Supporting Information

Vermeer and Rahmstorf (2009)

Data sources. Global mean sea level data from Church and White (1) and global temperature anomaly data from GISTemp (2) were processed as annual means, smoothed with a 15-year smoothing period ("embedding dimension") using Singular Spectrum Analysis, SSA (3) in order to reduce the extraneous impact of mostly short-term natural variability, in which we are not interested. After smoothing, data were binned into 15 year bins where not otherwise stated. This smoothing and binning was applied both to temperature and to sea level rate data. The choice of 15 years is the same as in Rahmstorf (2007) (4) and works well. It corresponds very roughly to a Nyqvist sampling period of 15 years, or a spectral (Fourier) low-pass cut-off period of 30 years. The precise choice of smoothing method is not critical to our results (see below).

Statistical analysis. The dual model fit was performed using a modified version of Eq. 2:

$$dH/dt = a\ ((T - T_0) + \lambda\ dT/dt),$$

with $\lambda = b/a$; λ was varied iteratively to maximize the Pearson correlation r of the linear fit yielding T_0 and a. This produces least-squares estimates a, T_0 and b for expressing dH/dt in terms of T and dT/dt. Then, when computing H from dH/dt by integration, an integration constant is introduced. The least-squares value for this constant is obtained by minimizing the sum of squared residuals for the fit to sea level H; note that H and dH/dt differ only in their auto-covariance behaviour, and the least-squares estimate is invariant to this. So we have also found a least-squares solution a, b, T_0, $H(t_0)$ for expressing H in terms of T and dT/dt.

Error bounds plotted in Figs. 1, 3 and 6 include the error contribution of the linear fit but not that of λ (or, equivalently, b). The autocorrelation introduced by smoothing was accounted for. The bounds represent one standard deviation and agree well with independent jackknife estimates. Standard deviations of fit parameters stated in the text were obtained by the delete-one jackknife applied to eight 15-year data bins. That of b obtained this way for the observational data (Figs. 3 and 6) was ± 1.0 cm K^{-1}. Together with a 21st Century temperature range of 1.4-6.1 K this yields uncertainties in 2100AD sea level of 1.4-6.1 cm, negligible against the other uncertainties.

In the text, Pearson correlation values are stated only informationally. We intentionally do not state confidence intervals or p-values, as these belong to *hypothesis testing*, which is not what we are doing. A p-value would only test against the null hypothesis "there is no relationship between global temperatures and the rate of sea level rise", which is not a very meaningful test since this null hypothesis is *a priori* very unlikely.

What we do here is *model intercomparison* by the Akaike information criterion, involving *a priori* two physically plausible models: the original Rahmstorf (2007) model, plausibly based on the degree-day notion of glaciologists, and the refined or dual model adding a rapid-response term $b\ dT/dt$, plausible on the basis of ocean mixed layer physics as shown by our studies here involving general circulation model output.

Artificial reservoir correction. The reservoir correction derived by Chao *et al*. (5) was used in a simple analytical form:

$$\Delta H = 1.65 \text{ cm} + (3.7 \text{ cm} / \pi) \arctan((t - 1978) / 13),$$

the integral of a Cauchy-Lorentz distribution approximating reservoir construction activity.

Ground water extraction sensitivity analysis. In the absence of a time series for this like we have for the artificial reservoir correction, we considered three educated guesses for this effect: a linear (in pumping rate) correction

$$\Delta H = - \alpha (t - 1870)^2 / \tau,$$

with $\alpha = 0.03$ cm a^{-1} and $\tau = 132$ a; and two different exponential corrections

$$\Delta H = -\alpha\tau \exp((t - 1870) / \tau) + \alpha (t - 1870) + \alpha\tau,$$

with either $\alpha = 0.00236$ cm a^{-1} and $\tau = 40$ a, or $\alpha = 0.012$ cm a^{-1} and $\tau = 80$ a.

All three are designed to start from a rate of zero and produce a high-end (worst case) rate of some 0.5...0.6 mm a^{-1} by 2000. In reality, estimates of this rate vary from 0.21 to 0.55 mm a^{-1} (6).

From Table S1 we see that the three scenarios have all very similar effects: mainly, a reduction of *a* by about 14%, while the quality of fit as expressed by *r* remains largely unaffected. Note that if the ground water extraction function were strictly proportional to sea level rise, then the effect on quality of fit would vanish exactly.

