

Interactive comment on “How unusual was autumn 2006 in Europe?” by G. J. van Oldenborgh

Anonymous Referee #1

Received and published: 4 July 2007

The goal of the manuscript is to demonstrate that the warm temperatures in Europe in autumn 2006 are a very unlikely outcome when one considers the probability distribution function of autumn temperatures in the observational record (excluding 2006). Even taking into account the generally rising temperatures in the 20th century, the autumn 2006 seems to be a clear outlier. The author also concludes that climate model simulations are not able to produce such improbable events, and therefore, some mechanisms present in nature are lacking in climate models.

In my opinion the goal of the manuscript is valuable and will contribute to increase our understanding of extreme events. I also find, however, that the manuscript can be technically improved; too frequently the formulations are imprecise, and in an analysis that is essentially statistical in nature, this raises some concerns. Perhaps more important, the manuscript virtually obviates uncertainties, and I think this is an aspect that should be considered with much more care. I try to explain my suggestions in the following.

Interactive
Comment

To estimate return periods, the data prior to 2006 are assumed to be drawn from a certain distribution - this would be the null-hypothesis against which autumn 2006 is confronted. However, this null hypothesis is formulated in a very loose manner. To assume that the data only show “interannual variability” is not a clear description of this null-hypothesis. No spectrum can only show a peak at frequencies of 1 year and be zero otherwise. The author means probably - although it is not completely clear to me - that the data are assumed to be Gaussian white noise, i.e. drawn from a Gaussian distribution and independent of one another. This process shows equal variability at all timescales, not only interannual, and the population autocorrelation function is zero. The gaussian hypothesis is hidden in the text, and it should be stated more clearly since it may be central to the results obtained. Were the distribution not gaussian, perhaps skewed to “the right”, extremes would be more probable, and the autumn 2006 might not such be clear outlier. A test for normality for the data prior to 2006 seems therefore necessary. The hypothesis of independence may also have some relevance, although perhaps a smaller one. As far as I understood, the data prior to 2006 are used to estimate a standard deviation, which is then used to estimate the return periods for the 2006 event assuming that the distribution is gaussian. However, if the data are not independent, the estimation of the standard deviation from the “usual” estimator is not correct: this would yield an underestimation of the true standard deviation (as the data are not independent, the risk of not sampling the full spread is larger). There exist some estimators of the variance from autocorrelated data, and I think that a test of the possible effect of this autocorrelation will be also helpful.

When trying to discount the effects of “global warming” (equation 1) , some additional problems may be lurking. First, no estimation of the uncertainties is given. If instead of the best estimate of the regression coefficient, one would take the upper limit of the 95% confidence interval, the return periods for the 2006 event would certainly diminish. Both probabilities- that of a higher regression coefficient and that of higher 2006 temperatures under the null-hypothesis, are coupled. But how large can the influence uncertainty in the regression be? Compounding this problem is the fact that the re-

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

gression in eq. 1 is performed with autocorrelated data, which again increases the uncertainty bounds, and by the fact that the residuals of the regression will be also probably autocorrelated, with the problem mentioned before of the estimation of their standard deviation. Furthermore, the estimation of a regression coefficient when the residuals are autocorrelated is problematic - a misspecified statistical model- and here a test a standard test for autocorrelated residuals (for instance Durbin-Watson) could be of help. Perhaps a more serious problem here is that the standard deviation of the residuals in eq. 1 is estimated from the calibration period, and this is again an under-estimation of the true standard deviation of the residuals. A better estimation could be achieved by estimating the regression in a calibration period and estimating the residuals in a validation period.

A good-of-fit analysis would be also necessary, i.e. a test of whether the relationship between global and local temperatures is linear. Indication of a non-linear fit might also increase the probability that the residual in 2006 falls, after all, within the pre-2006 probability distribution.

In summary, there exists a variety of technical issues that have not been considered, and which almost without exception could contribute to increase the probability of occurrence of the the 2006 event under the null-hypothesis. Perhaps, after all, this event remains quite improbable, but these issues should be quantified.

I am aware that all these points are interconnected and that a “clean” estimation is probably quite convoluted. I would suggest to explore some type of Montecarlo approach in which these aspects are all incorporated simultaneously. For instance, the author may want to check whether a bootstrap estimation of the probability distribution of the data prior to 2006 yields very different results - and wide confidence intervals- as by simply estimating the standard deviation of a white noise processes. For the results derived from the regression 1, perhaps a bootstrap approach for the estimation of the regression coefficient in a calibration period, following for each bootstrap sample by estimation of the standard deviation of the residuals in an independent period, could

perhaps be more correct than the solution presented in the manuscript. Perhaps some of these methods have already been use in the manuscript - see caption figure 3- but there have not been clearly explained. These are just some suggestions, and other approaches may be feasible and more accurate. A supplementary material covering these technical issues should be very helpful for the reader.

When the model gets more complex to include some regional physical processes (eq. 2-4) , the statistical problems may increase. Certainly, the predictors are autocorrelated - this widens the again the confidence interval of the regression coefficients. The underestimation of the residual variance in the calibration period becomes now a more serious issue. The variance of the residuals in calibration and validation will probably now differ more strongly, since more predictors allow for a better artificial fit in the calibration. To illustrate this potential problem, one may consider the case with 50 predictors: the residual variance will be vanishing small, and the 2006 event will appear as very unlikely, but this would be just an artifact. The physical basis of the regression model 2-4 may be debated, and other authors may exclude some of them and include others, although I in my view the choice of predictors seems reasonable. However, the skill of the model must be estimated in a period independent of the calibration. Only in this way one can judge the real validity of the predictors, and the residual variance be estimated.

I have some mixed feelings about section 5 - application to climate simulations. I think it can be very valuable to explore whether climate simulations show a change in the probability distribution of extreme local temperatures that is not related to simple changes in the mean. This is indeed shown in Figure 14, but I am not so sure that the preceding paragraphs are really necessary. Is the information about the the estimation of the trends in global, or even regional, temperature by the different models relevant here? I would delete these paragraphs and, on the other hand, expand the more interesting last paragraph in section 5 and concentrate in the changes of extreme value statistics over time within the simulations, if any. This would amount to explain more

[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)[Discussion Paper](#)

clearly Figure 14. Some of the issues raised for the observations reappear here - for instance the problem of a 'gaussian fit' for the estimation of return periods.

Some particular points: Abstract: 'current circulation' means current mean atmospheric circulation? Page 813, 'pre-instrumental reconstruction indicate' -> 'pre-instrumental reconstructions indicate Page 814, define the acronyms GHCN/CAMS Page 815, Cold Ocean/Warm Land pattern. Another factor that may contribute to enhance European temperature changes relative to the global mean is the larger warming in the Northern Hemisphere Page 815, 'the lower bound of the 95% CI' . How has this been estimated? Page 815, 'even in a linearly warming climate'. This is inaccurate, I think the author means 'in a local climate linearly related to global temperature' Page 816, 'M is a memory term of past circulation' I think it is a memory term of past local temperature. Page 816 'As, Aw, B are fixed by the interannual variability' . The author means probably that these parameters are most strongly influenced by the interannual variability. Page 817, 'in Fig. 11 the *estimated* contribution of the various terms *for the 2006* event

Figure captions. Include units in all figures. Caption 2 September-November 2006 Caption 8, 'As in Figure 3', but Figure 3 has a very different appearance. Please expand the caption fully. Caption 13. As suggested above, I would delete these paragraphs and corresponding figures, but in case they are retained, please mention the method that was used to estimate the linear trends.

Interactive comment on Clim. Past Discuss., 3, 811, 2007.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper