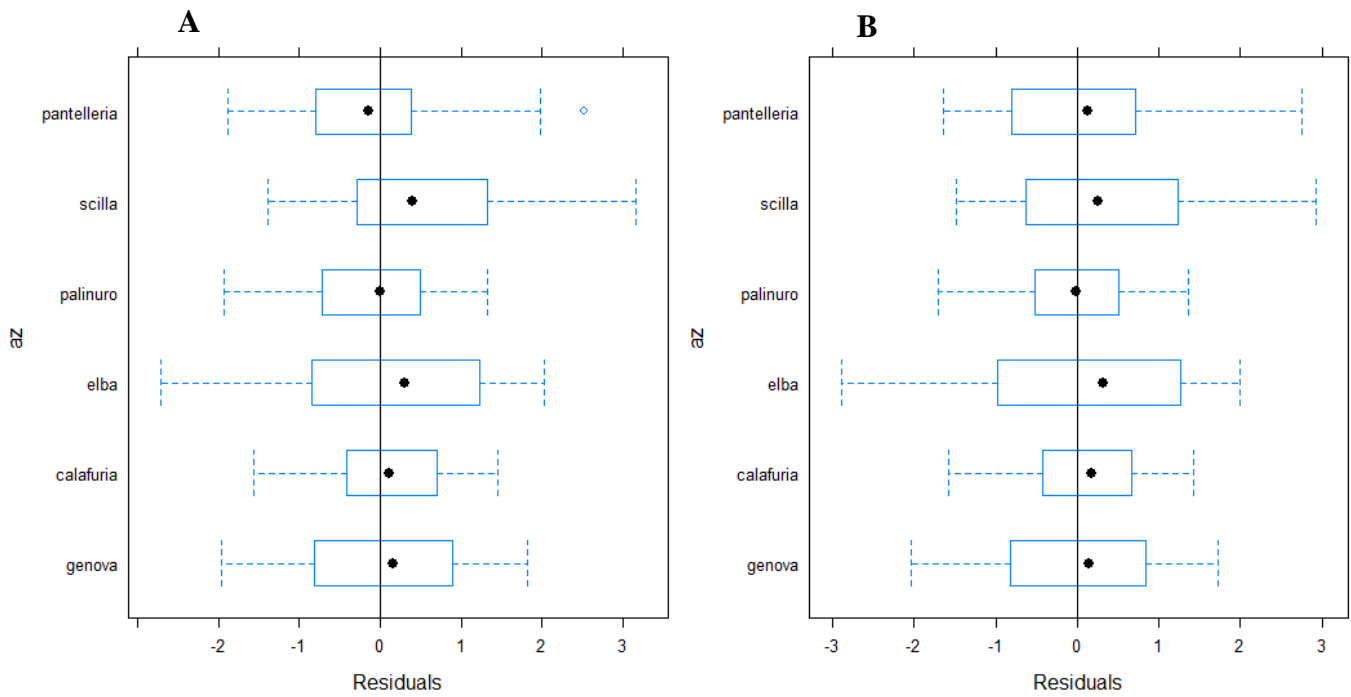
Figure 5.34 Observed vs fitted plot using the covariates with the two parameterizations for *Leptosammia pruvoti* data



A= Standard parameterization B= New parameterization

Figure 5.35 describes the distribution of the residuals according to the sites and shows that both the parameterizations have the worst performance with Scilla and Elba even if the new parameterization seems better performing with Scilla.

