**Climate
of the Past**

Exploring errors in paleoclimate proxy reconstructions using Monte Carlo simulations: paleotemperature from mollusk and coral geochemistry

M. Carré[1], J. P. Sachs[2], J. M. Wallace[3], and C. Favier[1]

[1]UM2-CNRS-IRD, UMR5554, Institut des Sciences de l'Evolution, Université Montpellier 2, Pl. Eugène Bataillon,
34095 Montpellier, France
[2]University of Washington, School of Oceanography, P.O. Box 355351, Seattle, WA 98195, USA
[3]University of Washington, Department of Atmospheric Sciences, P.O. Box 351640, Seattle, WA 98195, USA

Correspondence to: M. Carré (matthieu.carre@univ-montp2.fr)

Abstract. Quantitative reconstructions of the past climate statistics from geochemical coral or mollusk records require quantified error bars in order to properly interpret the amplitude of the climate change and to perform meaningful comparisons with climate model outputs. We introduce here a more precise categorization of reconstruction errors, differentiating the error bar due to the proxy calibration uncertainty from the standard error due to sampling and variability in the proxy formation process. Then, we propose a numerical approach based on Monte Carlo simulations with surrogate proxy-derived climate records. These are produced by perturbing a known time series in a way that mimics the uncertainty sources in the proxy climate reconstruction. A freely available algorithm, MoCo, was designed to be parameterized by the user and to calculate realistic systematic and standard errors of the mean and the variance of the annual temperature, and of the mean and the variance of the temperature seasonality reconstructed from marine accretionary archive geochemistry. In this study, the algorithm is used for sensitivity experiments in a case study to characterize and quantitatively evaluate the sensitivity of systematic and standard errors to sampling size, stochastic uncertainty sources, archive-specific biological limitations, and climate non-stationarity. The results of the experiments yield an illustrative example of the range of variations of the standard error and the systematic error in the reconstruction of climate statistics in the Eastern Tropical Pacific. Thus, we show that the sample size and the climate variability are the main sources of the standard error. The experiments allowed the identification and estimation of systematic bias that would not otherwise be detected because of limited modern datasets. Our study demonstrates that numerical simulations based on Monte Carlo analyses are a simple and powerful approach to improve the understanding of the proxy records. We show that the standard error for the climate statistics linearly increases with the climate variability, which means that the accuracy of the error estimated by MoCo is limited by the climate non-stationarity.

1 Introduction

Reconstructions of the past climate from proxy records involve a wide range of uncertainties at every step of the process. These uncertainties and the subsequent error bar in the reconstruction of a paleoclimatic variable need to be understood and quantified in order to properly interpret the reconstructed variability and to perform meaningful comparisons with climate model outputs. In a recent overview of methods used in high resolution paleoclimatology, Hughes and Ammann (2009) concluded that "*the study of the processes by which climate proxy records are formed [...] should be accorded high priority*".

Corals and mollusks are privileged archives for high resolution paleoceanographic studies and especially for El Niño Southern Oscillation (ENSO) reconstructions (Cole and

Fairbanks, 1990; Cole et al., 1993; Welsh et al., 2011). Short-term windows of monthly sea surface temperature (SST) series can be reconstructed from these accretionary archives using paleo-temperature proxies such as Sr/Ca (Beck et al., 1992; Marshall and McCulloch, 2002; Corrège et al., 2004; Rosenheim et al., 2004) or $\delta^{18}O$ (Epstein et al., 1953; Grossman and Ku, 1986; Böhm et al., 2000; Carré et al., 2005) serially measured along the growth axis. For these archives, the proxy development work has been mainly concentrated on the calculation of empirical regression models (or transfer functions) linking the geochemical proxies to the environmental variables. The error bar is then estimated for each calculated data point from the scattering of the calibration dataset. As for the reconstructed climate statistics, like the mean or the variance, the confidence in the reconstructed value is generally subjectively linked to the quality of the transfer function, but is not quantified with an error bar, except for a few studies in which the propagation of known errors was calculated for a reconstructed mean value (Abram et al., 2009) or for a reconstructed trend (Nurhati et al., 2011). Tudhope et al. (2001) estimated the ENSO-related variance in successive stages of the Late Quaternary from coral $\delta^{18}O$ records and could only address qualitatively the question of the statistical representativeness of their signal. A small number of studies focused on the uncertainties of coral-derived reconstructions. Evans et al. (1998) studied the effect of observational errors and the network of sparse coral records for the reconstruction of global SST. Brown et al. (2008) proposed that the variance of anomalies was the most robust measure of ENSO activity, and also pointed out that short records may reflect sampling of natural ENSO variability rather than a response to external forcing. Thompson et al. (2011) compared coral $\delta^{18}O$ records with pseudo-corals linearly calculated from climate model output and suggested that differences may be due to uncertainties in the proxy model and in the way coral record environmental variables. Sources of uncertainties are numerous (de Villers et al., 1995) but the way they affect reconstructions is unclear. There is growing agreement in the paleoclimate science community on the need for better methods to evaluate the uncertainties in climate proxy records (Jones et al., 2009). However, a full assessment of paleoclimate reconstruction uncertainties from the statistical analysis of modern dataset is strongly limited for corals and mollusks by the small size of datasets due to the practical constraints that characterize these archives (limited number of sites, scarcity, protection laws, analytical costs). In this case, we propose an approach based on Monte Carlo simulations with numerical surrogate climate proxy records.

Specifically, we provide a ready-to-use, parameterize-yourself, open access algorithm called MoCo available for Matlab® and R (Supplement) for estimating systematic and standard errors of the mean and variance of the annual SST, and the mean and variance of the SST seasonality reconstructed from mollusk and coral geochemistry. This algorithm is used in a case study using two instrumental SST time series from the Eastern Tropical Pacific to characterize and quantitatively evaluate the sensitivity of systematic and standard errors to categories of (1) sampling, (2) stochastic uncertainty sources, (3) archive-related biological limitations, and (4) climate non-stationarity when reconstructing the time series statistics from samples of short mollusk records.

Monte Carlo simulations have been used in previous studies (Briskin and Harrell, 1980; Ballentine and Hall, 1999; Touchan et al., 1999; Meibom et al., 2003; Kaufman, 2003; Evans et al., 2007) to estimate the error of a paleoclimate reconstruction. The method is thus not novel but its use has been limited to the estimation of a raw uncertainty value. Here we develop a methodological framework to quantitatively explore how systematic and standard errors of reconstructed statistics build from multiple sources, and to improve the understanding of the proxy signal, and eventually the quality of the coral and mollusk-based paleoclimate reconstructions. This technique is conceptually very simple compared to the full probabilistic modelling studies using Bayesian inferences that have been developed by statisticians for climate field reconstructions (Jones et al., 2009). It is intended for use as an intermediate method, realistic enough to provide reliable assessments of paleoclimate errors, while being technically and conceptually accessible to a broad community in paleoclimate science. We only consider the case where the statistics for the mean and variance of the climate state are estimated for a discrete time interval (similarly to Tudhope et al, 2011). We do not address the case of time series data being used to assess temporal changes in variability (similarly to Cole et al., 1993 or Cobb et al., 2003) where chronological uncertainties would also need to be addressed.

2 The MoCo program

As in all surrogate proxy (also referred to as "pseudo-proxy") studies, the basic principle is to use a realistic climate time series, sample and perturb it in a way that mimics the real sources of uncertainties (Mann and Rutherford, 2002). The inputs to use the MoCo program are (1) a monthly "target" time series from the studied locality, (2) a linear non-biased relationship linking the proxy to the reconstructed variable (thereafter referred to as the proxy model), and (3) the parameters defining the perturbations (more details in Sect. 3). MoCo randomly draws a sample of short windows out of the target time series and perturb them to simulate the sources of uncertainties characteristic of mollusks and corals. The perturbed chunks are considered as surrogate proxy-derived climate records and used together to estimate the statistics of the target time series. MoCo calculates the error of this simulated reconstruction by comparing the pseudo-reconstructed value with the "true" value, which is known from the original
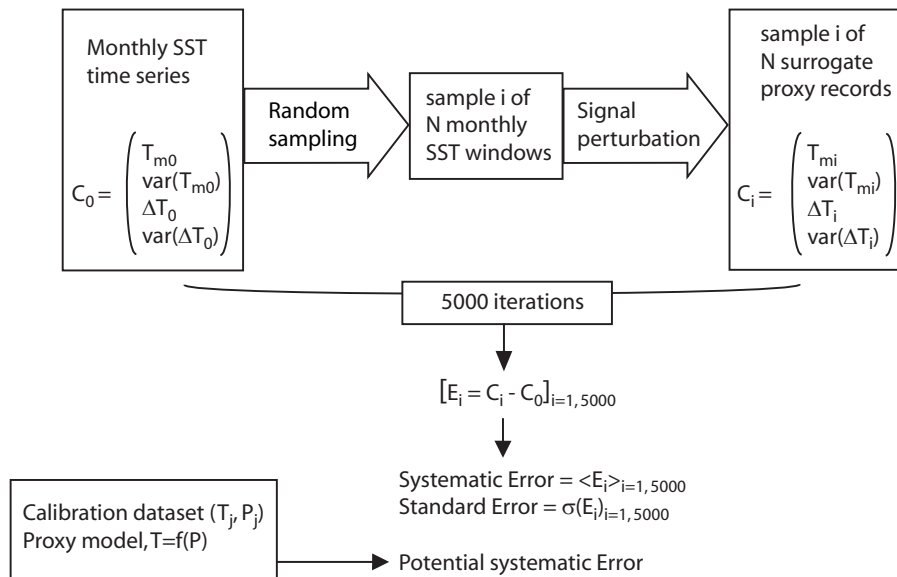
Fig. 1. Conceptual representation of the calculation of reconstruction errors in the MoCo algorithm.

non perturbed time series. By iterating this process thousands of times (Monte Carlo analysis), MoCo yields an estimate of the systematic error (the average error) and of the standard error (the standard deviation of the error) for the type of reconstruction that was simulated (Fig. 1).

Throughout the article, the variable of the original time series is the water temperature T, but, with slight modifications of MoCo, the method is suitable for any variable such as carbonate $\delta^{18}O$, salinity or pH. We explore the reconstruction of the statistics of the target climate, T_0, V_{T0}, Δ_0, and $V_{\Delta 0}$, which are respectively the average and the variance of the annual temperature, and the average and variance of the temperature annual amplitude, defined as follows for a N_0 year long monthly time series:

$T_0 = \frac{1}{N_0} \cdot \sum_{i=1,N_0} T_i$, where T_i is the mean temperature of year i.

