

Comment on "A Semi-Empirical Approach to Projecting Future Sea-Level Rise"

Torben Schmith,^{1*} Søren Johansen,² Peter Thejll³

Rahmstorf (Reports, 19 January 2007, p. 368) used the observed relation between rates of change of global surface temperature and sea level to predict future sea-level rise. We revisit the application of the statistical methods used and show that estimation of the regression coefficient is not robust. Methods commonly used within econometrics may be more appropriate for the problem of projected sea-level rise.

Rahmstorf (*I*) convincingly argued for the use of semi-empirical models for estimating sea-level response to future warming of the climate system. He hypothesized that the rate of global sea-level change is proportional to the global surface temperature departure from its equilibrium value. This hypothesis was statistically tested on observational data and a correlation coefficient of 0.88 was reported along with an associated *P* value of 1.6×10^{-8} and a regression slope of 3.4 mm/year per °C. We argue that this statistical analysis is based on an inappropriate application of statistics, in that the trend in both series is evident, thus violating basic assumptions of the statistical methods used. This could give misleading conclusions about inference (2) due to a spurious correlation (3) and, as such, casts doubt on the projected range of future sea-level rise.

To illustrate this problem, we reperformed the analysis of Rahmstorf (*I*), with methodological details as follows: As in (*I*), we used annually averaged global mean temperature and sea level. Nonlinear trends of both series were determined

¹Centre for Ocean and Ice, Danish Meteorological Institute, Copenhagen, Denmark. ²Institute of Mathematical Sciences, University of Copenhagen, Copenhagen, Denmark. ³Danish Climate Centre, Danish Meteorological Institute, Copenhagen, Denmark.

*To whom correspondence should be addressed. E-mail: ts@dmi.dk

as the first reconstructed component in a singular spectrum analysis (SSA) with an embedding dimension of 15 years. Before the SSA analysis, both series were extended forward and backward by linear extrapolation based on the nearest 15 years of data. The nonlinear trend of mean sea level was subsequently differentiated by calculating, at each point, the slope of a ± 5 -year least squares fit to obtain a "rate of sea level change" series. The correlation coefficient ρ between two time series x_t and y_t with deterministic time trends is defined as

$$\rho = \frac{E\{[x_t - E(x_t)][y_t - E(y_t)]\}}{\sqrt{E\{[x_t - E(x_t)]^2\}E\{[y_t - E(y_t)]^2\}}}$$

where *E* denotes expectation value. The correlation coefficient thus measures the degree to which there are coincident departures of the two time series from their respective expectation values. When estimating the correlation coefficient between filtered versions of temperature and rate of sea-level change, Rahmstorf (*I*) substituted the expectation values by the sample average. This assumes stationarity of the series (4), which is obviously violated by the two series. As an illustration, when redoing the analysis, we approximated the expectation values of the two series by the more realistic choice of a linear trend. The estimated correlation coefficient then drops to 0.68. In addition, the corresponding regression

coefficient increases from 3.3 mm/year per °C to 5.8 mm/year per °C. This nonrobust result underscores that the issue of correct statistical modeling is not an academic one and raises questions about the model put forward in (*I*).

Next, there is the point of establishing the significance of the correlation coefficient found, that is, "how likely is it to get the result by pure chance?" Rahmstorf appears to have estimated the *P* value of the correlation coefficient (1.6×10^{-8}) using a Student's *t* distribution assuming 24 degrees of freedom. The number of degrees of freedom apparently comes from assuming that the 24 bins of 5-year length are statistically independent. However, both data series were low-pass filtered in (*I*) "by computing nonlinear trend lines, with an embedding period of 15 years." Because of the autocorrelation introduced by the averaging procedure, neighboring 5-year bins can not be assumed to be statistically independent. A better approximation would be to set the number of degrees of freedom equal to the effective sample size calculated as $120/15 = 8$ (4). Using our value of 0.68 for the correlation coefficient, we get a corresponding *P* value of 0.97.

Finally, Rahmstorf used a *t* test for making inferences about the correlation coefficient. This is based on the assumption of independent and identically distributed (i.i.d.) data, yet the data analyzed in (*I*) are, due to the trend, not i.i.d. This reservation also applies to the estimated confidence interval and, therefore, the range of projected sea-level changes by the year 2100. A thorough analysis of the problem would include the application of difference stationary time series analysis methods (5), for which there is a rich tradition in the field of econometrics. Such analysis may also be helpful to other problems in climate science.

References

1. S. Rahmstorf, *Science* **315**, 368 (2007).
2. T. H. Kim, Y.-S. Lee, P. Newbold, *Econ. Lett.* **83**, 257 (2004).
3. G. U. Yule, *J. R. Stat. Soc. Ser. A* **89**, 1 (1926).
4. H. v. Storch, F. W. Zwiers, *Statistical Analysis in Climate Research* (Cambridge Univ. Press, Cambridge, 1999).
5. J. D. Hamilton, *Time Series Analysis* (Princeton Univ. Press, Princeton, NJ, 1994).

30 March 2007; accepted 10 August 2007
10.1126/science.1143286