**Climate
of the Past
Discussions**

Interactive
Comment

Interactive comment on "Multiscale regression model to infer historical temperatures in a central Mediterranean sub-regional area" *by* N. Diodato et al.

N. Diodato et al.

giannibellocchi@yahoo.com

To the Editor of Climate of the Past

Subject: reply to Reviewer #1, manuscript cp-2010-91

Dear Editor,

The authors have made a number of important changes to the text, revised on the basis of the Reviewer's report. Point by point response to Reviewers' comments follows.

We believe that the manuscript has been considerably improved by the insightful comments and suggestions of the Reviewer. After addressing all criticisms, the text con-

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

tains evidence of progress in the manuscript.

Sincerely, The authors

Interactive comment on "Multiscale regression model to infer historical temperatures in a central Mediterranean sub-regional area" by N. Diodato et al.

Anonymous Referee #1 Received and published: 21 January 2011

Reviewer. The manuscripts presents a method to estimate regional temperatures in the past centuries based on a documentary records and previous large-scale reconstructions based on proxy data.

Authors. Yes, this is correct interpretation of our manuscript.

R. I found the manuscript quite confusing. I had to struggle to understand its overall structure and some technical details. Perhaps the English usage, which often seems strange, would not help the reader to understand what the authors are trying to say. I have the impression that the manuscript has not undergone a careful final reading – there are some unfinished sentences, and the language is sometimes imprecise. In my opinion, the manuscript needs an extensive revision to clarify the content of the study.

A. Evidence of improvements is documented in the revised manuscript, after addressing the criticisms raised during the review.

R. The abstract is my opinion not very informative. It contains very general sentences that are not needed in an abstract (' To reconstruct sub-regional European climate over the past centuries, several efforts...), too detailed technical information (the autocorrelation of the residuals) , and on the other hand does not include which are the large-scale reconstruction used or a basic description of the statistical model.

A. The authors agree that, in the original submission, the first two sentences were too generic and have been omitted. Technical information about autocorrelation has been left out as well. Few more details about the model structure have been added (lines

15-27).

" Abstract

This paper has exploited, for Southern and Central Italy (Mediterranean Sub-Regional Area), an unprecedented historical dataset as an attempt to model seasonal (winter and summer) air temperatures in pre-instrumental time (back to 1500). Combining information derived from proxy documentary data and large-scale simulation, a statistical downscaling approach in the form of multiscale–temperature regression (MTR)–model was developed to adapt larger-scale estimations (regional component) to the sub-regional temperature pattern (local component). It interprets local temperature anomalies by means of monthly-based Temperature Anomaly Scaled Index in the range -5 (very cold conditions in June) to 2 (very warm conditions). The modelled response agrees well with the independent data from the validation sample (Nash-Sutcliffe efficiency coefficient >0.60). The advantage of the approach is not merely increased accuracy in estimation. Rather, it relies on the ability to extract (and exploit) the right information to replicate coherent temperature series in historical times. "

R. The introduction uses often the term modelling in a potentially confusing way. I guess that most readers of Climate of the Past would interpret the word 'modelling' as climate modelling, unless sit is specified, or it becomes clear from the context, that the authors are referring to statistical modelling. However, the introduction does not clearly specify what type of modelling the authors are referring to.

A. This is now specified towards the end of Introduction (lines 95-100).

" In this study, we have considered an alternative approach to address the statistical modelling of temperature variability, based on documentary records and previous large-scale reconstructions. In particular, a documentary-based technique was developed based on multiscale temperature regression (MTR)–model at sub-regional level. An area covering Southern and Central Italy and named in this paper Mediterranean Sub-regional Area (MSA) is the focus of the investigation. "

R. "al., 2005; Jones et al., 2009). However, as pointed out by Riedwyl et al. (2009), the issue of downscaling to small spatial and temporal scales has become a priority in order to achieve a better understanding of sub-regional climates. Brewer et al. (2007) investigated tree-ring sites to support the reconstruction of historical droughts in Mediterranean areas during the last 500 years. However, temperature series have not been modelled for this region so far. Moreover, continuous and homogeneous instrumental series cannot be extended before the 19th century (Camuffo et al., 2010). On the other hand, high-resolution climate information is increasingly needed for the study of past, present and future climate changes (Vrac et al., 2007)" The text flow changes suddenly in the middle of the paragraph. It starts with the issue of regional reconstructions, and turns to the the issue of existing proxies and which variables have been reconstructed.