$V_0 = \sigma^2(T_i)_{i=1,N_0}$

$\Delta_0 = \frac{1}{N_0} \cdot \sum_{i=1,N_0} \Delta_i$, where Δ_i is the annual amplitude ($T_{max} - T_{min}$) of year i.

$V_{\Delta 0} = \sigma^2(\Delta_i)_{i=1,N_0}$

T_0 and Δ_0 define the average climate, whereas V_{T0} and $V_{\Delta 0}$ are considered as measures of the interannual climatic variability. These statistics are estimated by T_m, Δ, V_T, and V_Δ respectively, calculated from a random sample of N surrogate climate proxy series of N_Y years, cumulating N_S years ($N_S = N \cdot N_Y$), with the following equations:

$T_m = \frac{1}{N_S} \cdot \sum_{i=1,N_S} T_i; \quad V_T = \frac{N_S}{N_S - 1} \cdot \sigma^2(T_i)_{i=1,N_S}$

$\Delta = \frac{1}{N_S} \cdot \sum_{i=1,N_S} \Delta_i; \quad V_\Delta = \frac{N_S}{N_S - 1} \cdot \sigma^2(\Delta_i)_{i=1,N_S}.$

T_m, Δ, V_T, and V_Δ are unbiased estimators of T_0, Δ_0, V_{T0}, and $V_{\Delta 0}$.

2.1 Different types of error

We explore both systematic and standard errors, S_E and σ_E of the estimators T_m, Δ, V_T, and V_Δ, defined as follows:

$S_E(T_m) = E(T_m) - T_0; \quad \sigma_E(T_m) = \sigma(T_m - T_0)$

$S_E(\Delta) = E(\Delta) - \Delta_0; \quad \sigma_E(\Delta) = \sigma(\Delta - \Delta 0)$

$S_E(V_T) = E(V_T) - V_{T0}; \quad \sigma_E(V_T) = \sigma(V_T - V_{T0})$

$S_E(V_\Delta) = E(V_\Delta) - V_{\Delta 0}; \quad \sigma_E = (V_\Delta) = \sigma(V_\Delta - V_{\Delta 0}).$

The esperance and standard deviation of the estimators are calculated from a population of 5000 values obtained from 5000 random samples of surrogate climate proxy records (Fig. 1). Identifying and estimating systematic errors may allow us to correct the reconstruction and improve its accuracy. A quantitative estimate of the standard error is also essential to determine a threshold of significance in the amplitude of the climate proxy variations.

Defining the error in a paleoclimate reconstruction from a local archive like a coral is not trivial because it depends on the climate information sought. An ideal proxy would provide the exact temperature in a precise location and thus be considered as error-free, but if the aim is to have regional scale information, the proxy signal would still be noisy owing to micro-environment effects. Weather also contributes to

the noise inherent in climate statistics. Thus, part of the uncertainty in the reconstruction is related to random sampling in time and space and is thus independent of the quality of the empirical regression model linking the proxy to the climate.

The formation of the proxy record involves a complex chain of physical and biological processes (for instance, mechanisms of Strontium incorporation into coral aragonite) that introduce non-climate-related stochasticity and limitations in the climate-proxy relationship (Allison et al., 2001, 2005; Meibom et al., 2003). The scatter inherent in in situ calibration datasets partly captures this stochastic variability but does not allow the exploration of its full range or the characterization of the uncertainty produced by different sources. Stochastic parameters may contribute to the standard error in the reconstruction of climate statistics as well as to systematic errors as we will show.

Paleoclimate reconstructions also involve systematic errors that cannot be estimated and corrected for, and could be referred to as potential systematic errors. They include errors related to the proxy calibration model. Considering that the mechanisms behind such errors are identical for all specimens in the archive (which is generally assumed), then the model would be identically wrong for all the climate calculations. Therefore, the paleoclimate error due to the imperfection of the proxy calibration model does not contribute to the standard error, but instead is more comparable to a systematic error (although its value might be linearly dependent on the proxy variable). It belongs thus to the category of potential systematic errors which are systematic errors for which only a potential range of values can be statistically estimated.

Here we only consider the potential systematic error S_P related to the uncertainty of the linear regression model between the reconstructed variable T and the proxy P: $T = \alpha \cdot P + \beta$. S_P is not estimated through a Monte Carlo simulation. $\pm S_P$ defines a 95 % confidence interval calculated as follows:

$$S_P(T_m) = \sqrt{E_\beta^2 + (T_m - T_C)^2 \cdot E_\alpha^2}; \ S_P(V_T) = V_T \cdot \frac{E_\alpha}{|\alpha|}$$

$$S_P(\Delta) = \Delta \cdot E_\alpha; \ S_P(V_\Delta) = V_\Delta \cdot \frac{E_\alpha}{|\alpha|}$$

where T_C is the mean temperature of the linear model calibration set. E_α and E_β are the uncertainty associated to the coefficients α and β so that $\alpha \pm E_\alpha$ and $\beta \pm E_\beta$ define the 95 % confidence interval of α and β. They are calculated by the following equations:

$$E_\alpha = t \cdot \frac{\sigma_T}{\sigma_P} \cdot \sqrt{\frac{(1 - R^2)}{N_C - 2}}; \ E_\beta = t \cdot \sigma_T \cdot \sqrt{\frac{(1 - R^2)}{N_C - 2}}$$

where N_C is the number of data points in the calibration dataset, σ_T and σ_P the standard deviations of T and P respectively in the calibration dataset, R is the Pearson's correlation coefficient, and t is the value of the Student variable

at the 0.05 confidence level and $N_C - 2$ degrees of freedom. MoCo yields estimates of the standard error, systematic error, and calculates the potential systematic error associated to the proxy linear model.

2.2 The Monte Carlo simulation

Monte Carlo techniques are useful for the analysis of complex stochastic systems (Metropolis and Ullam, 1949; Hastings, 1970). In our study, we simulate the reconstruction of climate statistics from samples of corals or shells from the same area. The shells of the sample are not strictly contemporaneous but belong to the same studied time period. The shell samples are simulated by samples of surrogate climate proxy records that are obtained by sampling a known sea surface temperature (SST) time series and perturbing each individual of the sample with a suite of random noises and constraints. The estimator of the climate statistics calculated from a sample surrogate climate proxy record is only the product of one random sample and of one realization of the infinite number of values and combinations that noise can take. In the Monte Carlo simulation, this process (Fig. 1) is repeated many times (5000 iterations in the experiments presented here) in order to have a representative sample of the range of responses of the estimators T_m, Δ, V_T, and V_Δ and robust estimates of the systematic and standard errors as defined in the previous section.

2.3 Using the MoCo application

The MoCo program is available as online supplementary material and can also be downloaded at http://www.isem.univ-montp2.fr/carre_matthieu, along with a parameter file and the MoCo_readme.txt file which contains step by step instructions for users. Two versions are available: in R and in Matlab® language. The algorithm is parameterized by the user according to the specifics of the study. The input parameters are listed in Table 1. The inputs required for the calculation of the potential systematic error related to the proxy model uncertainty are described in Sect. 2.1. In the next section, we only describe the inputs required for the Monte Carlo simulation.

3 Inputs to the algorithm

3.1 A linear proxy calibration model

The linear empirical relationship linking the geochemical proxy (e.g. Sr/Ca or $\delta^{18}O$) to the climate variable (SST) is thereafter referred to as the proxy model: $T = \alpha \cdot P + \beta$ (Sect. 2.1). As in all paleoclimate reconstructions, the proxy model is assumed to be unbiased and to be valid in the past. The statistics of the calibration dataset (detailed in Table 1) are used to calculate the potential systematic error related to the calibration scattering (Sect. 2.1). The parameters α and

Table 1. MoCo input parameters and output variables.

Parameters	Description	Type
	Input parameters	
	Monte Carlo analysis	
N	Number of specimen per sample	Integer
N_Y	Number of years spanned by individual records	Integer
T_S	Biological superior limit for skeletal growth	Real (T unit)
T_I	Biological inferior limit for skeletal growth	Real (T unit)
$gbi, i = 1$ to 12	Does skeletal growth occur during month i?	0/1
gap	How many random 1-month growth gaps per year ?	Integer [0, 12]
σ_S	Standard deviation of spatial T variations	Real (T unit)
σ_W	standard deviation of weather monthly noise	Real (P unit)
σ_C	standard deviation of carbonate micro-heterogeneity	Real (P unit)
σ_A	Analytical error (1σ)	Real (P unit)
$Niter$	Number of iteration of the Monte Carlo analysis	Integer
	Proxy model calibration	
$Alpha$	Slope of the linear proxy model	Real (T/P unit)
$Beta$	Intercept of the linear proxy model	Real (T unit)
σ_T	Standard deviation of T in calibration dataset	Real (T unit)
σ_P	Standard deviation of P in calibration dataset	Real (P unit)
R	Pearson's correlation coefficient in calibration dataset	Real in [0, 1]
N_C	Number of datapoints in calibration dataset	Integer
$T0$	Average value of T in calibration dataset	Real (T unit)
	Output Variables	
	Systematic errors	
$EsystTm$	Systematic error for T_m	Real (T unit)
$EsystvTm$	Systematic error for V_T	Real (T^2 unit)
$Esystdelta$	Systematic error for Δ	Real (T unit)
$Esystvdelta$	Systematic error for V_Δ	Real (T^2 unit)
	Standard errors	
$EstTm$	Standard error for T_m	Real (T unit)
$EstvTm$	Standard error for V_T	Real (T^2 unit)
$Estdelta$	Standard error for Δ	Real (T unit)
$Estvdelta$	Standard error for V_Δ	Real (T^2 unit)
	Potential systematic error related to the proxy model (95 % confidence level)	
$RETm$	Error bar for T_m	Real (T unit)
$REvTm$	Error bar for V_T	Real (T^2 unit)
$REmdelta$	Error bar for Δ	Real (T unit)
$REvdelta$	Error bar for V_Δ	Real (T^2 unit)

β of the proxy model allow a direct isometry between the proxy space and the temperature space.

3.2 A target time series

When using MoCo, a climate time series is first chosen that will be used as the "target climate". As will be discussed later, the characteristics of the target time series have a large influence on the estimation of the reconstruction uncertainty. The number of years N_0 of the time series should be larger than the total number of years N_S of the proxy records sample. A N_S/N_0 ratio lower than 0.2 would keep the average overlap rate in the random samples under 10 %. The length of the target time series should also be chosen according to the period the sample is expected to be representative of. For instance, if coral

records are being used to study millennial scale variability of ENSO, a time series of about 1000 years should be used as a target in the Monte Carlo simulation. Such a long time series could be either provided by climate models or by a very old modern coral, or could be statistically synthesized. Instrumental time series may be used when century-scale periods are studied from short-lived proxy archives such as one-year long mollusk shells. Here, proxy records and target time series have a monthly resolution.