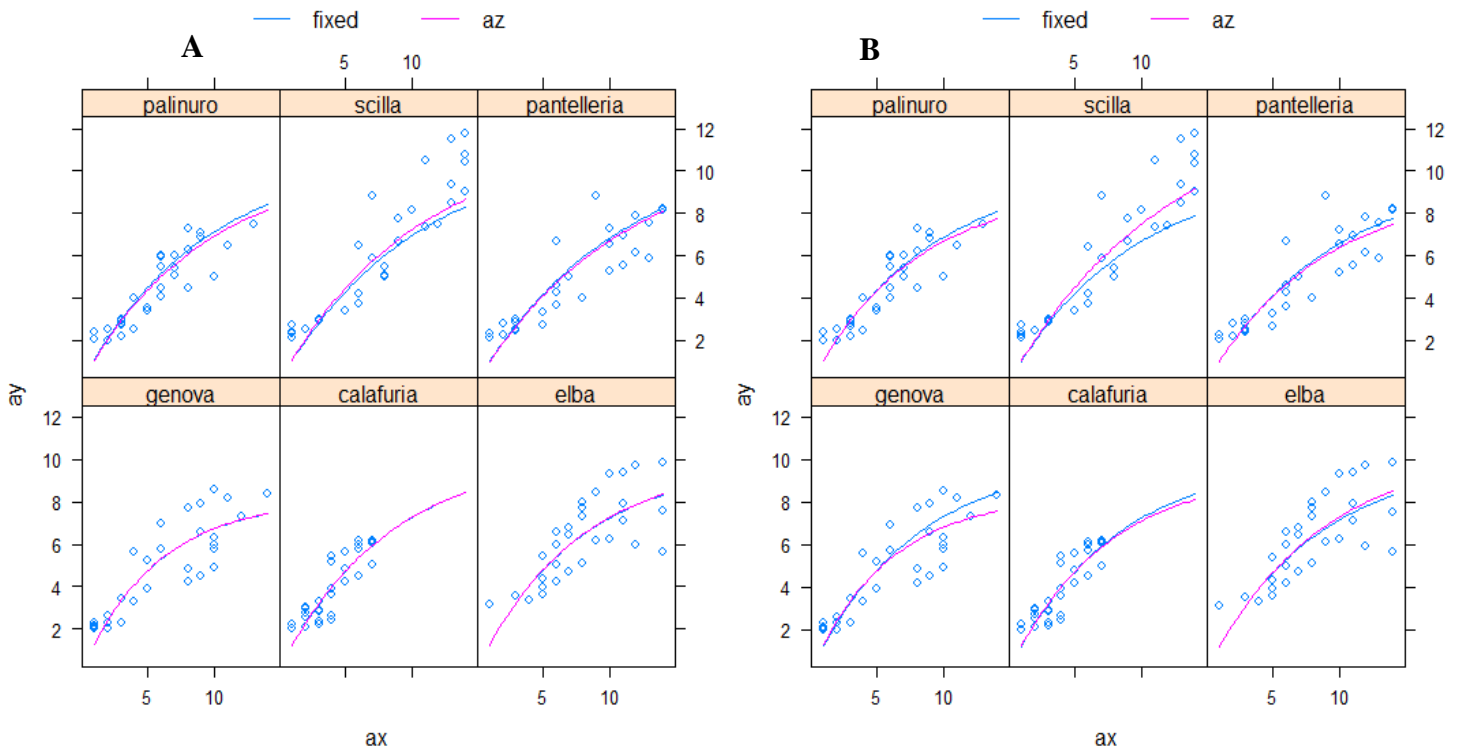
Figure 5.35 Boxplot of residuals by sites using the covariates with the two parameterizations for the *Leptosammia pruvoti* data



A= Standard parameterization B= New parameterization

Figure 5.36 shows the predicted curve with the intra-site random effect accommodations: the biggest difference between the two models is in the prediction of the coral growth in Scilla confirming the better performance of the new parameterization.

Figure 5.36 Observed data, population prediction (fixed) and within-site predictions (z) with the two parameterizations for the *Leptosammia pruvoti* data



A= Standard parameterization B= New parameterization

6 Conclusions

The motivating problem of this thesis concerns the estimation of the growth curve of solitary corals that follow the nonlinear Von Bertalanffy Growth Function (VBGF). The most common parameterization of the VBGF for the corals is based on two parameters: the ultimate length L_∞ and the growth rate k . One of the aims of this thesis is to find a more reliable method for estimating these parameters than those currently used and able to capture the influence of environmental covariates. The main problem of the methods currently used is the forced linearization of VBGF. Therefore a preliminary step for improving the model was the use of a nonlinear least square method to estimate the two parameters. A further decisive improvement was expected by introducing the hierarchical nonlinear model which has the appealing features of taking into account the influence of the collection sites and the possible intra-site measurement correlation and variance heterogeneity. Further on, the hierarchical nonlinear model can manage the influence of environmental factors and all the reliable information that might influence coral growth. The models considered were tested on two databases relating two different solitary corals i.e. *Balanophyllia europaea* and *Leptopsammia pruvoti*, collected in six different sites having different environmental conditions. The nonlinear least square model introduces a substantial improvement in the results compared to those obtained by the linearized methods used by biologists. The use of nonlinear hierarchical models including environmental covariates gives a further marked improvement in the fit suggesting it is a suitable method to study the growth of solitary corals. Nevertheless, the theory of the energy balance in growth ascertains the linear correlation of the two parameters and the independence of the ultimate length L_∞ from the influence of environmental covariates, so a new parameterization was introduced based on the ultimate length and a parameter c which explicitly describes the part of the growth accountable to site-specific conditions such as environmental factors. In this thesis we explored the possibility of estimating these parameters characterizing the VBGF new parameterization via nonlinear Least squares regression. The second step was to improve the model considering a nonlinear hierarchical model; a further improvement was reached by introducing the possibility of managing the intra-site variance structure and autocorrelation into the model. Finally, the presence of environmental covariates was introduced into the model. Also in this case the nonlinear least squared regression highlighted a general improvement with respect to the traditional methods, but the most