A. The text was re-arranged. Issues about regional and sub-regional reconstructions have been grouped before talking about existing proxies and variables to be reconstructed (lines 64-74).

" The issue of downscaling to small spatial and temporal scales must be made a priority in order to achieve a better understanding of sub-regional climates (Riedwyl et al., 2009). Documentary proxies' investigation remains a reliable approach to trace back the temperature extremes before the advent of instrumental recording of meteorological data (Brázdil et al., 2005; Jones et al., 2009). Brewer et al. (2007) investigated tree-ring sites to support the reconstruction of historical droughts in Mediterranean areas during the last 500 years. However, temperature series have not been modelled for this region so far. Moreover, continuous and homogeneous instrumental series cannot be extended before the 19th century (Camuffo et al., 2010). On the other hand, high-resolution climate information is increasingly needed for the study of past, present and future climate changes (Vrac et al., 2007). "

R. "tions across European land areas back to 1500. In particular, Luterbacher et al. (2004) developed separate multiple regression equations between each leading princi-

pal component (PC) time series of the proxy records and all the leading PC time series of the instrumental data" In my understanding Luterbacher et al. applied the regression equations with the in-strumental PCs and predictands and the proxy PCs as predictors. It would be more adequate to say that 'the regression equations between each principal component of the instrumental data and all leading principal component of the proxy records

A. The Reviewer's understanding is correct. The text was changed accordingly (lines 48-50).

" In particular, Luterbacher et al. (2004) developed separate multiple regression equations between each principal component (PC) of the instrumental data and all leading PC of the proxy records. "

R. "Several authors such as Luterbacher and Xoplaki (2003), Pauling et al. (2003), and Ge et al. (2005) suggested that pre-modern instrumental weather indices may be promising to enrich climate reconstructions at regional or local scales. DiïňÃerent sets of..." This paragraph is also confusing. It mixes different questions: the possibility to reconstruct past climate from documentary indices- this is not related to the small vs large scale issue; the empirical relationships between proxies and temperatures, which is also a general requirement, not specific of the problem at hand; the possibility to use a linear method or a non-linear method – also not specific to the regional problem; and finally the local correlation between proxy and temperature and the estimation of uncertainties. I honestly do not see the connection between all these issues

A. Previous paragraph was rearranged according to a new, logical outline (lines 75-94): 1) importance of pre-modern instrumental weather indices for climate reconstructions (with no reference to spatial scales), 2) examples from the international literature, 3) limitations in the approaches used in the cited studies, 4) the strength (and gaps) in current evidence of using documentary sources. The latter opens up to the mode in which this study is introduced (lines 95-97).

" Several authors such as Luterbacher and Xoplaki, (2003), Pauling et al. (2003), and Ge et al. (2005) suggested that pre-modern instrumental weather indices may be promising to enrich climate reconstructions. Different sets of proxy-variables have indeed been used to find out relationships between predictors and predictands in high-resolution climate time reconstructions (e.g. Wang et al., 1991; Briffa et al. 2002; Larocque and Smith, 2005; Moberg et al., 2005; Diodato, 2007; Davi et al., 2008). Many of these reconstructions depend on empirical relationships between proxy records and climate data. Comparing linear algorithms and neural networks, Helama et al. (2009) proved that both the approaches are reliable for temperature reconstruction. Although regression-based techniques have been used with considerable success for climate reconstructions, they can engender bias in the estimates if not employed with care (Robertson et al., 1999; Moberg et al., 2005; von Storch et al., 2005). Moreover, these relationships are seldom based on a training process capable to capture all the possible data combinations that occur when extrapolation is performed (i.e. reconstruction period). With reference to dendroclimatological studies, correlation between tree-ring proxies and temperature data was found to only explain about 50% of the (Liang et al., 2008; Helama et al., 2009; Tan et al., 2009). Documentary data series are expected to better correlate with temperature, the overall explained variance being of about 70% (Leijonhufvud et al., 2008; Dobrovolná et al., 2010). However, there are few estimates of uncertainty in documentary based climate reconstructions (Moberg et al., 2009). In this study, we have considered an alternative approach to address the statistical modelling of temperature variability, based on documentary records and previous large-scale reconstructions. "