3.3 Random sampling of target time series

Once the target time series is identified, N time windows of N_Y years are randomly drawn from it. N represents the sample size or the number of specimens that were analyzed to study the paleoclimate. N_Y is the typical duration (number of years) of the proxy records. Two approximations are made in this process: (1) the proxy record time resolution is assumed to be monthly and constant; and (2) all specimens of the sample are assumed to have the same duration.

3.4 Signal perturbation

In this step, the random sample of N windows is perturbed by stochastic noise and by archive-related limitations to simulate the effects of the real-world uncertainty sources. Overlap of the windows within a sample is possible but is generally limited since the size of the windows (N_Y years) is much smaller than the target time series (Sect. 3.1). Every window in the sample is independently perturbed. The product is a group of N monthly windows of surrogate proxy-derived SST. Specific details about the perturbations is given below.

3.4.1 Spatial heterogeneity

Corals and mollusks, as sessile organisms, record only a micro-scale environmental variability that may differ from the regional environmental variability. This effect is especially significant in coastal environments where large spatial heterogeneity can occur. In MoCo, this stochastic noise is represented by a random parameter with a normal distribution $N(0, \sigma_S)$. A random value is drawn for every specimen and added uniformly to the climate variable T over the whole time window. σ_S should be estimated by field measurements according to the studied geographic scale.

3.4.2 Monthly noise

MoCo provides for three additional sources of month-to-month noise: (1) the geochemical analysis standard error, (2) the weather scale variability during the life time of the specimen, and (3) biogenic carbonate heterogeneity at the μm scale of microsampling owing to vital effects and diagenesis. These three types of noise follow here the normal distributions N(0, σ_A), N(0, σ_W) and N(0, σ_C), respectively, where the standard deviations are expressed in the proxy unit

(e.g. mmol mol^{-1} for coral Sr/Ca). These three sources of uncertainty add in quadration because they are independent and yield a month-to-month noise with normal distribution $N(0, \sigma_m)$ ($\sigma_m^2 = \sigma_A^2 + \sigma_W^2 + \sigma_C^2$). For every monthly value of every time window, a random value is drawn, converted to the temperature scale, and added to the monthly SST value. All the stochastic noise sources are here considered to be normally distributed because they are expected to have symmetric distributions, and each is comprised of a large number of stochastic processes.

3.4.3 Limitations of the biological archives

The archives considered in this study, corals and mollusks, are living organisms. Their biology constrains their growth and thus the way in which they record the environment. Every species is defined by a range of physico-chemical tolerances beyond which they stop precipitating new carbonate skeletal material. If it is possible that the reconstructed variable may represent a growth limitation (like temperature or salinity) the effect of these limits should be explored and quantified. In MoCo, T_S and T_I represent temperature limits above and under which precipitation of skeletal material stops. The values beyond these threshold in a sample are discarded.
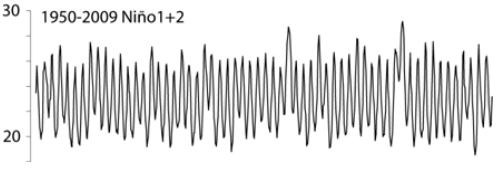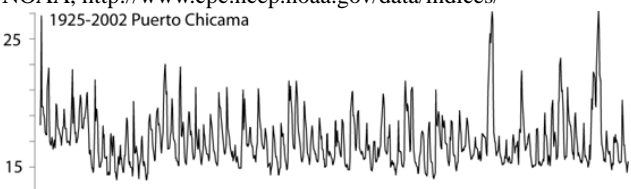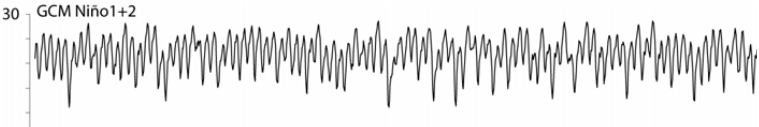
Some species systematically stop growing at a precise period of the year because their resources are exclusively dedicated to reproduction. This implies a systematic growth break in the record that may affect the final calculated averages or variance. The typical monthly growth pattern of the species is defined in MoCo by parameters gb1 to gb12 (Table 1).

Finally, growth breaks may occur randomly because of storms, predation, or sickness. In the MoCo program, the parameter *gap* allows the choice of occurrence of zero to 12 random growth breaks per year, of a monthly duration each (Table 1).

4 Sensitivity experiments

Six sensitivity experiments were performed using the MoCo algorithm to explore the influence of specific parameters on the systematic and standard errors calculated by MoCo. In every experiment, MoCo was used repeatedly with different settings to produce a series of error values that were plotted against the studied parameter. Experiments were thus designed to explore under different angles the influence of (1) the sample size (exp. 1, 2, and 6), (2) stochastic perturbations (exp. 1, 2, and 3), (3) the biological limitations of the archive (exp. 4 and 5), and (4) the target time series (exp. 3, 4, 5, and 6) on the systematic and standard errors when reconstructing the statistical characteristics of the target time series using the estimators T_m, Δ, V_T, and V_Δ. To perform these experiments we chose the illustrative case study

Table 2. Characteristics of monthly SST time series used as target in sensitivity experiments. Only the first 100 years of the GCM Niño1+2 SST time series are shown. The mean and variance of the annual SST (T_m and V_T), and the mean and variance of the seasonal amplitude of SST (Δ and V_Δ) were calculated for the complete series as described in Sect. 2.

Monthly SST target time series	T_m	V_T	Δ	V_Δ
1950-2009 Niño1+2 NOAA, http://www.cpc.ncep.noaa.gov/data/indices/	23.0	0.8	6.2	0.8
1925-2002 Puerto Chicama Quispe Arce (1993), http://jisao.washington.edu/data_sets/#time_series	17.1	1.6	4.5	4.3
GCM Niño1+2 Servonnat et al. (2010)	26.8	0.3	3.4	0.8

of temperature reconstructions in the Eastern Tropical Pacific from mollusk shell oxygen isotopes. We used the empirical proxy model established by Grossmann and Ku (1986) for biogenic aragonite (Eq. 2), which is often considered as the definition of isotopic equilibrium for biogenic aragonite and has been widely used for paleoclimate studies from aragonitic mollusk shells:

$$T\left(^{\circ}\mathrm{C}\right) = 19.73 - 4.34\left(\delta^{18}\mathrm{O}_{\mathrm{arag.}} - \delta^{18}\mathrm{O}_W\right).$$

$\delta^{18}\mathrm{O}_{\mathrm{arag.}}$ and $\delta^{18}\mathrm{O}_W$ are expressed in ‰ versus the V-PDB and V-SMOW international standards respectively (Coplen, 1996). This proxy model is used here for the case study. Any other proxy model could be used with MoCo. The biological characteristics of the species (growth breaks, temperature tolerance range, length of record...) are varied in the experiments to evaluate their influence on the reconstruction skills.

Three sea surface temperature (SST) time series were used as "target" climatology: (1) the 1925–2002 monthly in situ instrumental record from Puerto Chicama, Peru (Quispe Arce, 1993, dataset available at http://jisao.washington.edu/data_sets/#time_series), (2) the 1950–2009 monthly SST time series of the Niño1+2 region (NOAA Climate Prediction Center, http://www.cpc.ncep.noaa.gov/data/indices/), and (3) the 1000-year long monthly SST time series of the Niño1+2 area from the preindustrial control simulation of the IPSL_CM4v2 coupled ocean atmosphere general circulation model (GCM) (for details about the simulation, see

Servonnat et al., 2010). The target time series are presented in Table 2. The first 2 time series were used in all experiments except for the experiment 2 in which the GCM time series was used (Table 3). The parameterization of the sensitivity experiments are summarized in Table 3.

4.1 Experiment 1: influence of random sampling and stochastic noise

The first experiment was designed to test the effect of sampling on the standard and systematic errors and compare it to the effect of three stochastic noises that were turned off or on (Fig. 2). Realistic values were assigned to σ_S (spatial heterogeneity) and σ_m (month-to-month noise) based on field measurements on the Peruvian coast. The effect of random 1-month growth gaps was also tested (Table 3). In this experiment, shell records span one year and no biological limitations were included.