Table S1. Ground water extraction effect on estimated parameters of fit.

Scenario	T_0	*a*	*b*	*r*
None	-0.41	0.56	-4.9	0.992
Linear	-0.41	0.48	-4.9	0.988
Exp (40)	-0.44	0.49	-5.1	0.985
Exp (80)	-0.42	0.50	-5.0	0.988

r is the Pearson correlation coefficient of fit.

The change in *a* seen for these worst-case scenarios is well within *a*'s 2σ uncertainty bound, as will be the change in projected sea level rise by 2100 caused by this *indirect effect*, the propagated change in fit parameters. The *direct effect*, caused by ongoing ground water pumping over 1990-2100, will act to further reduce this; conservatively assuming it to remain constant at 0.6 mm a^{-1}, this effect will be at least 6.6 cm over this time span.

We conclude that for the present state of our knowledge the ground water extraction effect yields an uncertainty contribution to projected sea level rise, and to the fit parameters themselves, that is likely small compared to, or at worst similar to, the formal uncertainty derived from the parameter fit.

Akaike information criterion. In order to compare the performance of both the original model, and the dual model, in reconstructing sea level over 1880-2000, we first computed the (sea level, not sea level rate) residual sum of squares (*RSS*) for both fit alternatives, re-centering *H* as noted above to eliminate the arbitrary integration constant and arrive at the least-squares fit. We used the Akaike Information Criterion (7, 8) in its small-sample version:

$$AIC_c = 2K + n \ln (RSS / n) + 2K (K + 1) / (n - K - 1).$$

Here $n = 8$ is the sample size and *K* the number of model parameters, including the intercept T_0, the integration constant $H(t_0)$, and the error variance. We find:

	Original model (*a* only)	**Dual model (*a* and *b*)**
K	4	5
***RSS* (cm^2)**	1.053	0.098
AIC_c	5.11	4.75
Parameters	a, T_0, $H(t_0)$, error variance	a, b, T_0, $H(t_0)$, error variance

The preferred model is that with the smallest AIC_c value. Both models are statistically plausible, their Akaike weights being 55% and 45%. What this means is that the extra parameter in the dual model does not lead to overfitting. Model selection, as pointed out in (8), should not be based on statistics only, but also consider the physics of the process being modeled.

Akaike used in the above form assumes independently distributed errors for the models compared; its applicability becomes questionable (8) if reality differs strongly from this assumption.

We performed an independence check on the residuals both from the original and from the dual fit model using the Ljung-Box test. It was found that the original model fit had significantly auto-correlated residuals, with a Ljung-Box statistic of 8.236 for three degrees of freedom, yielding $p = 0.04$. The dual model fit, however, had its most significant Ljung-Box statistic, 5.232, at four DoF, for $p = 0.26$. Figure S1 shows both the residuals of the two fits and their respective autocorrelation functions.

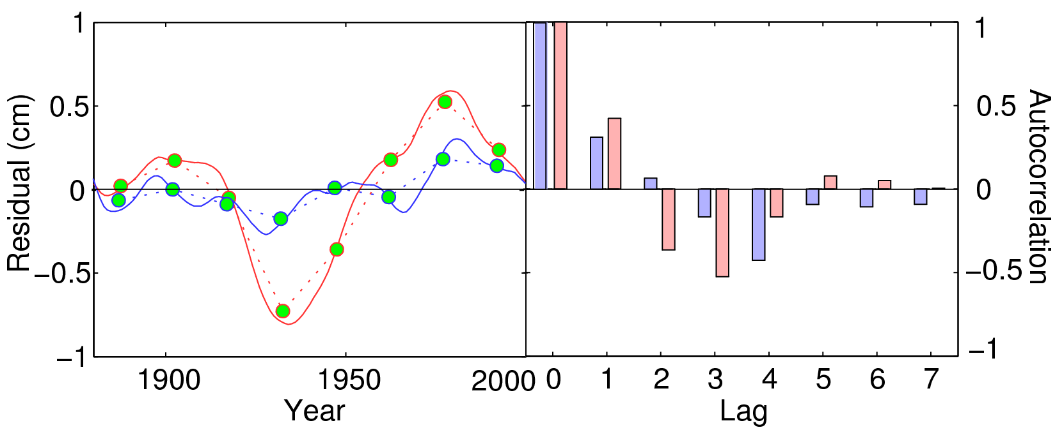

Fig. S1. (*Left*) sea level residuals after fit for the original (*a* only) model (red) and for the dual (*a* and *b*) model (blue). Drawn curves: annual residuals after recentering. Bullets and dotted lines: 15-year binned values. (*Right*) sea level residuals autocorrelations, original (*a* only) model (red) and dual (*a* and *b*) model (blue).