R. "In this study, we have considered an alternative approach to address the training- and extrapolation issue. In particular, a documentary-based technique was developed" The 'training and extrapolation issue' has not been mentioned before in the manuscript. It remains unclear what the authors would be addressing

A. This "training and extrapolation issue" was replaced by the "statistical modelling of

temperature variability" (lines 95-96).

R. "Regional temperature data (hereafter called TR) were derived from Luterbacher et al. (2004) for Europe over 1500–2002. The data, upscaled at about 0.25-degree grid resolution (â′Lij35–50 km) from historical instrumental series and multi-proxy data" The original Luterbacher et al reconstructions are on a 0.5x0.5 grid

A. Right, the authors have corrected their blunder (line 122).

R. "duce reliable outcomes (i.e. time-series reconstruction). Two distinct climate periods (1867–1903 and 1972–2002) were included in the calibration dataset (68 records in The authors mean 68 time years, not records. There are several time series over this period and the manuscript deals with summer and winter means"

A. Right, "records" was replaced by "years" (line 146).

R. "Information held in the written documentary sources was extrapolated to derive temperature related indices. DïïnËĞAËŸerent types of indices have been proposed in historical" I think 'extrapolation' is not the right word here and its use here is also confusing

A. The term "extrapolated" was replaced by "extracted" (line 164).

R. The description of the method to derive the index is not complete. The authors write that their method deviate from the standard 7-level scale, but they do not specify in which sense their method is different, and why it captures extreme months. better

A. The look-up table method has a geometric interpretation, represented by an additional figure (Fig. 2). The variability around the average temperature of the MSA is shown, which explains the codes used to represent temperature anomalies. The text now comments on this, also based on geographical, climate factors (lines 184-208).

" Fig. 2. Geometric interpretation of monthly values of the Temperature Anomalies Scale Index (TASI) for winter and summer (see Table 1a for details). Black line: mean

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

seasonal temperatures; red lines: reference values for positive temperature anomalies; blue lines: references values for negative temperature anomalies.

The geometric interpretation of the classification process is shown in Fig. 2. The asymmetric profile for winter and summer seasons is a bi-dimensional simplification based on observations and documentary-proxy data. For the study-area, positive (red line) and negative (blue line) temperature anomalies result asymmetrically arranged around the mean seasonal values (black line). The latter are long-term average temperatures calculated, for the study-area, from the European database of Luterbacher et al. (2004). In the case of negative anomalies, the baseline is the freezing point of water (0 °C). A baseline for all seasons was not set to reproduce positive anomalies. In this case, in fact, temperature extremes are dictated by the Mediterranean latitudes. Although this region presents a twofold climate regime, where both tropical and mid-latitude aspects play a role, the latitudinal radiative flux stands out as the main factor determining the temperature. Advective transport off northern Africa can also occasionally affect the Mediterranean, but the seasonal variations are well marked (e.g. Schiano et al., 2000; Lionello et al., 2006) and, notably, temperatures in winter are never as high as summer values. Negative anomalies were assigned to cover the gap between the mean value and the freezing point, which is only sporadically (or never) approached in summertime (N/A). In winter (December, January, and February), values of -1 (cold) / +1 (warm) and -2 (very cold) / +2 (very warm) are consistent with temperature values deviating up to three and four times the standard deviation, respectively. Abrupt jumps from "very cold" (-2) to "freezing" (-4) in winter are due to the lack of appreciative intermediate states during the calibration period. In the case of positive anomalies, a similar scheme is reproduced for summer season (June, July and August). Negative anomalies are instead doubled (July-August) or tripled (June) compared to winter, because most evidence of "cold" and "very cold" conditions in the historical sources only refers to cooling to temperatures well below the seasonal mean. "