4.2 Experiment 2: should long or short records be preferred?

Experiment 2 was designed to explore the advantages and drawbacks of using short or long records for a reconstruction considering a constant sample size (i.e. a fixed total number of years N_S), and realistic stochastic perturbations of the proxy signal (Fig. 3). No biological limitations were included. The record length N_Y ranged from 1 to 200 years
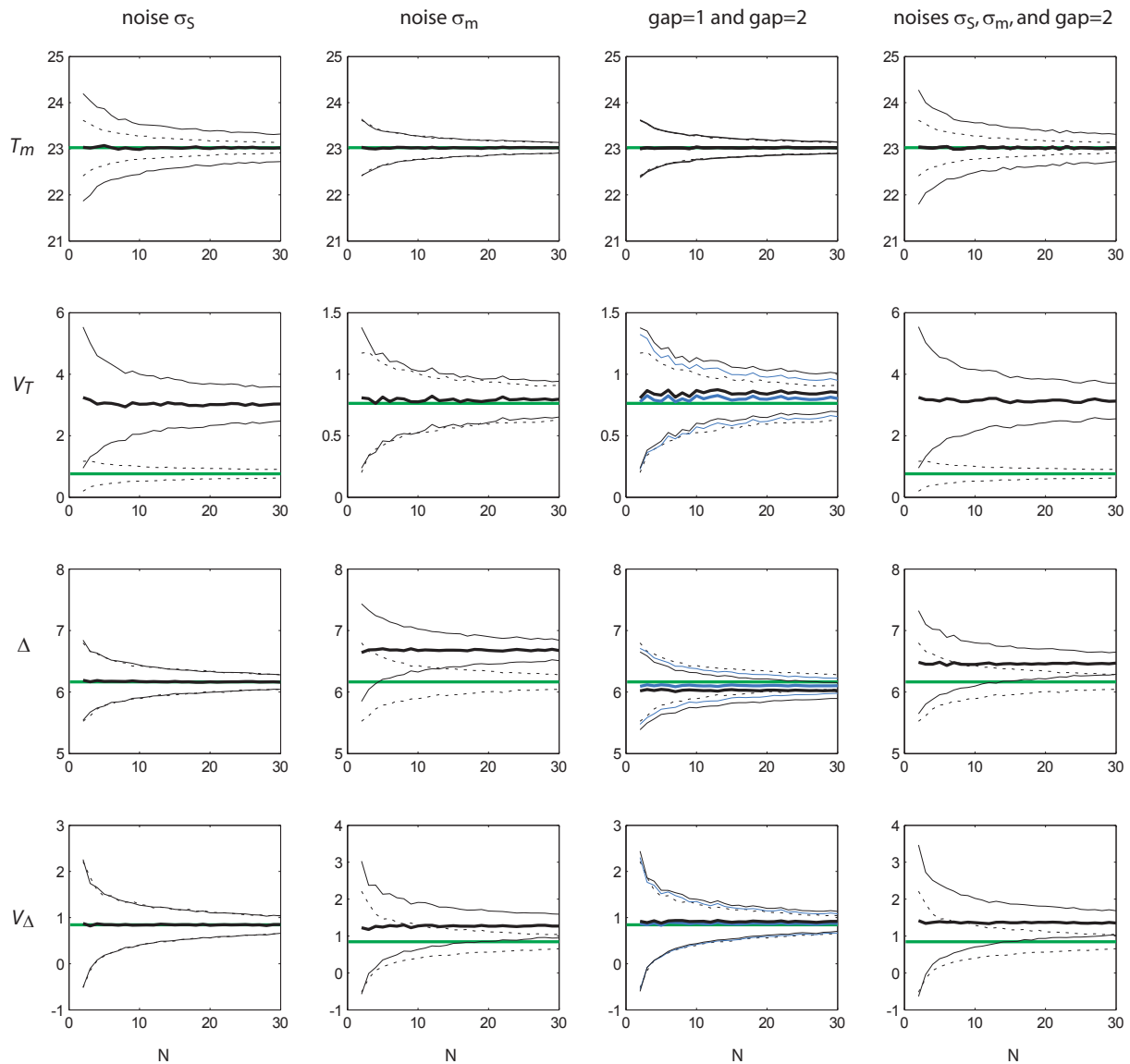
Fig. 2. Results of experiment 1 with the 1950–2009 Niño1+2 SST time series. Mean values of T_m, V_T, Δ, and V_Δ (black bold lines) calculated from 5000 iterations of surrogate proxy simulations (MoCo algorithm), and compared to the true values (green) of the target time series, versus the sample size (e.g. number of shells). Systematic errors are indicated by the difference between the mean calculated value and the true value. Dotted lines show the standard error interval ($\pm 1\,\sigma$) for an ideal proxy (no signal perturbation) versus the sample size. Thin black lines show the standard error interval ($\pm 1\,\sigma$) for surrogate proxies with stochastic noise. In the first three columns, the effects of spatial variability (σ_S), monthly variability (σ_m), and the occurrence of random growth breaks (blue: 1 per year, black: 2 per year) are investigated separately. Their effects are combined in the fourth column.

as would be expected for coral-based records, and the total number of years $N_S = N \cdot N_Y$ recorded by the sample was held constant at 200 years. No instrumental record was long enough for this experiment so a 1000 year long pre-industrial OA-GCM simulation of the Niño1+2 SSTs was used as a monthly target time series. Although this time series is not realistic, it is useful for this sensitivity test.

4.3 Experiment 3: influence of month-to-month noise

Experiment 3 explored further the effect of the individual data point quality degraded by the monthly noise, characterized by σ_m, which includes the analytical error, weather scale noise, and skeletal carbonate heterogeneity (Sect. 3.3.2). Here, all perturbations were turned off and all parameters were fixed except for σ_m. In MoCo, this noise is added to the SST time series as a random monthly series characterized
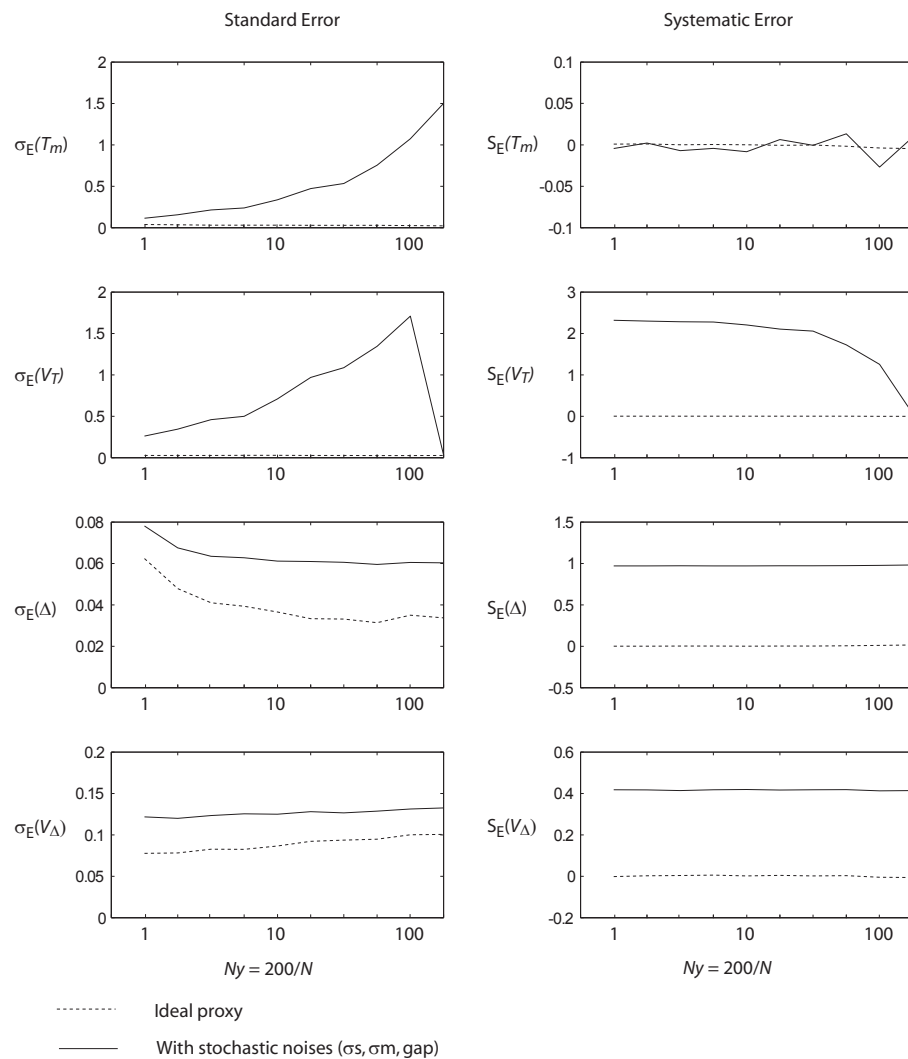
Fig. 3. Results of experiment 2 with the 1950–2009 Niño1+2 SST time series. Standard error (left panel) and systematic error (right panel) obtained for the reconstruction of T_m, V_T, Δ, and V_Δ using the MoCo algorithm, versus the length (number of years) of the proxy record, considering a constant number of 200 recorded years (from 200 one-year old shells to one 200-years old shell). Results using an ideal proxy (dotted line) were compared to results involving stochastic noise (black line).

by a standard deviation of $\alpha \cdot \sigma_m$. In this experiment, $\alpha \cdot \sigma_m$ increased gradually from 0 to σ, the standard deviation of the instrumental SST time series. The error values obtained from MoCo were plotted against the noise/signal ratio $R = \alpha \cdot \sigma_m / \sigma$. Simulations were performed with samples of 20 one-year long shells, and two different target time series (Fig. 4).

4.4 Experiment 4: influence of a yearly growth break

Experiment 4 was designed to test the effect on reconstruction errors of yearly growth breaks always occurring at the same period (such as spawning growth breaks) (Fig. 5). All other perturbations were turned off and all parameters except *gbi* were fixed. We only considered the case of a single

one-month growth break per year, and compared the effect of varying its month of occurrence. Simulations were performed with samples of 20 one-year long shells, and with two different target time series.

4.5 Experiment 5: influence of temperature tolerance range

In experiment 5, we explored the effect of skeletal growth temperature threshold on the errors of reconstruction of two target time series. In this experiment, an upper (lower) temperature threshold was imposed so that temperatures beyond that threshold were not recorded. The upper (lower) threshold value decreased from the maximum (minimum) temperature $T_{\max}$ ($T_{\min}$) of the target time series to $T_{\max} - 10\,°\mathrm{C}$

Table 3. Parameter setting in sensitivity experiments 1 to 5. Gray cells indicate varying parameters.

Parameters	Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5	Exp. 6
			Parameters of Monte Carlo analyses			
N	1 to 30	1 to 200[b]	20	20	20	10 to 30
N_Y	1	1 to 200	1	1	1	1
T_S	100	100	100	100	$T_{\max} - 10$ to $T_{\max}^c$	100
T_I	0	0	0	0	$T_{\min}$ to $T_{\min} + 10^c$	0
$gbi, i = 1$ to 12	1	1	1	0/1	1	1
gap	0/1	1	0	0	0	1
σ_S	0/1.5	1.5	0	0	0	1.5
σ_m^a	0/0.14	0.17	0 to 0.5	0	0	0.17
$Niter$	5000	5000	5000	5000	5000	5000
			Target SST time series			
1950–2009 Niño1 + 2	x		x	x	x	
1950–2002 Puerto Chicama	x		x	x	x	x
GCM Niño 1+2		x				

[a] $\sigma_m^2 = \sigma_A^2 + \sigma_W^2 + \sigma_C^2$; [b] $N \cdot N_Y = 200$; [c] $T_{\max}$ and $T_{\min}$ are the maximum and minimum temperature values of the monthly SST time series used as a "target".

($T_{\min} + 10\,^{\circ}\mathrm{C}$) by 0.1 °C increments. The standard and systematic error values calculated by MoCo under these conditions were plotted against the threshold values (Fig. 6). All other perturbations were turned off and all parameters except T_S and T_I were fixed. Simulations were performed with samples of 20 one-year long shells.