It may surprise the reader familiar with the long-term persistence (LTP) present in geophysical time series like global mean temperature and sea level, that we do not find significant autocorrelation here. The reason is that here we study one as a function of the other, not both as functions of time. If the LTP in both is due to the same cause, it will cancel from their interdependence.

Our choice of Akaike is conservative in that, e.g., the Bayesian Information Criterion (*BIC*) would favour the dual model even more.

Cross-validation. A further test of the robustness of our result is using part of the data to predict the part that was withheld, as in Rahmstorf (2007) (4). We did this with the observational data 1880-2000, using first the first half (only four 15-year bins) and then the second half, to predict sea level for the other half. The results are included in Fig. 3; numerical results below. With only four calibration data points, performance is naturally degraded; even so, the error in sea level in the year 2000 is only 3 cm. For a good explanation of the RE and CE statistics shown in the table, see (9).

	T_0 (K)	a (cm a^{-1} K^{-1})	b (cm K^{-1})	RE	CE
Full data	-0.41	0.56	-4.9		
First half	-0.37	0.72	-6.2	0.99	0.32
Last half	-0.55	0.42	-3.2	0.99	0.91

A more powerful cross-validation, though for model-generated data, is that shown in Fig. 2, where calibration for 120 years is used to predict sea level changes over a millennium.

Aliasing of τ and *b*. To demonstrate this effect we start from the Taylor expansion up to the second time derivative of temperature:

$$dH/dt = a\, T(t + \tau) + b\, dT(t + \tau)/dt = a\, T(t) + (a\tau + b)\, dT(t)/dt + (\tfrac{1}{2} a\tau^2 + b\tau)\, d^2T/dt^2.$$

From this we see that the expansion in delayed temperature T and temperature rate dT/dt, with coefficients a and b, is equivalent to an expansion in undelayed T and dT/dt, but with coefficients a and $b'=(a\tau + b)$, provided that the second-derivative term vanishes:

$$(\tfrac{1}{2} a\tau^2 + b\tau) = 0 .$$

This happens trivially for $\tau = 0$, and nontrivially for $\tau = -2b / a$, for which $b' = -b$.

To test this hypothesis, we produced a plot (Figure S2) of Pearson correlation values for delayed dual model fits to the observation data, on the space of (b,τ) value pairs. Inspection of the plot shows that there are indeed two equivalent maxima: one corresponding to $\tau = 0$ and $b = -4.5$ cm K^{-1}, and the other to $\tau = -13$ years and $b = +4.5$ cm K^{-1}. We propose that the latter solution is the physically realistic one: a combination of a positive b coefficient, contributed to by thermosteric ocean surface layer expansion and possibly by the surface mass balance response of smaller glaciers (10), offset by a large delay of the response masquerading as a negative b coefficient. This aliasing works only for certain ratios of a, b and τ, and only thanks to the powerful smoothing (15 years) used on the data, making further terms in the above Taylor expansion negligible.

In the four panels of Figure S2 we investigate the influence of different smoothing parameters and bin sizes, ranging from 13 to 17 years. The choice of this parameter is tightly constrained: any shorter value will make natural variability, which we are not interested in, dominant; any longer will lose information. This is a model choice, not a tunable parameter.