suitable models were obtained considering the hierarchical nonlinear approach involving the presence of environmental covariates. The results giving the best fit of each of the two parameterizations were similar, even if a very slight improvement was observed in the new one. This one is nevertheless, more suitable from a theoretical point of view when considering environmental covariates. In fact, their influence is correctly ascribed only to the parameter sensitive to them, i.e. parameter c . The advantages of this new parameterization are considerable only if the measurement of environmental covariates is accurate. The lack of accuracy in measuring covariates might cancel the advantage of the new parameterization. The results obtained show, for both parameterizations, a decisive improvement when compared to the traditional methods; nevertheless, other models might be examined introducing, for example, a Bayesian approach.

Bibliography

- Arendt, J. «Ecological correlates of body size in relation to cell size and cell number: patterns in flies, fish, fruits and foliage.» *Biological Reviews* , 2007: 82:241–256.
- Arino, A. & Reuer, J. « Designing and Renegotiating Strategic Alliance Contracts.» *Academy of Management Executive*, 2004: 3: 37-48.
- Bartlett, AA. «Arithmetic of growth: methods of calculation.» *Population and Environment*, 1993: 14(4): 359–387.
- Bartlett, AA. «Arithmetic of growth: methods of calculation II.» *Population and Environment*, 1999: 20(3): 215–246.
- Bartlett, AA. «The essential exponential! For the future of our planet.» *Center for Science, Mathematics and Computer Education, University of Nebraska-Lincoln*, 2004: 302.
- Basso, N.G. and A.I. Kehr. «Postmetamorphic growth and population structure of the frog *Leptodactylus latinasus* (Anura: Leptodactylidae).» *Stud. Neotrop. Fauna Envir.*, 1991: 26: 39-44.
- Bates, D. M. and D. G. Watts. *Nonlinear Regression Analysis and Its Applications*. New York: Wiley, 1988.
- Bates, D. M., J. M. Chambers. «Nonlinear Models.» In *Statistical Models in S*, di J. M. Chambers. Cole, Pacific Grove, California: Wadsworth & Brooks, 1992.
- Beverton, R. J. H., e S. J. Holt. *On the Dynamics of Exploited Fish Populations*. Fishery Investigations Series II Volume XIX, Ministry of Agriculture, Fisheries and Food, 1957.
- Brett, J.R.. «Environmental factors and growth.» In *Fish Physiology, VIII.* , di W.S., Randall, D.J., Brett, J.R. Hoar, 599–667. London: Academic Press, 1979.
- Burnaby, T. P. «Growth-invariant discriminant function and generalized distances.» *Biometrics*, 1966: 22:96-110 .
- Cailliet G. M., Smith W. D., Mollet H. F., Goldman K. J. «Age and growth studies of chondrichthyan fishes: the need for consistency in terminology, verification, validation, and growth function fitting.» *Environmental Biology of Fishes* , 2006: 77:211-228.
- Cairns, S.D. «Deep-water corals.» *Sea Frontiers*, 1977: 23(2): 84-89.
- Calder, W.A. *Size, function and life history*. Cambridge, Massachusetts: Harvard University Press, 1984.
- Clarke, L. M., Thorrold, S. R. & Conover, D. O. «Population differences in otolith chemistry have a genetic basis in *Menidia menidia*.» *Canadian Journal of Fisheries and Aquatic Sciences*, 2011: 68:105–114.
- Conover, D.O. and Munch, S.B. «Sustaining fisheries yields over evolutionary time scales.» *Science*, 2002: 297:94-96.
- Davidian M., Giltinan D.M. *Nonlinear models for repeated measurement data*. London: Chapman & Hall, 1995.
- DeAngelis, D. L., and Gross L. J. *Individual-Based Models and Approaches in Ecology*. New York: Routledge, Chapman and Hall, 1992.