4.6 Experiment 6: influence of the time series variability

Experiment 6 was designed to explore the effect on the standard errors of changes in the time series climate variability. We used MoCo with the Puerto Chicama time series and 11 additional time series with decreasing variability. These additional time series were produced by removing El Niño and La Niña years from the Puerto Chicama time series. Since most of the interannual variability in this area is produced by ENSO, the values of V_T and V_Δ decreased when more El Niño or La Niña years were removed from the original time series, while the other general characteristics of the climate are conserved. The experiment was performed using 10, 20, and 30 shells of 1 year (Fig. 7). Realistic stochastic perturbations were included but no biological temperature threshold (Table 3).
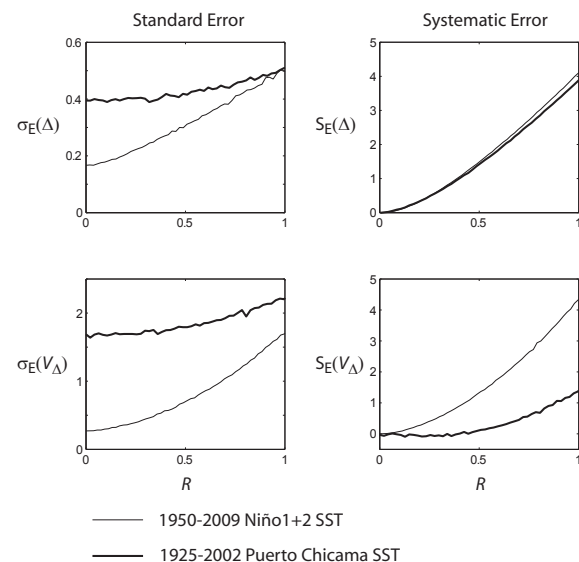

Fig. 4. Results of experiment 3. Standard error and systematic error for ΔT and var(ΔT) versus the standard deviation of the monthly stochastic perturbation. The effect was investigated using the 1950–2009 Niño1+2 (thin line) and the 1925–2002 Puerto Chicama (bold line) time series.
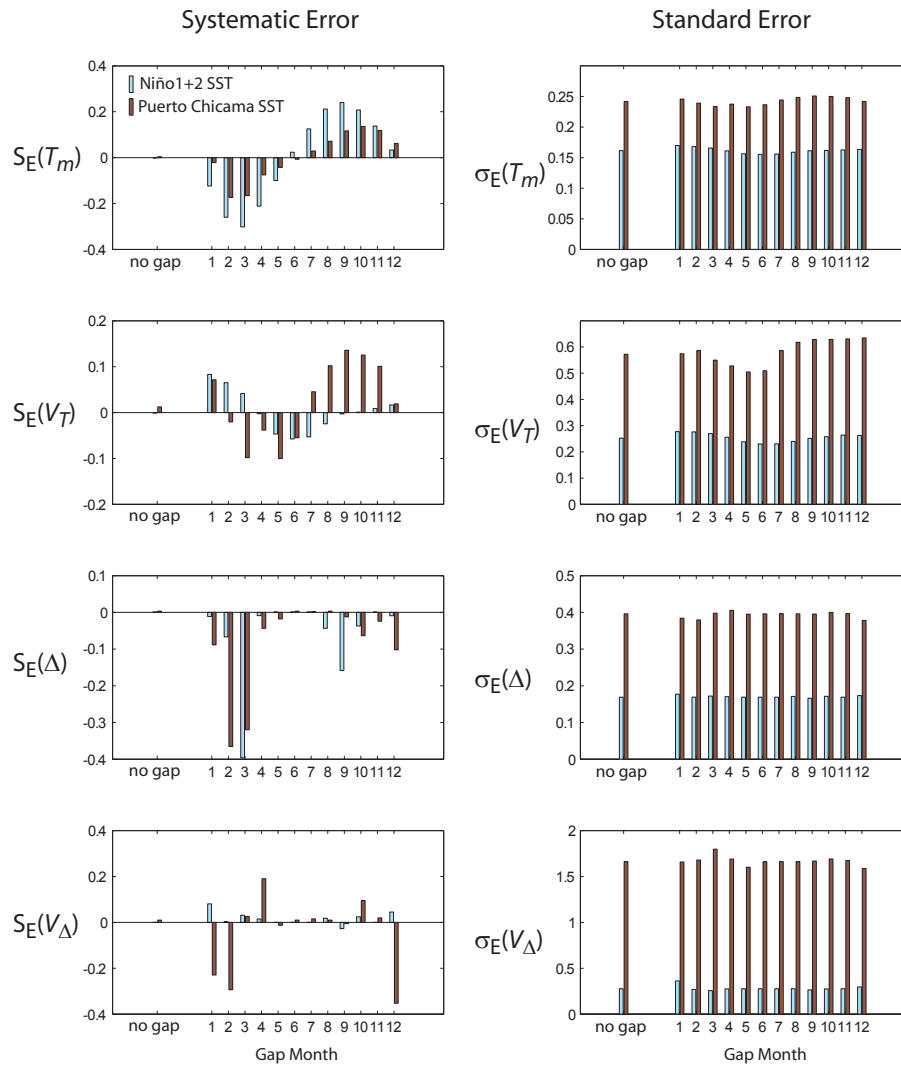
Fig. 5. Results of experiment 4. Standard and systematic errors for proxy reconstruction of T_m, V_T, Δ, and V_Δ of Niño1+2 SSTs (blue) and Puerto Chicama SSTs (red), versus the month of systematic growth hiatuses.

5 Results

5.1 Potential systematic error

The MoCo program calculates the potential systematic error bar (at 95 % confidence level) due to uncertainties in the linear proxy model calibration. This error is not estimated from surrogate proxy records but is simply calculated from the equations of error propagation presented in Sect. 2.1. This error bar not only depends on the proxy model calibration dataset but also on the target climate. When calculating the mean temperature, for instance, the error bar increases with the difference between the reconstructed conditions and the calibration dataset mean value. The mean value T_0 of Grossmann and Ku's (1986) dataset was 10.5 °C. When using the Puerto Chicama time series ($T_m = 17.1$ °C), the error bar for

T_m reconstruction due to the proxy model only is ± 2.1 °C (95 % confidence level). It increases to ± 3.8 °C with the Niño1+2 time series ($T_m = 23.0$ °C). These uncertainties are so large because the target temperature range is far from the temperature calibration range. If the mean value of the target time series was 11 °C, the uncertainty at 95 % confidence level would only be ± 0.4 °C. The proxy model-related error bar for the mean seasonal amplitude (Δ) was also significant for Puerto Chicama (± 1.4 °C) and for Niño1+2 (± 1.9 °C) time series. This confirms the importance of local specific calibration works to minimize this type of errors. For the interannual variability variables V_T and V_Δ, the relative uncertainty related to the proxy model is constant and equal to $E_\alpha/|\alpha|$ (see Sect. 2.1), which is here 7 %. These error bars are different from the standard errors studied thereafter and should therefore be represented separately.
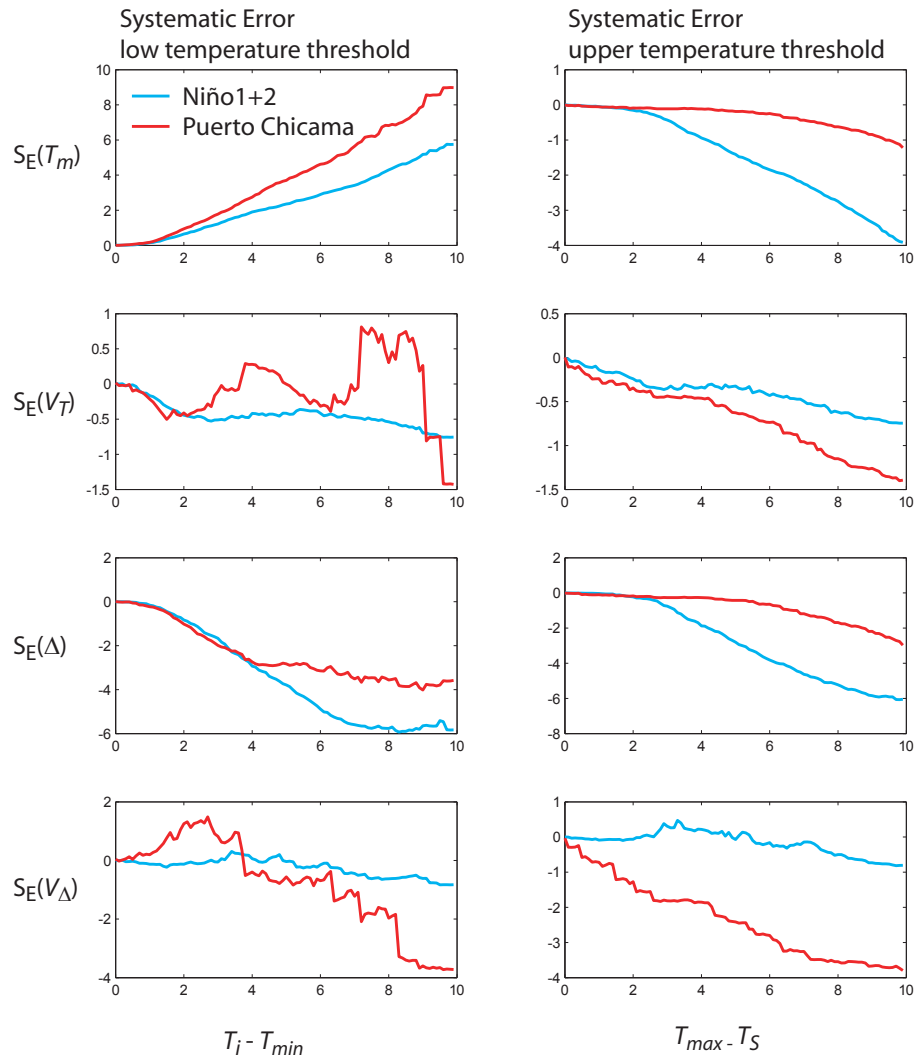
Fig. 6. Results of experiment 5. Systematic errors for T_m, V_T, Δ, and V_Δ due to growth temperature limits. The effects of inferior (superior) temperature limits are shown on the left (right). In the upper panels are indicated the temperature ranges of Niño1+2 SSTs (blue) and Puerto Chicama SSTs (red). The darker intervals show the range of temperature recorded by the archive limited by T_I (left panels) and T_S (right panel).

5.2 Random sampling (Exp. 1 and 2)

In experiment 1, the effect of the sample size is quantitatively estimated and compared to the effect of stochastic noises σ_m, σ_S and *gap* (Fig. 2). The effect of sampling only is represented in Fig. 2 by the "ideal proxy" curves obtained with all the perturbations turned off. It appears that random sampling is one of the main sources of the standard error. This standard error decreases rapidly with the sample size and becomes relatively insignificant for T_m and Δ when N reaches 20. On the other hand, the standard error for V_T and V_Δ due to sampling remains relatively significant up to $N = 30$. The strong effect of the sample size on the standard errors is also evidenced by the results of experiment 6 (Fig. 7).