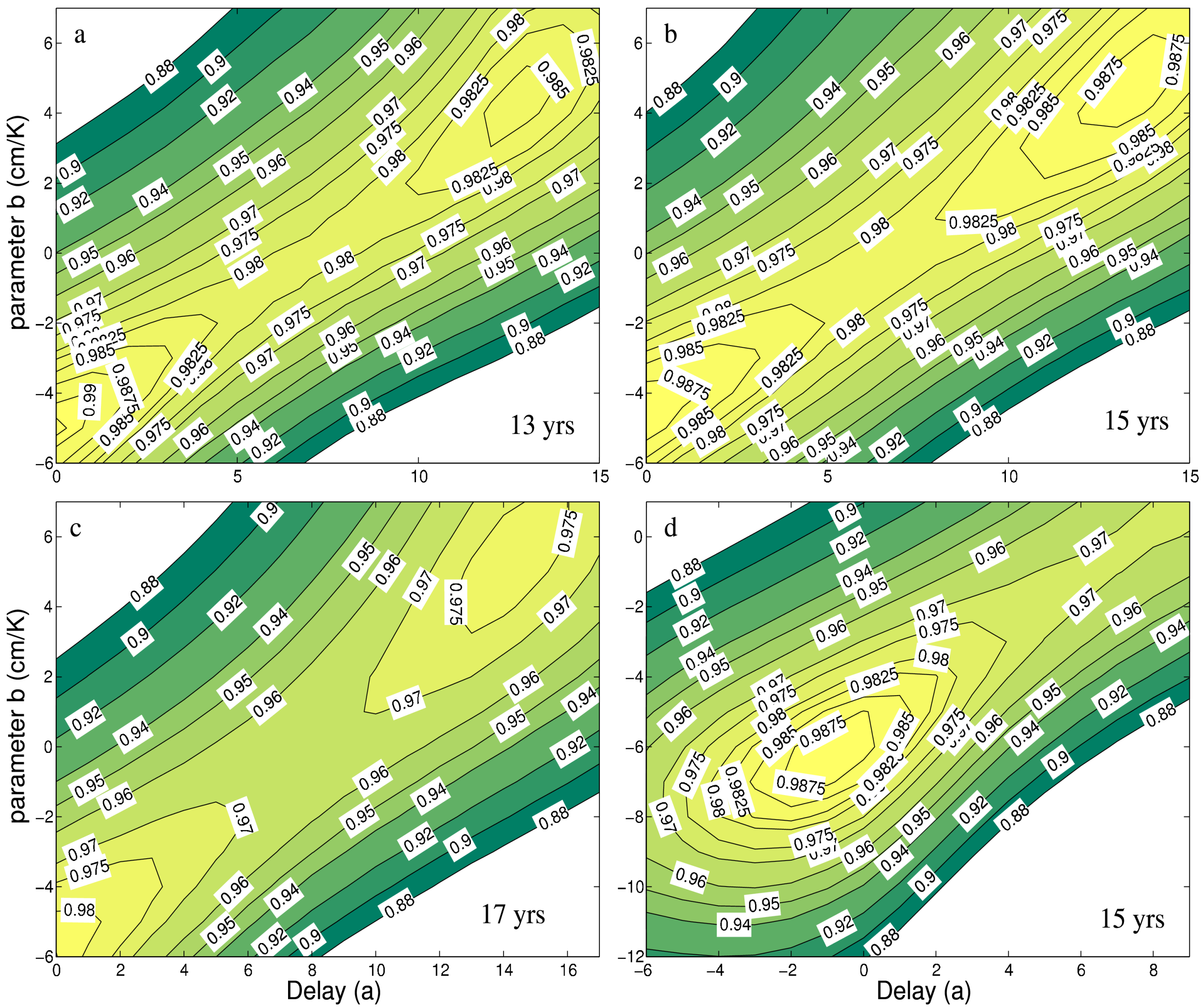

Fig. S2. Pearson correlation *r* as a function of delay τ and coefficient *b*, computed for a dual model fit. In each panel, a different smoothing/binning parameter was used as indicated. (*a*): 8 bins of 13 years, total 104 years; (*b* and *d*): 7 bins of 15 years, total 105 years; (*c*): 6 bins of 18 years, total 108 years.

Note that when doing delayed fits of this kind, it is necessary to use only a subset of the available set of 120 annual observation values. This means that the results of the fits are not directly comparable to the fit in the main paper using the full data set, or, indeed, to each other. As can be seen by visually intercomparing the panels, the details of the pattern found vary somewhat with this subset selection. However, the diagonal ridge and its central saddle point are robust against varying the smoothing parameters, and also there is no clear dependence of the location of the top right-hand maximum on the smoothing parameter value chosen. Thus, this maximum can in our judgement not be an artefact of smoothing.

