Hydrol. Earth Syst. Sci. Discuss., 8, C5527–C5532, 2012
www.hydrol-earth-syst-sci-discuss.net/8/C5527/2012/

Hydrology and
Earth System
Sciences
Discussions

Interactive comment on "Monitoring and quantifying future climate projections of dryness and wetness extremes: SPI bias" *by* F. Sienz et al.

R. T. Clarke (Referee)

clarke@iph.ufrgs.br

Received and published: 3 January 2012

General comment.

This paper compares results from comparing estimates of the Standardized Precipitation Index SPI derived from fitting four probability distributions (Weibull, Burr Type III, exponentiated Weibull, and generalized Gamma) regarded as alternatives to the standard Gamma distribution. These five distributions are fitted to observed calendar-monthly precipitation in (i) the England and Wales precipitation time series 1766-2007; (ii) the CRU observed high-resolution (0.5o) precipitation data set for Europe and the contiguous USA; and to (iii) simulated monthly precipitation derived from the coupled ocean-atmosphere climate model ECHAM5/MPI-OM. From the latter data-set, precipitation estimates are analysed from a 500-year control period produced from conditions pertaining in the year 1860, and from post-1860 periods under scenario A1B. Comparisons between the five distributions are achieved by means of the Akaike Information Criterion (AIC) which allows for the different numbers of parameters in the five probability distributions. The Authors conclude that Weibull-type distributions almost always gave better fits to the sequences of observed and estimated monthly precipitations than the widely-used Gamma distribution, and that SPI calculated from the Gamma distribution overestimated extreme dryness, and under-estimated extreme wetness.

The Authors' treatment is extremely thorough; the paper is very well-written, and the material is well-presented, giving a succinct account of results from a great deal of work. However in the interests of broadening the discussion on a very challenging paper, the Authors are invited to address the following points.

Specific points.

1. The SPI calculation described in the paper consists of (i) fitting a probability distribution to a set of "data" (the quotation marks being a reminder that some of the sequences analysed were precipitation totals estimated from climate models: more on this later); (ii) transforming the fitted distribution to the standard Normal distribution. Selecting a different probability distribution to be fitted as the first step obviously influences results given by the procedure, and the choice of distribution is what the paper is really about. The Authors' text says "Deviations from standard normal properties occur due to problems in the estimation procedure or, even more important, due to the wrong distribution assumption" (Reviewer's italics). This implies that there is a "correct" distribution that should be selected for a particular application, but this Reviewer would argue that there is never just one correct distribution. Whatever probability distribution is used, whether it be Gamma, Weibull or any other, it is no more than an approximation to the complex reality by which Nature produces the sequences of monthly precipitations that are measured in the field. Thus the Reviewer's first point is that all probability distributions are wrong, although some may be more wrong than others.

2. This leads to a second point. The seven categories defined by the six N(0,1) quantiles are described (Authors' Table 1) as extremely dry, severely dry, moderately dry, normal, moderately wet, severely wet and extremely wet. Thus transformed values lying between the N(0,1) quantiles -1 and +1, for example, corresponding to the cumulative probabilities 0.159 and 0.841, are described as normal; in fact the seven categories are defined by the N(0.1) quantiles corresponding to the cumulative probabilities 0.023, 0.067, 0.159, 0.841, 0.935 and 0.977. But since (see point 1 above) no distribution fitted to the "data" is "correct", why go to the trouble of selecting and fitting a distribution at all? Why not simply calculate the quantiles, corresponding to cumulative probabilities 0.023, 0.067, 0.159, 0.841, 0.935 and 0.977, of the empirical distribution given by the "data"? These would then define seven categories which could be given the labels extremely dry, moderately dry . . .. extremely wet. Much calculation would then be avoided, and the need to search for a "best" distribution would not arise. Although the Authors show that Weibull-type distributions give better fits than Gamma, other distributions may be suggested in the future which give "better" fits than Weibull. The Authors fitted their distributions whenever the number of "data" values was fifty or more. When exactly fifty values are ranked in increasing order of magnitude, the smallest would estimate the quantile of the "true" distribution used by Nature, corresponding to a cumulative probability of 0.02, close to the cumulative probability 0.023 used in the SPI calculation. Similarly the largest value would estimate the quantile in Nature's distribution corresponding to a cumulative probability of 0.98 (one in fifty values, or 2%, would be greater than or equal to xmax). This Reviewer would therefore argue that a simpler approach would be simply to label the smallest and largest values in a sample of, say, 50 January precipitations, as extremely dry and extremely wet, with the intermediate categories defined by quantiles calculated from the empirical distribution of the 50 values. When there are more than 50 values remaining for analysis (see point 4 below), the six quantiles could be calculated from the "data" in the normal way.