In experiment 2, we test whether reconstructions from long proxy records are more reliable than reconstructions obtained from short proxy records considering a fixed total number of years recorded. N_S is kept constant and equal to 200 in order to test only the influence of the record length in the sampling and not of the total sample size. For mean temperature (T_m) reconstructions, a large sample set of short records provides a lower standard error than a few long records because local spatial temperature heterogeneity is better averaged out with numerous specimens. In a similar way, the effect of spatial variability (σ_S) tends to overwhelm the climatic variability V_T when the number of records is small, unless it is calculated from a single record (Fig. 3, $N = 1$, $N_Y = 200$). Short or long records do not yield significantly different results for systematic errors except for V_T.
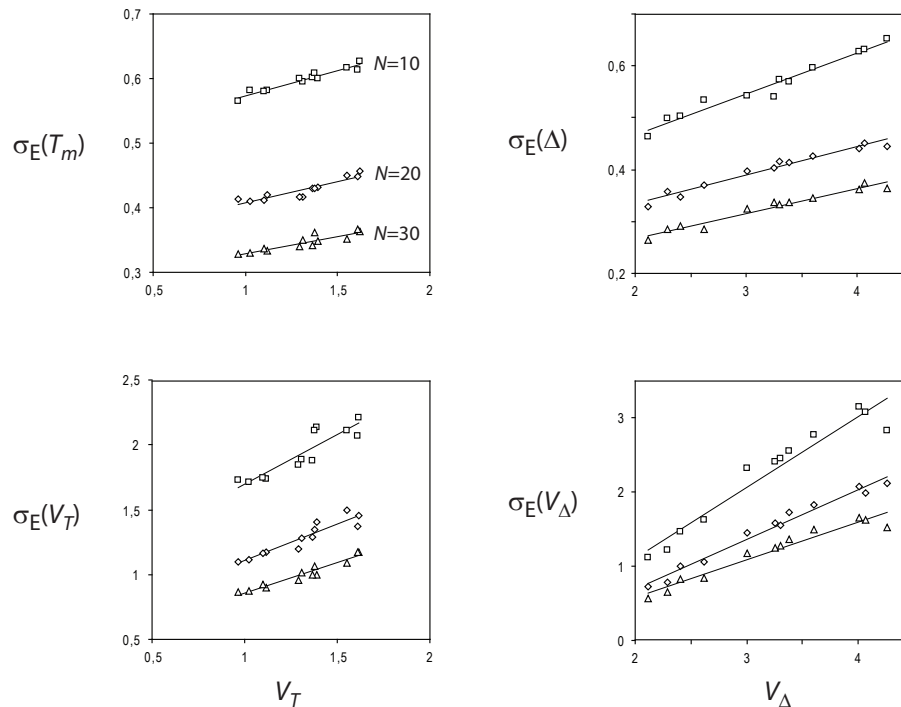
Fig. 7. Results of experiment 6. Left panels: Linear relationships between T_m and V_T standard errors obtained from MoCo versus V_T in 12 time series modified from the Puerto Chicama time series. Right panels: Linear relationships between Δ and V_Δ standard errors obtained from MoCo versus V_Δ. For every variable, a least-square linear correlation was obtained for reconstructions using 10 shells (squares), 20 shells (diamonds), and 30 shells (triangles).

V_T is overestimated when using short records because the spatial variability adds up to the climatic variability, but this effect decreases when using fewer longer records. Finally it appears that intermediate values of N and N_Y (here 20 10-years old shells) would yield the best compromise for accuracy and precision in the reconstruction of T_m and V_T. The reconstruction skills for Δ, and V_Δ are not significantly affected by the record length in our experiment.

5.3 Effects of stochastic noise (Exp. 1, 2, and 3)

Three kinds of stochastic perturbations (see Sect. 4) are applied to the climate signal in experiments 1, 2, and 3 in order to produce surrogate climate proxy records. In the first experiment, their influence is observed separately and compared to ideal proxy reconstruction errors. As expected, spatial variability greatly affects the standard error of the T_m reconstruction (Fig. 2) and this effect decreases when N increases (Fig. 3). It also induces a systematic positive bias for V_T reconstruction which also decreases when N increases (Fig. 3).

In our case study, the monthly stochastic perturbation characterized by σ_m involves monthly water $\delta^{18}O$ variability, carbonate $\delta^{18}O$ analytic error, and shell carbonate heterogeneity. The monthly stochastic perturbation in experiment 1 does not significantly affect the T_m and V_T reconstructions but it produces an unexpected overestimation of

the annual amplitude Δ and of its variance V_Δ (Fig. 2). Its effect on Δ and V_Δ reconstructions is further explored in experiment 3 (Fig. 4) using two different target SST time series. The maximum value of σ_m corresponded to a noise/signal ratio $R = 1$, which represents an extremely noisy proxy record. The systematic positive bias on Δ estimate is $\sim$0.1 °C when $R = 0.1$ and increases to 3 °C when $R = 0.8$. For other parameters, the response depends on the target time series. For $R = 0.4$, the systematic error for V_Δ is about 0.8 (or 100 %) for the Niño1+2 time series, while it is almost null for the Puerto Chicama time series. Standard errors for Niño1+2 SSTs are more sensitive to the monthly noise than those for Puerto Chicama SSTs. This can be explained by the higher interannual climate variability of Puerto Chicama (Table 2), which makes the noise-related variability relatively smaller.

Random growth breaks have no significant impact on the standard error for any of the variable (Fig. 2). They induce a slight positive bias for V_T and a slight negative bias for Δ. These biases are due to the time series properties since they are not observed when using the Puerto Chicama time series (not shown).

5.4 Effects of biological limitations (Exp. 4 and 5)

Growth hiatuses may occur every year at approximately the same date for breeding or other reasons (Sato et al., 1999). In experiment 4 (Fig. 5), we showed that the date of the growth break has little impact on standard errors but may produce systematic errors. As expected, the mean annual temperature would be underestimated (overestimated) if the growth breaks occur in the warmest (coldest) period. Here, the maximum systematic error for T_m was $-0.3\,°C$. The annual amplitude may be largely underestimated if the growth hiatuses occur during seasonal extrema. In our experiment the systematic bias for Δ reached $-0.4\,°C$, with a systematic growth break in March for the Niño1+2 SST time series. Proxy reconstructions of variances were affected by growth hiatuses in a much less predictable way. For Puerto Chicama SSTs, maximum systematic errors reached about 8 % for V_T when growth breaks occurred in September. The largest underestimation (8 %) for V_Δ as obtained for growth breaks occurring in December, January, or February. This is because V_Δ is strongly related to El Niño events whose maximum development occurs during these months. Again, the error was strongly dependent on the target time series.

Temperature tolerance for skeletal growth is an especially important biological limitation that may induce significant systematic biases in paleoclimate reconstructions. These biases were explored with two different time series in experiment 5 (Fig. 6). While it is obvious that upper (lower) temperature limits cause underestimation (overestimation) of the mean temperature T_m, and that Δ is in both cases underestimated, our experiment permits calculation of a quantitative estimate of this systematic biases. In experiment 5, the systematic error responses to lower and upper limits were not symmetrical. For Niño1+2, the effect of the lower temperature limit began to be significant for T_m and Δ when it reached $\sim1.5\,°C$ above the time series minimum temperature, whereas the effect of the upper limit began to be significant when it reached $\sim3\,°C$ under the time series maximum temperature. The biases produced for V_T and V_Δ reconstructions remained relatively low for Niño1+2 while, for Puerto Chicama, they rapidly became significant and had unpredictable profiles, switching from positive to negative values (Fig. 6), especially for the lower temperature limit. This is due to the particular variability of coastal SSTs in Peru characterized by a weak annual cycle of $\sim4\,°C$ periodically perturbed by very strong El Niño anomalies that can reach up to $9\,°C$ in austral summer. Therefore, when the lower temperature threshold reaches $4\,°C$ above the minimum SST, most of the annual cycle is not recorded, so that the variability is restricted to the highly irregular El Niño-related summer anomalies. On the other side, the upper threshold cuts the largest El Niño anomalies and systematically induces an underestimation of the interannual variability V_T or V_Δ.

5.5 Influence of the target time series (Exp. 3, 4, 5, and 6)

The results of experiments 3, 4, and 5 showed that the behavior and values of standard and systematic errors strongly depend on the climate time series used as a target in the experiment because they depend on the probability density function of the temperature. The Puerto Chicama time series has a much wider distribution than the Niño1+2 time series (Table 2) so that standard error due to random sampling is much larger for Puerto Chicama (Fig. 4, $\sigma_m = 0$; Fig. 5). In the experiments 3, 4, and 5, the error responses to perturbations were strongly modulated by the characteristics of the target time series. The same proxy perturbation may induce strong errors with one time series and be insignificant with another. For instance, the upper growth temperature limit had no significant effect on the reconstruction of T_m and Δ until it reached $\sim6\,°C$ under the maximum value of the Puerto Chicama SST (Fig. 6) because of the highly asymmetric distribution of temperature due to El Niño phenomenon at this location. These results show that the choice of the target time series when using MoCo to estimate paleoclimate reconstruction errors is a critical step. The standard and systematic error estimated with the IPSL_CM4v2 GCM Niño1+2 time series are therefore not reliable because the seasonal cycle and the ENSO variability are not correctly represented (Table 2).

The experiment 6 showed that the standard error of all the proxy climate statistics increases almost linearly with the interannual variability in Puerto Chicama (R^2 values ranges from 0.81 to 0.98). The dependence, indicated by the slope of the linear regressions, is about 10 times lower for the standard error of the reconstructed T_m and Δ, than for the reconstructed variance V_T and V_Δ. The relationships are also strongly affected by the sample size N. This result means that the reconstruction of past ENSO activity indicated by the variance has an error bar that varies with the ENSO activity itself.

6 Discussion

6.1 Implications for paleoclimate error representation

Three types of errors have been distinguished in this study that should all be treated distinctively. Systematic errors are meant to be corrected for when detected and quantified, and thus are not supposed to be represented. Large modern datasets are required to detect and calculate a robust estimate of a bias in a proxy climate reconstruction. The number of modern coral records in a location is always small because of limited specimen availability and/or analytical costs. MoCo provides an alternative way to detect and quantify biases and thus improve the accuracy of coral and mollusk reconstructions provided the parameterization and the target time series are realistic in the specific context of the study.

The standard error is due to sampling, stochasticity at different time and space scale, and the biological limitations that have been explored in our experiments. This error bar exists even if the proxy model is ideal. The standard errors estimated for different shell or coral samples are independent.

The errors related to the proxy model calculated in different samples of the same type of archive are not independent since they originate from the same proxy model. For this reason, this type of error is fundamentally different from the standard error.

Depending on the context, these two errors may be combined or represented separately. Generally, if estimating the error bar for a single period or event, both errors may be added in quadrature. Nurhati et al. (2009, 2011) combined the analytical standard error with the uncertainty related to the proxy calibration to estimate the error bar of a reconstructed temperature trend through the 20th century. However, if several successive samples are analyzed to study long-term climate variability, the standard error and the potential systematic error may be represented separately in paleoclimate results as in Carré et al. (2012) for a more complete representation of the reconstruction uncertainty.