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

R. "These classes were allocated by an asymmetric matrix in order to take into account temporal" I do not understand what the authors mean by 'asymmetric matrix' in this context. Why is a matrix - as a mathematical object needed to allocate the classes ?. Perhaps the authors mean 'asymmetric table' as shown in Table 1 ? For a matrix to be symmetric or asymmetric, it should be first a square matrix, which is not the case of Table 1. I also have problems with the definition of the categories and of their index coding in Table 1. I agree that 'freezing in June' is indicative of a much colder anomalies than in 'freezing in March' for instance. But why is 'cold in June' linked to a more negative anomaly than 'cold' in February'? Arent the categories defined as deviations from the normal ? By the same token, 'warm in January' should be linked to a higher value of the index than 'warm in July'. All in all, the table seems to be not justified, or seems to be the result of a quite subjective assessment

A. The term "matrix" may indeed be ambiguous, and "table" replaced it (lines 178 and 183). The asymmetry between winter and summer, as well as between positive and negative anomalies, is now represented in Fig. 2.

R. "In some experimental situations, it is possible to measure more than one response for each case. This is also the case of temperature, which needs multi-scale predictors" Again, I am confused. What is an experimental situation (a measurement ?) , what is a 'response' and what is a 'case' ?

A. The term "experimental" was eliminated. Regional versus sub-regional temperatures is our case-response issue. The text was rearranged for clarification purposes (lines 217-230).

" In this study, regional temperatures (case) from Luterbacher et al. (2004) are the basis for modelling sub-regional temperatures (response). In this situation, it is possible to have more than one response for each case. Thus, a central problem in the analysis of multiresponse situations, is finding a function that combines several responses to determine more realistic estimates. This is also the case of air temperature, for which
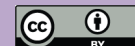
multi-scale predictors are needed to model over different space- and time-domains (after Bates and Watts, 2007). In this way, the information collected (regional temperature data) was downscaled to reasonably approximate the behaviour of the disturbance terms (or stimulus variables) driving the temperature measurements at sub-regional scale. These approximations reside on the general assumption that sub-regional air temperature depends on two disturbance terms: regional-synoptic forcing and local weather conditions. The regional scale can drive the general temperature trend, while area-specific temperatures are met by local conditions. Weather variables and climate indices were both used as predictors as basis of the multi-scale regression model. "

R. "In some experimental situations, it is possible to measure more than one response for each case. This is also the case of temperature, which needs multi-scale predictors to be modelled over diïňĂerent space- and time-domains. In the analysis of these ex-" I do not see why temperature should need multiscale predictors ? If a thermometer happens to be located on the spot, one would need just the information provided by this thermometer, and nothing else.

Our approach is a multi-scale model-based method. It uses the Luterbacher-based component on a regional scale and local conditions represented by documentary-proxy data on a sub-regional scale. Fig. 1a documents the lack of predictors in our study-area (after Luterbacher et al., 2004). This is why regional estimates have been improved by integrating the information conveyed by local sources. Fig. 1b (a new format is provided) shows winter-temperature correlation patterns between a hypothetical thermometer placed in Northern Italy and others placed elsewhere in central Mediterranan Europe (lines 121-128). The correlation is visibly degrading from North to South as result of lack of predictors.

" Fig. 1. a): Geographical setting of the Mediterranean Sub-regional Area (MSA, squared) with the location of temperature sites (red circles), and documentary monthly-resolved data (blue dots) used by Luterbacher et al. (2004) to reconstruct the regional seasonal temperatures over Europe since 1500 AD; b): Winter temperature

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

correlation patterns (values rendered in white are not significant, p>0.05) between one grid-point of Northern Italy (46° North, 12° East) and grid-points over central Mediterranean Europe (the MSA is squared), as processing by Climate Explorer with E-OBS version 3.0 gridded dataset (http://eca.knmi.nl/download/ensembles/ensembles.php) for the period 1950-2010; c); Winter temperature pattern averaged over 1961-1990 in the MSA, as arranged by LocClim FAO software at 10-km resolution (http://www.fao.org/sd/2002/EN1203a_en.htm).