- Epstein, P.R and Selber, J. *Oil: A Life Cycle Analysis of its Health and Environmental Impacts*. Boston: The Center for Health and the Global Environment. Harvard Medical School, 2002.
- Ernande, B., Dieckmann, U. and Heino, M. «Adaptive changes in harvested populations: plasticity and evolution of age and size at maturation.» *Proceedings of the Royal Society of London Series B-Biological Sciences*, 2004: 271:415-423.
- Fabens, A. J. «Properties and fitting of the Von Bertalanffy growth curve.» *Growth*, 1965: 29: 265-289.
- Ferreira, B. P. & Russ, G. R. «Age validation and estimation of growth-rate of the coral trout, *Plectropomus leopardus*, (Lacepede 1802) from Lizard Island, Northern Great Barrier Reef.» *Fish. Bull*, 1994 : 92:46-57.
- Fox, L. R. «Cannibalism in natural populations.» *Annual Review of Ecology and Systematics* , 1975: 6:87-106.
- Fraser, D.F., J.F. Gilliam, M.J. Daley, A.N. Le & G.T. Skalski. «Explaining leptokurtic movement distributions: Intrapopulation variation in boldness and exploration.» *Am. Nat.* , 2001: 158:124-135.
- Fraser, S. E., Keynes, R.J. and Lumsden, A. «Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. .» *Nature* , 1990: 344:431-435.
- Fujiwara, M., Kendall B. E., and Nisbet R. M. « Growth autocorrelation and animal size variation.» *Ecology*, 2004: Letters 7:106-113.
- Galland, J. A. and S. J. Holt. «Estimation of growth parameters for data at unequal time intervals.» *Journ. Cons. CIEM*, 1959: 25 (1): 47-49.
- Gallant, A. R. *Nonlinear Statistical Models*. New York: Wiley, 1987.
- Galluci, V. F. & Quinn, T. J. « Reparameterizing, fitting, and testing a simple growth-model.» *Trans. Am. Fish. Soc.* , 1979: 108:14-25.
- Genning, C., Chinchilli, V. M. and Carter, W. H. «Response surface analysis with correlated data: A nonlinear model approach.» *Journal of the American Statistical Association*, 1989: 84: 805-809.
- Goffredo S., Caroselli E., Mattioli G., Zaccanti F. «Growth and population dynamic model for the non-zooxanthellate temperate solitary coral *Leptopsammia pruvoti* (Scleractinia, Dendrophylliidae).» 2010: 157: 2603-2612 .
- Goffredo S., Chadwick-Furman N. E. «Comparative demography of mushroom corals (Scleractinia, Fungiidae) at Eilat, northern Red Sea.» *Marine Biology*, 2003: 142: 411-418.
- Goffredo S., Chadwick-Furman N. E. «Abundance and distribution of mushroom corals (Scleractinia, Fungiidae) on a coral reef at Eilat, northern Red Sea. .» *Bulletin of Marine Science*, 2000: 66: 241-254.
- Goffredo S., Lasker H. R. «An adaptive management approach to an octocoral fishery based on the Beverton-Holt model.» *Coral Reefs*, 2008: 27: 751-761.