6.2 Understanding the proxy

A sensor-specific precision of reconstruction is often attributed to the paleoclimate proxy estimates (for example Beck et al., 1992; Anand et al., 2003). Our numerical experiments simulating the process of paleoclimate reconstruction from coral or mollusk shell geochemistry show that the error cannot be considered a constant value characteristic of the proxy. Stochastic uncertainties and biological limitations significantly affect the resulting climate reconstruction, in different manners depending on the studied statistics (T_m, V_T, Δ, V_Δ) and on the studied location. General qualitative relationships between the perturbations, the target time series characteristics and the reconstruction skills are described in the result section. However, standard and systematic errors are so sensitive to the noise parameterization and the target time series that we recommend reproducing the sensitivity of experiments described here with a parameterization specific to the study in order to get a quantitative understanding of the proxy records.

The results of the experiments yield an illustrative example of the range of variations that climate reconstruction errors may undergo, and bring to light their complexity. Although classic calibration-validation techniques are useful for a first approximation test of corals and mollusks reliability as archives, they are not well-suited for identifying the causes of reconstruction errors, estimating their relative contribution, or understanding how errors accumulate from a multitude of sources.

While the influences of several sources of error were qualitatively predictable, some perturbations produced significant

unpredictable systematic bias (Figs. 4–6). Beyond the findings of error sources for coral and mollusk-based reconstructions, our study demonstrates that numerical simulations based on Monte Carlo analyses are a simple and powerful approach to improve the proxy calibration process. A thorough understanding of the proxy record errors is essential for the interpretation of paleoclimate records from proxies derived from accretionary skeleton geochemistry.

6.3 Quantifying errors: contribution and limitations

Quantifying errors in paleoclimate reconstructions is essential for accurate and meaningful proxy-proxy and proxy-model comparisons. In the case of mollusks and corals, modern datasets are necessarily limited by costs and material availability, which limits statistical techniques to accurately estimate reconstruction uncertainties. The MoCo algorithm was designed to provide quantitative estimates of the three kinds of errors identified in Sect. 2.1 for coral and mollusk based climate reconstructions. It implies that the linear proxy model has been previously validated. The accuracy of the systematic and standard error quantification using MoCo may be affected by several limitations, as we now describe.

The model for the paleoclimate reconstruction process implies simplifications including: (1) proxy records in a sample are considered of equal duration and of constant monthly resolution, and (2) stochastic noise was represented by stationary normal distributions. Furthermore, the parameterization of MoCo by the user requires field measurements and knowledge of the organism ecology and growth, characterized by fixed values that have their own uncertainty and variability. Uncertainties and non-stationarity in σ_S, σ_m, T_I or T_S estimates may affect the quantification of errors.

The standard and systematic errors calculated by MoCo are not true errors but only estimates from simulations based on the assumption that the temperature distribution of the paleoclimate is similar to the distribution of the time series used in the simulation. The main limitation is thus related to the non-stationarity of climate, although this issue is not peculiar to our method but is general in paleoclimatology since uncertainties are necessarily assessed in modern conditions. By definition, the true target climate is unknown, and necessarily different from the target time series used in the MoCo simulation. Therefore, the estimation of the standard and systematic errors may be biased by the difference of variability between the true climate and the target time series. In the selection process for the simulation target time series, it is recommended to seek the most realistic time series.

However, this bias can be minimized for the standard errors. Experiment 6 showed that in a defined location, the standard error is linearly correlated to the interannual variability through a relationship that is strongly modulated by the conditions of the reconstructions and especially the sample size. Therefore, MoCo should be used repeatedly as in experiment 6 to estimate the linear relationships specific to

the paleoclimate study. The standard error of a proxy reconstruction can then be calculated using this relationship and the reconstructed values of V_T and V_Δ. This method is valid if we consider that in a defined location, the interannual or decadal variability may change significantly through time but the larger climatic features responsible for these modulations remain similar, at least on multi-millennial timescales. For example, in the tropical Pacific, the characteristics of the climate variability can be considered conservative in a defined location and its amplitude is modulated by the activity of ENSO and the Pacific Decadal Oscillation (PDO). Therefore, based on experiment 6, if $V_\Delta = 3\,°\mathrm{C}^2$ is reconstructed in the Peruvian coast from 20 1-year long mollusk shells, the linear relationship estimated in experiment 6 (Fig. 7) yields a value of standard error of $0.4\,°\mathrm{C}^2$ for Δ and $1.3\,°\mathrm{C}^2$ for V_Δ. This method uses the information of the reconstruction itself to narrow the potential biases due to climate non-stationarity but is not perfect because the reconstruction is necessarily not exact and because it assumes the stationarity of the variability vs. standard error relationship.

Very few coral and mollusk paleoclimate studies include a proper error calculation. One of the most complete attempt was made by Abram et al. (2009) who estimated the standard error of mean temperatures estimated from Sr/Ca coral records combining the variability between several corals from the same area (which corresponds to the parameter σ_S in our study) and a measured standard error related to analytical precision and variability within corals (which is similar to our parameter σ_m). Although a large part of the standard error is due to these parameters, MoCo represents a significant improvement since it involves additional error sources for a more realistic estimate. It also offers the possibility to quantitatively estimate the influence of individual error sources. The improvements of this method are: (1) the identification of 3 independent categories of errors, (2) the quantification of the standard error due to uncertainty sources that were so far not considered, and (3) the identification and estimation of systematic biases that would not otherwise be detected because of the limited modern dataset. It thus offers the possibility of correcting the proxy-based climate from these biases for a more accurate reconstruction.

6.4 Extending applications

Sensitivity experiments were based on an illustrative case study of SST reconstruction from mollusk or coral δ^{18}O in an environment where water δ^{18}O is reasonably constrained. These experiment could be similarly performed to improve the understanding of other proxies including temperature reconstructions from Sr/Ca ratios in corals (Beck et al., 1992; De Villiers et al., 1994; Marshall and McCulloch, 2002) and sclerosponges (Swart et al., 2002; Rosenheim et al., 2004), Mg/Ca ratios in coralline algae (Kamenos et al., 2008), and pH from coral δ^{11}B (Rollion-Bard et al., 2011).

In many environments, calcium carbonate δ^{18}O variations yield a mixed signal between water temperature and water δ^{18}O variations related to freshwater input. Changes in the climate variability are thus estimated by the δ^{18}O variability in the fossil coral compared to modern corals (Tudhope et al., 2001). Under such conditions, the uncertainties of the proxy reconstructions should be evaluated in the space of the proxy variable ($\delta^{18}\mathrm{O}_{\mathrm{carb}}$), which would require very little modifications of the MoCo since the climate and the proxy space are linearly linked by the proxy regression model (Sect. 3.1). A target time series of $\delta^{18}\mathrm{O}_{\mathrm{carb}}$ would be required, that could be either provided by a very long modern coral of the same area or calculated from temperature and water δ^{18}O time series. This approach, referred to as forward modelling, has been applied to corals by Thompson et al. (2011) and a variety of proxies, such as planktonic foraminifera δ^{18}O (Schmidt, 1999), or stalagmite δ^{18}O (Baker and Bradley, 2010). The strength of forward modelling that incorporates Monte Carlo analyses for paleoclimate proxy calibration was showed by Evans (2007) through a case study with wood cellulose δ^{18}O. To our knowledge it has never been applied to corals or mollusks. MoCo-type algorithms would also be useful for exploring the uncertainties of salinity or water δ^{18}O reconstructions based on the combination of coral δ^{18}O and Sr/Ca ratios, since the perturbations of both proxies add with different types of perturbations and propagate in an unpredictable way.

This study does not apply to time series reconstructions used to assess temporal changes in variability. Further work in a similar methodological framework is needed to study the effect of age uncertainty, stochasticity and proxy biological limitations on the correlation between the reconstructed time series and target time series.

7 Conclusions

Reconstructions of paleoclimate statistics (annual mean, mean seasonal amplitude and interannual variability) from coral and mollusk geochemistry were so far performed either without quantified error bars or with error estimates that do not involve all the sources of uncertainty. We defined here the error bar due to the proxy calibration as a potential systematic error, and defined two additional and independent categories of errors: the standard error and the systematic error that build from multiple sources and are potentially larger than those estimated from empirical calibration scatter. We provided an open access program, MoCo, available for R and Matlab® (Supplement and also at http://www.isem.univ-montp2.fr/carre_matthieu), that calculates the error related to the proxy linear model, and estimate with a Monte Carlo simulation the systematic biases, and the standard errors for the reconstruction of climate statistics from mollusks and corals. MoCo involves realistic surrogate proxy-derived climate records, and was designed to improve

the accuracy and the quantification of coral and mollusk-based SST reconstructions. We showed in an illustrative case study that surrogate proxy techniques associated with Monte Carlo analyses are powerful tools to improve the understanding and calibration of mollusk and coral proxy records. Sensitivity experiments showed the significant and often unpredictable influence of random sampling, stochastic variability in the proxy formation, archive biological limitations, and climate characteristics. These numerical experiments are a fast and efficient technique for a quantitative assessment of the relative contribution of these influences in the systematic and standard errors of the paleoclimate reconstruction.

The accuracy of error quantification by MoCo is mainly limited by: (1) the simplification of the simulated climate reconstruction process, (2) the variability of the parameters used in MoCo, (3) the potential non-stationarity of the proxy model, and (4) the non-stationarity of climate between the modern reference used in the simulation and the studied period. We propose a technique to constrain this latter source of bias using the value of the reconstructed variability and the relationship between the standard error and the variability that can be estimated with MoCo. While the errors estimated by MoCo are only the product of simulations and may be imperfect, this method is still a significant improvement for quantitative reconstructions of climate statistics from mollusks and corals. Its conceptual simplicity should allow it to be adapted for a wide range of applications in paleoclimate research involving, among other archives, corals, mollusks, sclerosponges and coralline algae.

**Supplementary material related to this article is available online at:
http://www.clim-past.net/8/433/2012/
cp-8-433-2012-supplement.zip**.

References

Abram, N. J., McGregor, H. V., Gagan, M. K., Hantoro, W. S., and Suwargadi, B. W.: Oscillations in the southern extent of the Indo-Pacific Warm Pool during the mid-Holocene, Quaternary Sci. Rev., 28, 2794–2803, 2009.

Allison, N., Finch, A. A., Sutton, S. R., and Newville, M.: Strontium heterogeneity and speciation in coral aragonite: implications for the strontium paleothermometer, Geochim. Cosmochim. Acta, 65, 2669–2676, 2001.

Allison, N., Finch, A. A., Newville, M., and Sutton, S. R.: Strontium in coral aragonite: 3. Sr coordination and geochemistry in relation to skeletal architecture, Geochim. Cosmochim. Acta, 69, 3801–3811, 2005.