For the curious reader we made a separate plot (panel *d*) including negative τ values, although we hold these to be unphysical. Note that one cannot choose a broad range of delay values for one plot as this would use only a very short subset of annual values.

Sensitivity to smoothing and binning parameters. We studied the effect of changing the smoothing parameter and binning time span on the dual model fit obtained; see table. It is clear that, although the fit parameter *b* is somewhat sensitive to changing these parameters, the projections for 2100AD are not sensitive, changing by only +/- 3 cm over a factor-2 change in smoothing and binning.

Smoothing (yrs)	Bin size (yrs)	T_0 (K)	*a* (cm a^{-1} K^{-1})	*b* (cm K^{-1})	*r*	2100AD (min B1, cm)	2100AD (max A1FI, cm)
10	10	-0.42	0.53	-3.68	0.95	79	176
15	15	-0.41	0.56	-4.91	0.99	81	179
20	20	-0.41	0.57	-5.22	0.99	83	182

Millennium fits to other models. We performed least-squares parameter fits to data generated by the models ECHO-G (11) and ECBilt-CLIO (12). Results are displayed in Figures S3 and S4.

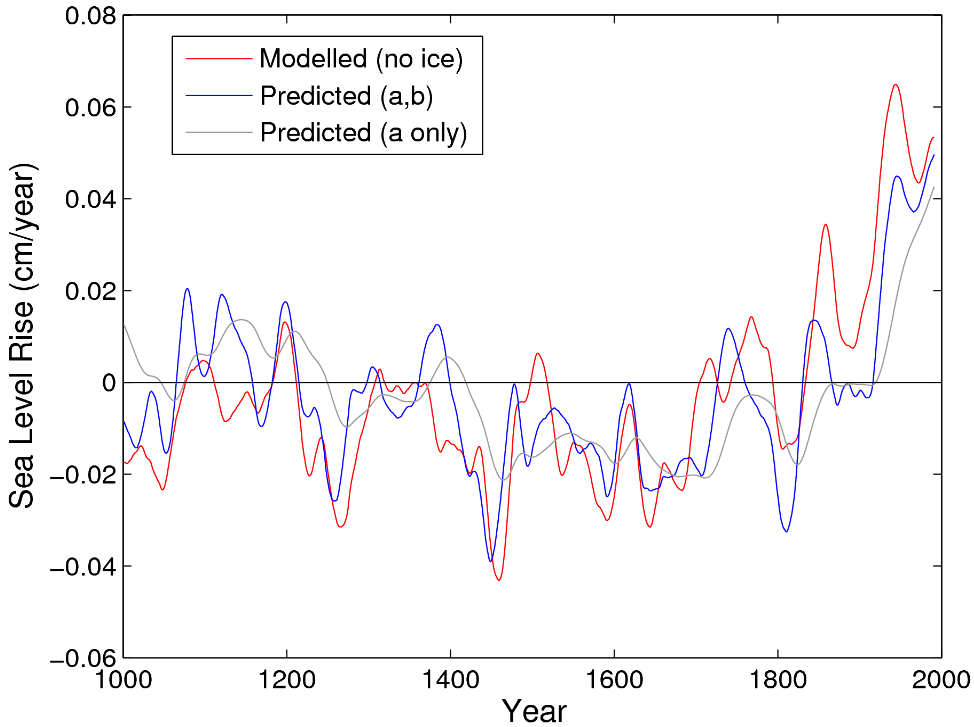

Fig. S3. Millennium fit to data from ECHO-G. Fit parameters: *a* = 0.05 cm a^{-1} K^{-1}, *b* = 2.2 cm a^{-1}, variance explained 66%.