- 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.*, 2007: 152: 351–361.
- Gompertz, B. «On the nature of the function expressive of the law of human mortality.» *Philosophical Transactions of the Royal Society*, 1825: 115:513-585.
- Grigg R.W., Polovina JJ, Atkinson MJ. «Model of a coral reef ecosystem III. Resource limitation, community regulation, fisheries yield and resource management.» *Coral Reefs*, 1984: 3: 23-27.
- H., Draper N R & Smith. *Applied regression analysis*. New York: Wiley, 1998.
- Harper, J.L. *Population Biology of Plants*. 1st. London: Academy Press, 1977.
- Harrell Yee S., Barron M.G. «Predicting coral bleaching in response to environmental stressors using 8 years of global-scale data.» *Environmental Monitoring and Assessment* (Environmental Monitoring and Assessment), 2010: Volume 161, Issue 1-4, pp 423-438.
- Helser, T. E., H. Lai. «A Bayesian hierarchical meta-analysis of fish growth: with an example for North American largemouth bass, *Micropterus salmoides*.» *Ecological Modelling*, 2004: 178:399-416.
- Höök, M., Aleklett, K. «Trends in U.S. recoverable coal supply estimates and future production outlooks.» *Natural Resources Research*, 2011: 3: 189-208.
- Hubbert, M.K. «Nuclear energy and the fossil fuels.» *Spring Meeting of the Southern District, American Petroleum Institute*, 1956: 7–9.
- Immland AK, Nilsen T, Folkvord A. «Stochastic simulation of size-variation in turbot: possible causes analysed with an individual based model.» *J Fish Biol*, 1998: 53:237–258.
- Immland, A.K., Jonsdottir, ODB. «Is there a genetic base for growth in cod? .» *Fish Fish.*, 2002: 3:36 – 52.
- Janoschek, A. «Das reaktionskinetische Grundgesetz und seine Beziehungen zum Wachstum- und Ertragsgesetz.» *Statist. Vjschr.*, 1957: 10: 25-37.
- Kendall, B. E., and Fox, G. A. «Variation among individuals and reduced demographic stochasticity.» *Conservation Biology*, 2002: 16:109–116.
- Kleiber, M. «Body size and metabolic rate. .» *Physiol. Rev.*, 1947: 27:511.
- Kooijman, S. A. L. M. & Troost, T. A. «Quantitative steps in the evolution of metabolic organisation as specified by the dynamic energy budget theory .» *Biol.Rev.*, 2007 : 82, 1–30.
- Kooijman, S. A. L. M., Sousa, T., Pecquerie, L., van der Meer, J. & Jager, T. « From food-dependent statistics to metabolic parameters, a practical guide to the use of dynamic energy budget theory .» *Biol. Rev.*, 2008: 83: 533–552.
- Kooijman, S. A. L. M. *Dynamic energy and mass budgets in biological systems*,. Cambridge: 2nd rev. ed. Cambridge Univ. Press, 2000.
- Lv, Q. and J.W. Pitchford. «Stochastic von bertlanffy models, with applications to fish recruitment.» *J. Theor. Biol.*, 2007: 244: 640-655.

- McClanahan, T. R., J. E. Cinner, N. A. J. Graham, T. M. Daw, J. Maina, S. M. «Identifying reefs of hope and hopeful actions: Contextualizing environmental, ecological, and social parameters to effectively respond to climate change.» *Conservation Biology*, 2009: DOI: 10.1111/j.1523-1739.2008.01154.x.
- MOREAU, J. «Mathematical and biological expression of growth in fishes: recent trends and further developments.» In *The age and growth of fish*, di R.C. SUMMERFEIT e G.E. HALL, 81-113. The Iowa State Univ. Press, 1987.
- Peters, R.H. *The ecological implication of body size*. Cambridge: Cambridge University Press, 1983.
- Pfister, C.A., , Stevens F.R. «The genesis of size variability in plants and animals .» *Ecology* , 2002: 83: 59-72.
- Pilling, G. M., Kirkwood, G. P., and Walker, S. G. «An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters.» *Canadian Journal of Fisheries and Aquatic Sciences*, 2002: 59: 424–432.
- Pinheiro, J. C. and D. M. Bates. *Mixed-Effects Models in S and S-Plus*. New York: Springer-Verlag, 2000.
- Pütter, A. «Studien über physiologische ähnlichkeit. VI. Wachstumsähnlichkeiten.» *Pflüg. Arch. ges. Physiol.*, 1920: 180:298–340.
- Radetzki, M. *A Handbook of Primary Commodities in the Global Economy*. Cambridge: Cambridge University Press, 2007.
- Rao, C. R. «The theory of least squares when the parameters are stochastic and its application to the analysis of growth curves.» *Biometrika*, 1965: 52: 447-458.
- Richards, F.J. «A flexible growth curve for empirical use.» *Journal of Experimental Botany*, 1959: 10(2), 290–301.
- Ricker, E. W. «Handbook of computations for biological statistics of fish populations.» *Fish. Res. Board Can. Bull.*, 1958: 119-300.
- Ricker, W.E. . «Growth rates and models ,» In *Fish physiology Bioenergetics and growth Vol 8, chapter 11.*, di D.J. Randall & J.R. Brett W.S. Hoar, 678-743. New York: Academic Press Inc, 1979.
- Rilling, G. C. and E. D. Houde. « Regional and temporal variability in growth and mortality of bay anchovy, *Anchoa mitchilli*, larvae in Chesapeake Bay.» *Fish. Bull., U.S.* , 1999: 97:555-569.
- Ross, J. L., Stevens, T. M. & Vaughan, D. S. « Age, growth, mortality, and reproductive-biology of red drums in North Carolina waters .» *Trans. Am. Fish. Soc.* , 1995: 124:37–54.
- Ross, S. W., F. C. Rohde, and D.G. Lindquist. «Endangered, threatened, and rare fauna of North Carolina. Part II. A re-evaluation of the marine and estuarine fishes.» *Occasional Papers of the North Carolina Biological Survey*, 1988: 7- 20 .
- Roy, Richard F. Potthoff and S. N. «A Generalized Multivariate Analysis of Variance Model Useful Especially for Growth Curve Problems.» *Biometrika*, 1964: 51: 313-326.
- Sainsbury, K.J. « Effect of individual variability on the von Bertalanffy growth equation.» *Can. J. Fish. Aquat. Sci.* , 1980: 3(2): 241-7.