Anand, P., Elderfield, H., and Conte, M. H.: Calibration of Mg/Ca thermometry in planktonic foraminifera from a sediment trap time series, Paleoceanography, 18, 1050, doi:10.1029/2002PA000846, 2003.

Baker, A.,and Bradley, C.: Modern stalagmite δ^{18}O: Instrumental calibration and forward modelling, Global Planet. Change, 71, 201–206, 2010.

Ballentine, C. J. and Hall, C. M.: Determining paleotemperature and other variables by using an error-weighted, nonlinear inversion of noble gas concentrations in water, Geochim. Cosmochim. Acta, 63, 2315–2336, 1999.

Beck, J. W., Edwards, R. L., Ito, E., Taylor, F. W., Recy, J., Rougerie, F., Joannot, P., and Henin, C.: Sea-surface temperature from coral skeletal Strontium/Calcium ratios, Science, 257, 644–647, 1992.

Böhm, F., Joachimski, M. M., Dullo, W.-C., Eisenhauer, A., Lehnert, H., Reitner, J., and Wörheide, G.: Oxygen isotope of marine aragonite of coralline sponge, Geochim. Cosmochim. Acta, 64, 1695–1703, 2000.

Briskin, M. and Harrell, J.: Time-series analysis of the Pleistocene deep-sea paleoclimatic record, Mar. Geol., 36, 1–22, 1980.

Brown, J., Tudhope, A. W., Collins, M., and McGregor, H. V.: Mid-Holocene ENSO: Issues in quantitative model-proxy data comparisons, Paleoceanography, 23, PA3202, doi:10.1029/2007PA001512, 2008.

Carré, M., Bentaleb, I., Blamart, D., Ogle, N., Cardenas, F., Zevallos, S., Kalin, R. M., Ortlieb, L., and Fontugne, M.: Stable isotopes and sclerochronology of the bivalve Mesodesma donacium: potential application to peruvian paleoceanographic reconstructions, Palaeogeogr. Palaeocl., 228, 4–25, 2005.

Carré, M., Azzoug, M., Bentaleb, I., Chase, B. M., Fontugne, M., Jackson, D., Ledru, M.-P., Maldonado, A., Sachs, J. P., and Schauer, A. J.: Mid-Holocene mean climate in the south-eastern Pacific and its influence on South America, Quatern. Int., 253, 55–66, doi:10.1016/j.quaint.2011.02.004, 2012.

Cobb, K. M., Charles, C. D., Cheng, H., and Edwards, R. L.: El Niño/Southern Oscillation and tropical Pacific climate during the last millennium, Nature, 424, 271–276, 2003.

Cole, J. E. and Fairbanks, R. G.: The southern oscillation recorded in the δ^{18}O of corals from Tarawa atoll, Paleoceanography, 5, 669–683, 1990.

Cole, J. E., Fairbanks, R. G., and Shen, G. T.: Recent Variability in the Southern Oscillation: Isotopic results from a Tarawa Atoll coral, Science, 260, 1790–1793, 1993.

Coplen, T. B.: New guidelines for reporting stable hydrogen, carbon, and oxygen isotope-ratio data, Geochim. Cosmochim, Acta, 60, 3359-3360, 1996.

Corrège, T., Gagan, M. K., Beck, J. W., Burr, G. S., Cabioch, G., and Le Cornec, F.: Interdecadal variation in the extent of South Pacific tropical waters during the Younger Dryas event, Nature, 428, 927–929, 2004.

De Villiers, S., Shen, G. T., and Nelson, B. K.: The Sr/Ca-temperature relationship in coralline aragonite: Influence of variability in $(Sr/Ca)_{seawater}$ and skeletal growth parameters, Geochim. Cosmochim. Acta, 58, 197–208, 1994.

Epstein, S., Buchsbaum, R., Lowenstam, H. A., and Urey, H. C.: Revised carbonate-water isotopic temperature scale, Bull. Geol. Soc. Am., 64, 1315–1326, 1953.

Evans, M. N.: Toward forward modeling for paleoclimatic proxy signal calibration: A case study with oxygen isotopic composition of tropical woods, Geochem. Geophy. Geosys., 8, Q07008, doi:10.1029/2006GC001406, 2007.

Evans, M. N., Kaplan, A., and Cane, M. A.: Optimal sites for coral-based reconstruction of global sea surface temperature, Paleoceanography, 13, 502–516, 1998.

Grossman, E. L. and Ku, T.-L.: Oxygen and carbon fractionation in biogenic aragonite: Temperature effect, Chem. Geol., 59, 59–74, 1986.

Hastings, W. K.: Monte Carlo sampling methods using Markov chains and their applications, Biometrika, 57, 97–109, doi:10.1093/biomet/57.1.97, 1970.

Hughes, M. and Ammann, C.: The future of the past – an earth system framework for high resolution paleoclimatology: editorial essay, Climatic Change, 94, 247–259, 2009.

Jones, P. D., Briffa, K. R., Osborn, T. J., Lough, J. M., van Ommen, T. D., Vinther, B. M., Luterbacher, J., Wahl, E. R., Zwiers, F. W., Mann, M. E., Schmidt, G. A., Ammann, C. M., Buckley, B. M., Cobb, K. M., Esper, J., Goosse, H., Graham, N., Jansen, E., Kiefer, T., Kull, C., Küttel, M., Mosley-Thompson, E., Overpeck, J. T., Riedwyl, N., Schulz, M., Tudhope, A. W., Villalba, R., Wanner, H., Wolff, E., and Xoplaki, E.: High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects, Holocene, 19, 3–49, doi:10.1177/0959683608098952, 2009.

Kamenos, N. A., Cusack, M., and Moore, P. G.: Coralline algae are global palaeothermometers with bi-weekly resolution, Geochim. Cosmochim. Acta, 72, 771–779, 2008.

Kaufman, D. S.: Amino acid paleothermometry of Quaternary ostracodes from the Bonneville Basin, Utah, Quaternary Sci. Rev., 22, 899–914, 2003.

Mann, M. E. and Rutherford, S.: Climate reconstruction using "Pseudoproxies", Geophys. Res. Lett., 29, 1501, doi:10.1029/2001gl014554, 2002.

Marshall, J. F. and McCulloch, M. T.: An assessment of the Sr/Ca ratio in shallow water hermatypic corals as a proxy for sea surface temperature, Geochim. Cosmochim. Acta, 66, 3263–3280, 2002.

Meibom, A., Stage, M., Wooden, J., Constantz, B. R., Dunbar, R. B., Owen, A., Grumet, N., Bacon, C. R., and Chamberlain, C. P.: Monthly Strontium/Calcium oscillations in symbiotic coral aragonite: Biological effects limiting the precision of the paleotemperature proxy, Geophys. Res. Lett., 30, 1418, doi:1410.1029/2002GL016864, 2003.

Metropolis, N. and Ulam, S.: The Monte Carlo Method, J. Am. Stat. Assoc., 44, 335–341, 1949.

Nurhati, I. S., Cobb, K. M., Charles, C. D., and Dunbar, R. B.: Late 20th century warming and freshening in the central tropical Pacific, Geophys. Res. Lett., 36, L21606, doi:10.1029/2009GL040270, 2009.

Nurhati, I. S., Cobb, K. M., Charles, C. D., and Dunbar, R. B.: Correction to "Late 20th century warming and freshening in the central tropical Pacific", Geophys. Res. Lett., 38, L24707, doi:10.1029/2011GL049972, 2011.

Quispe Arce, J.: Variaciones de la temperatura superficial del mar en Puerto Chicama y del indice de oscilacíon del sur:1925:1992, Bull. Inst. Fr. Etudes Andines, 22, 111–124, 1993.

Rollion-Bard, C., Blamart, D., Trebosc, J., Tricot, G., Mussi, A., and Cuif, J.-P.: Boron isotopes as pH proxy: A new look at boron speciation in deep-sea corals using 11B MAS NMR and EELS, Geochim. Cosmochim. Acta, 75, 1003–1012, 2011.

Rosenheim, B. E., Swart, P. K., Thorrold, S. R., Willenz, P., Berry, L., and Latkoczy, C.: High-resolution Sr/Ca records in sclerosponges calibrated to temperature in situ, Geology, 32, 145–148, 2004.

Sato, S. I.: Temporal change of life history traits in fossil bivalves: a example of *Phacosoma japonicum* from the pleistocene of Japan, Palaeogeogr. Palaeocl., 154, 313–323, 1999.

Schmidt, G. A.: Forward Modeling of Carbonate Proxy Data from Planktonic Foraminifera Using Oxygen Isotope Tracers in a Global Ocean Model, Paleoceanography, 14, 482–497, 1999.

Schöne, B. R., Freyre Castro, A. D., Fiebig, J., Houk, S. D., Oschmann, W., and Kröncke, I.: Sea surface water temperatures over the period 1884–1983 reconstructed from oxygen isotope ratios of a bivalve mollusk shell (*Arctica islandica*, southern North Sea), Palaeogeogr. Palaeocl., 212, 215–232, 2004.

Servonnat, J., Yiou, P., Khodri, M., Swingedouw, D., and Denvil, S.: Influence of solar variability, CO_2 and orbital forcing between 1000 and 1850 AD in the IPSLCM4 model, Clim. Past, 6, 445–460, doi:10.5194/cp-6-445-2010, 2010.

Swart, P. K., Thorrold, S., Rosenheim, B., Eisenhauer, A., Harrison, G. A., Grammer, M., and Latkoczy, C.: Intra-annual variation in the stable oxygen and carbon and trace element composition of sclerosponges, Paleoceanography, 17, 1045, doi:1010.1029/2000PA000622, 2002.

Thompson, D. M., Ault, T. R., Evans, M. N., Cole, J. E., and Emile-Geay, J.: Comparison of observed and simulated tropical climate trends using a forward model of coral $\delta^{18}O$, Geophys. Res. Lett., 38, L14706, doi:10.1029/2011GL048224, 2011.

Touchan, R., Meko, D., and Hughes, M. K.: A 396-year reconstruction of precipitation in southern Jordan, J. Am. Water Resour. Assoc., 35, 49–59, 1999.

Tudhope, A. W., Chilcott, C. P., McCulloch, M. T., Cook, E. R., Chappell, J., Ellam, R. M., Lea, D. W., Lough, J. M., and Shimmield, G. B.: Variability in the El-Niño Southern Oscillation Through a Glacial-interglacial Cycle, Science, 291, 1511–1517, 2001.

Welsh, K., Elliot, M., Tudhope, A., Ayling, B., and Chappell, J.: Giant bivalves (Tridacna gigas) as recorders of ENSO variability, Earth Planet. Sc. Lett., 307, 266–270, 2011.