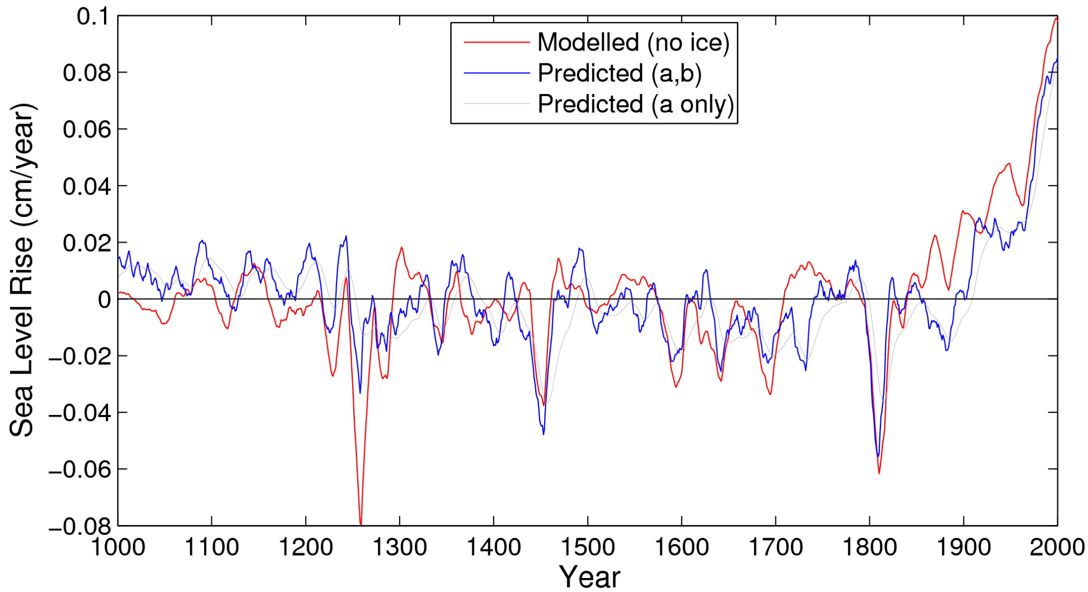

Fig. S4. Millennium fit to data from ECBilt-CLIO. Fit parameters: a = 0.13 cm $a^{-1}K^{-1}$, b = 1.8 cm a^{-1}, variance explained 65%.

References

1. Church, J. A. & White, N. J. (2006) A 20th century acceleration in global sea-level rise, *Geophys. Res. Lett.* 33:L01602.
2. Hansen, J., *et al.* (2001) A closer look at United States and global surface temperature change, *Journal Of Geophysical Research-Atmospheres* 106:23947-23963.
3. Moore, J. C., Grinsted, A., & Jevrejeva, S. (2005) New Tools for Analyzing Time Series Relationships and Trends, *Eos* 86:226,232.
4. Rahmstorf, S. (2007) Response to comments on "A semi-empirical approach to projecting future sea-level rise", *Science* 317:1866d.
5. Chao, B. F., Wu, Y. H., & Li, Y. S. (2008) Impact of Artificial Reservoir Water Impoundment on Global Sea Level, *Science* 320:212-214.
6. Milly, P. C. D., *et al.* (2009) in *Proceedings of the WCRP workshop 'Understanding sea level rise and variability'*, eds. Church, J. A., Woodworth, P., Aarup, T., & Wilson, S. (Blackwell Publishing), in press.
7. Akaike, H. (1974) New Look at Statistical-Model Identification, *Ieee Transactions on Automatic Control* AC19:716-723.
8. Burnham, K. & Anderson, D. (2002) *Model Selection and Multimodel Inference: A Practical Information-theoretic Approach* (Springer, Berlin).
9. Wahl, E. R. & Ammann, C. M. (2007) Robustness of the Mann, Bradley, Hughes Reconstruction of Surface Temperatures: Examination of Criticisms Based on the Nature and Processing of Proxy Climate Evidence, *Clim. Change* 85:33-69.
10. Jóhannesson, T., Raymond, C., & Waddington, E. D. (1989) Time-scale for adjustment of glaciers to changes in mass balance, *J. Glaciol.* 35:355-369.
11. von Storch, H., Zorita, E., & Gonzalez-Rouco, J. F. (2008) Relationship between global mean sea-level and global mean temperature in a climate simulation of the past millennium, *Ocean Dyn.* 58:227-236.

12. Goosse, H., Renssen, H., Timmermann, A., & Bradley, R. S. (2005) Internal and forced climate variability during the last millennium: a model-data comparison using ensemble simulations, *Quatern. Sci. Rev.* 24:1345-1360.