Regional temperature data (hereafter called TR) were derived from Luterbacher et al. (2004) for Europe over 1500-2002. The data, upscaled at about 0.5-degree grid resolution (∼50 km) from historical instrumental series and multi-proxy data (http://www.ncdc.noaa.gov/cgi-bin/paleo/eurotemp.pl), covers an area extending from 25° West to 40° East and from 35° to 70° North (Fig. 1a). From this map and from that depicted in Fig. 1b, it is also possible to observe the temperature-data missing over Southern Europe (including the MSA), as suggested by both data-density and correlation pattern. "

R. "In the analysis of these ex-periments, information from all the collected responses can be combined to provide parameters that are more accurate and, in turn, determine more realistic tempera- ture data (after Bates and Watts, 2007). In this way, the information collected was downscaled to reasonably approximate the behaviour of the disturbance terms in the temperature measurements. These approximations reside on the general assumption 5" I cannot understand this paragraph. What are the 'experiments' ? what are the disturbance terms ?(no statistical model has been laid out so far). why should the information be downscaled ? what is the scale of this information ?

A. The text was rearranged for clarification purposes (lines 217-230, see above).

R. "the disturbance terms have a fixed, unknown variance-covariance matrix for different responses. A model was written along this path, assuming M responses (measured

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

on each of N experimental runs) and dependence on P parameters, ïĄś, as referred to" what are the experimental runs ? The description of the statistical model is unnecessarily convoluted and disorganized. I guess that one of the indices denotes time, the other is some spatial index, but I am not sure. The terms used here are also unnecessarily unclear. Does 'stimulus variables' mean simply predictors ?

A. As these mathematical details (meant to describe the background used to develop our statistical model) are not essential in the context of this paper, this part is now simplified (lines 233-248). Interested readers may refer to the seminal literature.

" A statistical model of sub-regional temperature estimation was created with aims of prediction and explanation. For prediction, the model structure was generated based on Box and Draper (1972). In particular, a determinant parameter-estimation criterion for multiresponse data was derived upon the primary assumption that the disturbance terms of different cases are uncorrelated. A corollary assumption was that, in a single case, the disturbance terms have a fixed, unknown variance-covariance matrix for different responses. A model was written along this path, assuming multiple responses and dependence on a set of parameters, as referred to by Bates and Watts (2007): the temperature random variable is a function depending on some predictors by a set of parameters, and assuming the sum of the errors equal to zero. To contribute to the aim of explanation, influential predictors were identified and insight gained into the relationship between the predictors and the outcome based on climate history and modelling background. In this path, the temperature random variable comprises predicting variables at regional, $(.)R$, and sub-regional, $(.)SR$, scales (Fig. 3). Once regional and sub-regional components are identified, one can estimate the relationship between expected temperature and predictors. "

R. In equation 2, the authors write that the vector of temperatures Y contains the parameter matrices. I am completely lost here. Y is the multivariate temperatures , as stated in equation 1. How can it contain now the parameters matrices ?

A. This part is now simplified, as above.

R. The authors should define much more clearly the meaning of all symbols used in equations 1 and 2. X is in equation 1 'experimental settings' and in equation 2 'the actual data'. Y is in equation 1 the temperature random variable; in equation 2 it is the model estimate. .

A. This part is now simplified, as above.

R. I am sorry, but the description of the method and of the statistical model is confusing to the extreme. From what I could understand, I can only hypothesize that the method is based on some type of recursive fitting of something as a function of something.

A. This part is now simplified, as above. The approach employs a recursive procedure to calibrate the model (from line 259 down).

R. "ters were estimated using an iterative, knowledge-driven approach to bias correction steps (after Box et al., 1978). For instance, after a first run, it was found that regional temperatures (TR) were increasingly biased over historical times. Likewise, Mann et al. (2000) found a decreased number of spatial degrees of freedom in the earliest regional inferences (associated with significantly decreased variance). To account for" Why is a reduction of degrees of freedom associated with a bias in the estimation? I would understand that it was associated with an increase in the variance (not decrease, as the authors write).