Whilst cases do arise in the analysis of hydrological data where fitting a probability distribution is unavoidable, the calculation of SPI is not (in the opinion of this Reviewer)

one of them. One example where distribution-fitting cannot be avoided is where a hydrologist is required to estimate an annual peak discharge with 100-year return period from fifty years of record. Here the data are too few to allow the type of quantile calculation described above, and some additional assumption is required (namely, the assumption about which probability distribution to use to represent the dispersion and other characteristics of annual peak discharge. The hydrologist's choice of distribution will never be "the correct" one, just as there is no one "correct" distribution for monthly precipitation).

3. Since the Authors compared the five distributions using the AIC, which has a term calculated from the maximized likelihood (ML), this suggests that the parameters in each of the five distributions were estimated by maximum likelihood, although this is not stated explicitly in the text. The reader would like to know more about what this involved: what criteria were used to assess whether iterative calculations had converged to a solution; whether there were cases where no convergence was found; and if so, how many.

4. The Authors' text states (page 10643, line 24) "A threshold of 0.035 mm month-1 is used to separate months with and without precipitation. Distributions are calculated if at least 50 values remain." The reason for the choice of this threshold is not clear, and this Reviewer is unsure whether the use of SPI only for months when precipitation is non-zero is entirely logical. Suppose that a site in a very dry region had 100 years of record, and that the 100 monthly precipitations in the driest month at the site had 50 zero values, the remainder being small. Applying the SPI calculation only to the 50 non-zero precipitation totals, roughly one would lie below -2 on the Normal scale, and one would lie above +2, and these would presumably be classified as "extremely dry" and "extremely wet". In the opinion of this Reviewer, such a conclusion would be questionable in this case where half the months had no precipitation, and when precipitation was small in the remainder.

5. The following sentence (page 10643, line 25) says "The observed time series are

detrended if linear trends are present." But what can be the interpretation of SPI if linear trends exist in, say, the January, February…precipitation totals? And what is its interpretation of SPIs if linear trends have been removed? The Reviewer's opinion is that SPIs can be calculated only where each of the time series of January precipitations, February precipitations… is statistically stationary. If this is not the case, one approach might be to incorporate a linear time-trend in the location parameter of the fitted probability distribution for January, February… precipitation totals, so that x in the Authors' expression (1) for the gamma distribution would be replaced by x-μ-βt, where t is the year. Then, for any particular year t0, one would consider where, for that year, the observed January precipitation lay relative to the quantiles of the fitted gamma distribution for that year; if, say, it lay in its upper 2.3% tail, it would be considered as extremely wet for the particular year t0 in the trend development.

6. A final point concerns the use of statistical methods to analyse sequences produced by climate models generally, and by the ECHAM5/MPI-OM model in particular. In the opinion of this Reviewer, to use statistical methods to analyse such sequences as if they were derived from field measurement is open to question. Using ML to estimate the parameters of a univariate probability distribution (usually) assumes that data values are statistically independent, so that their probabilities multiply to give the likelihood function L(.). But estimates of monthly precipitation obtained from climate models cannot be considered as statistically independent, since each such value depends on estimates of whatever parameters (or "fudge factors") were used in the climate model. Each value will, in general, be correlated with all other values in the sequence. Furthermore, values in the sequence are deterministic in the sense that, starting from the same initial conditions, the same sequence will always be simulated. Whether the correlations between "data" are small enough to be ignored is not known or, at least, is never stated. To produce the monthly precipitation Π at a given site, Nature will (may?) use physical laws that express Π = F(X; Θ) where X are causative variables and Θ are parameters entering "her" physical laws. An analyst estimating monthly precipitation P from a climate model will take, as his model, P=f(x; θ) where, at best, x is a subset

of X, and θ a subset of Θ, and where f(.) will certainly be different from F(.). Having estimated the parameters θ by whatever procedure, the fitted sequence will be f(x;) + ε and the simulated sequence, of whatever length, will be obtained by substituting the appropriate causative variables x in f(x;). Hence, not only are the values in the sequence correlated through their mutual dependence on , but they also fail to include the "unexplained" variation given by the residuals ε, so that their dispersion is reduced. And this is apart from the fact that the function f(.) used by the analyst is at best an approximation to F(.) (just as his/her assumption of a Weibull or Gumbel distribution as a descriptor of monthly precipitation is at best a good approximation to its "true" distribution).

7. To conclude, the opinion of this Reviewer is that the routine use of standard statistical procedures to analyse sequences of estimates given by climate model simulations requires a good deal of caution, since the characteristic statistical independence required by many such standard procedures will not be satisfied. Such issues may or may not be important, but if they are not, the Reviewer would like to be assured that the Authors can explain why.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/8/C5527/2012/hessd-8-C5527-2012-supplement.zip

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 8, 10635, 2011.