- Sakamoto, Y., Ishiguro, M. and Kitagawa, G. *Akaike Information Criterion Statistics*. Dordrecht: D. Reidel, 1986.
- Schaalje, G. B., McBride, J. B., and Fellingham, G. W. «Approximations to Distributions of Test Statistics in Complex Mixed Linear Models Using SAS Proc MIXED.» *Proceedings of the SAS Users Group International 26th Annual Conference*. 2001. Paper 262-26.
- Schwartz, C. C. & Hundertmark, K. J. «Reproductive characteristics of Alaskan moose.» *J. Wildl. Manage*, 1993 : 57:454–468.
- Schwarz, G. «Estimating the Dimension of a Model.» *Annals of Statistics*, 1978: 6: 461-464.
- Seber, G. A. F. and Wild, C. J. *Nonlinear Regression*. New York: Wiley, 1989.
- Simon, J.L. *The Ultimate Resource*. Princeton: Princeton University Press, 1981.
- Solow, Robert M. «The Economics of Resources or the Resources of Economics.» *The American Economic Review*, 1974: 1-14.
- Soto, C. J., John, O. P., Gosling, S. D., and Potter, J. «Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample.» *J. Pers. Soc. Psychol*, 2011: 100, 330–348.
- Sparre, P. and S.C. Venema. «Introduction to tropical fish stock assessment .» *part I- Manual. FAO Fish. Tech. Pap.* , 1992: 306/1:1-376.
- Stolarski, J. , Meibom A., Przenioslo R. and Mazur M. «A Cretaceous scleractinian coral with a calcitic skeleton.» *Science*, 2007: 318:92-94.
- Strum, S.C. «Weight and age in wild olive baboons.» *Am J Primatol* , 1991: 25:219–237.
- Sumpter, J.P. «Control of growth of rainbow trout (*Oncorhynchus mykiss*).» *Aquaculture* , 1992: 100, 299–320.
- Uchmanski, J. «Individual variability and population regulation: an individual-based model.» *Oikos* , 2000: 90:539–548.
- Verhulst, P.F. «Notice sur la loi que la population poursuit dans son accroissement.» *Correspondance mathématique et physique*, 1838: 10: 113–121.
- Von, Bertalanffy L. «A quantitative theory of organic growth (inquiries on growth laws. II).» *Hum. Biol.*, 1938: 10:181-213.
- Vonesh, R. F. and Chinchilli, V. M. *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. New York: Marcel Dekker, Inc., 1997.
- Wishart, J. «Growth-rate determinations in nutrition studies with the bacon pig and.» *Biometrika*, 1938: 30: 16-28.
- Zibrowius, H. « Les Scleractiniares de la Mediterranee et de l Atlantique nord-oriental.» Monaco: Mem. Inst. ocean, 1980. 11: 284.