A. Some parts of the text were rearranged to clarify these issues (lines 269-276). Biased and imprecise estimates in historical times were due to the regional predictor (TR), overweighed in our first run. Iterative fitting of the data with successive downweighting (square root of TR) allowed for correcting the bias initially observed and capturing the full range of sub-regional scale variability.

" For instance, after a first run, it was found that regional temperatures (TR) introduced increasingly biased and imprecise estimates over historical times. Likewise, earliest

regional inferences in Mann et al. (2000) tended to be associated with decreased performance. To account for this non-invariance over the historical time-scale, a power law was assigned to TR with the exponent forced to be lower than one (and finally set equal to 0.5) to rebalance internally the quality of calibration. Such iterative fitting of the data allowed for correcting the bias initially observed and capturing the full range of sub-regional scale variability. "

For Southern Europe, where winter temperatures are more variable than Central Europe, regional estimates suffer from reduced precision. In summertime, estimated and observed variances are similar. However, correlation is weaker in Southern Europe between regional estimates and observed temperatures (correlation and Nash-Sutcliff coefficients of Table 2, new format; lines 313-323).

" In contrast, the regional model by Luterbacher et al. (2004) poorly reflects the variability of actual winter temperature in both seasons (circles in Fig. 5a), as also confirmed by the correlation coefficient and the Nash-Sutcliffe efficiency values (equal to 0.26 and –0.43, for winter, and 0.50 and –0.30 for summer, Table 2, validation dataset). In wintertime, regional estimates suffer from reduced precision in Southern Europe where temperatures are more variable than Central Europe. In summertime, when estimated and observed variances are similar, most assessments of the poor performance of regional estimates focus on the weak correlation with observations (Fig. 1b).

Table 2. Performance and autocorrelation statistics for (MTR)–model (Eq. 1) at the calibration and validation stages. Performance values over the validation set are also reported for the regional simulations. "

R. The rationale of equation 4 is very unclear. why is the regional mean temperature square rooted in the first term and not in the second term?.

A. TR appears to return a direct, non-linear effect (square root), and combined with the sub-regional anomalies, which correct the bias observed in the historical times (lines 279-282).

" TR appears in both the square root (power of 0.5) and linear term. In the first case, it returns a direct, non-linear effect, while in the brackets it crosses the sub-regional anomalies identified by the TASI to correct the bias observed in the historical times. "

R. Is omega the winter and summer regional climatological temperatures ?

A. It is a shift parameter of TR, which represents climatological boundary conditions for both seasons (line 287).

R. Why are the index anomalies added to the regional mean temperature?

A. To interpret the joint action of local and regional components, the summed TASI is combined linearly to the regional estimate (lines 284-288).

R. This would be roughly reasonable if the categories to define the index had been defined as deviations from a regional mean for each month., but it seems to me that the categories (e.g. cold) are defined as deviations from longterm mean, independently of whether the regional mean is also cold or warm. In that case, equation 4 would be adding the same information twice.

A. Monthly categories are defined as deviations from a regional mean in each season. See the new figure (Fig. 2).

R. Also, equation 4 is not units-consistent. The left-hand-side has temperature units (I guess). The first term on the right-hand-side has units of squared temperature, whereas the other terms on the right-hand-side seem to have units of temperature. The model would then produce different results when using different temperature units (e.g. Kelvin or Celsius).

A. Scale parameter k=1 ($°C2$) was added to assure unit consistency (lines 277-279).

" The scale parameter k ($°C2$) was initially set equal to one and, for reasons of parsimony as by Grace (2004), not treated as a free parameter because the initial value resulted in a fit that satisfied the criteria outlined above (Eq. 2). "

R. What are the units of beta and omega, given in the following paragraph ?

A. Beta is unit-less, omega has the same unit of TR (°C).

R. how can uncertain ranges in the estimation of these parameters be calculated? what is the uncertainty range in the final estimation of regional temperature?

A. The calibration process did not formally account for parameter uncertainty, because the iterative process employed achieved parameter estimation in more steps. Indeed this is a critical issue, now discussed with detail (lines 357-375).

" The scope of our modelling approach and model parameterization was restricted to capturing the temporal variability of seasonal temperature data in the study-area, and some limitations of the methodology should be acknowledged. Uncertainty ranges in the estimation of parameters were not formally accounted because parameter estimation was achieved in more steps, which makes confidence bounds for model parameters not easily quantifiable. The model error (mismatch between the observed and the modelled value) is however an indication of total model uncertainty (e.g. Shrestha and Solomatine, 2008), and Nash-Sutcliffe efficiency values of 0.6 can discriminate between bad and good performances (e.g. Lim et al., 2006). The efficiency values obtained in the validation stage (>0.8) thus indicate limited model uncertainty; likely associated with narrow parameter uncertainty. Since the results of model calibration were satisfactory, the robustness of the solution was relied on and sensitivity analysis was not added to the study. The reconstruction of temperatures series has thus used generic optimized parameters, which are crude estimates over multiple years. This ensures a generic representation for the MSA, with evidence of improved performance compared to previous estimates. Since geographical locations have characteristics that require specific model structures and local optimization, then the application of the model to other sub-regions may be limited by the ability to provide representative drivers and parameter values. "

R. "the calibration dataset, indication of possible correlation is produced at $0.01 < \ddot{A}$ą

< 0.05 significance level for winter only (Table 2). This may be due to some internal constraint in the calibration stage, probably related to the fact that winter temperatures in the regional dataset and model outputs are more similar in recent times (the period of years used for calibration) than it was in historical times. However, both calibration" I do not see how this could explain the presence of autocorrelation in the residuals.

A. Few lines and citations were added to clarify this issue, interpreted as the result of a functional misspecification problem (lines 337-349). Our tentative explanation is that the model likely represents some redundancy in the explanatory variables for the years used for calibration.

" The existence of the autocorrelation can be understood as the result of a functional misspecification problem (e.g. Green, 2003). This aspect is similar to the multicollinearity problem in linear regression, usually dealt with separately from autocorrelation, but also examined by its autocorrelation effect in the error term (e.g. Ramsey III et al., 2001). In our case, autocorrelation may be due to some internal constraint in the calibration stage, probably related to the fact that winter temperatures in the regional dataset and model outputs are more similar in recent times (the period of years used for calibration) than it was in historical times. The calibration dataset is from recent times (covering periods around the 20th century), when estimates from Luterbacher et al. (2004) better approach observed temperatures. Under such conditions, the model likely represents some redundancy in the explanatory variables that means, other predictors than the regional temperature component might not be effective in improving upon the sub-regional estimates. "

R. "The multi-scale regression approached here overcomes the inherent loss of variance in both early instrumental records and univariate least-squares calibration equations. In" This has not been shown here, as the study does not systematically analyse the variances of the reconstructions versus observations. Furthermore, the problem of variance loss in statistical reconstruction does not appear in all least-square-methods, but only in those where direct regression is used. It should not appear when total least

squares or inverse regression are used.

A. We refer here to the better ability of our model (based on regional and local components) to capture the actual variability than estimates only based on regional proxies (especially in winter). This is now discussed with details in the text (lines 295-302).

" In the temperature series supplied by Luterbacher et al. (2004), standard deviation (sd) for winter increases in more recent years, i.e. after the LIA (sd=0.96 against 0.74 for 1739-1783). This contrasts with the instrumental observations, for instance those performed by Domenico Cirillo in the 18th century (sd=1.1) and documented by the Meteorological Diaries of the Royal Society of London for the Kingdom of Naples (Derham, 1733-1734). The reconstructed series based on Eq. (1) gives sd∼1.0 for both recent and historical times. For summertime, sd∼0.6 was registered for 1739-1783 in the regional dataset, also approached by the reconstructed series. "

Please also note the supplement to this comment:
http://www.clim-past-discuss.net/6/C1473/2011/cpd-6-C1473-2011-supplement.pdf

Interactive comment on Clim. Past Discuss., 6, 2625, 2010.