Hydrol. Earth Syst. Sci. Discuss., 9, C7026–C7042, 2013
www.hydrol-earth-syst-sci-discuss.net/9/C7026/2013/

Hydrology and Earth System Sciences Discussions

Interactive comment on "Elusive drought: uncertainty in observed trends and short- and long-term CMIP5 projections" *by* B. Orlowsky and S. I. Seneviratne

B. Orlowsky and S. I. Seneviratne

boris.orlowsky@env.ethz.ch

Received and published: 11 March 2013

Reply to the review by Dr. Henny van Lanen

We are very grateful to Dr. Henny van Lanen for this constructive and overall positive review. We appreciate particularly the detailed comments which prove a careful reading, which make us confident that any potential issues with our manuscript have been detected reliably. We put our answers to his comments in blue colour right under the

unmodified comments from his review.

1. Study is on several places too qualitative, too descriptive. A quantitative analysis should support the remarks/conclusions made. Readers cannot derive this from the maps only. This applies to: (i) intercomparison of maps on economical, physical exposure, crop land distribution, and related to population density. I believe that a quantitative measure is required to support the conclusion "are very similar" (e.g. 13778, lines 23-24).,

   Agreed. We will provide pattern correlation coefficients to quantify the similarity of the maps.

   (ii) Future projections of two drought indicators in the CMIP5 ensemble ... show that some of these hot spots are consistently projected to become even drier during the 21st century... On the other hand, some of the drought hot spots of the recent past are projected to become wetter, ... (13779, lines 8-17),

   We will provide the median change under the RCP8.5 for the last 20-year period from Figs. 7 and 8 to give the order of magnitude of the change.

   (iii) "This is further supported by the relatively consistent increases in the observational datasets (Fig. 2c)" (13780, lines 22-23),

   As a reaction on the concerns about our way of testing statistical significance, we have increased the strength of the significance level to 0.5%. As a consequence less regions show consistent signals, in particular for the observation-based datasets, and this sentence will thus be removed from the text.

and (iv) "First, the three observational datasets correlate reasonably well, although the amplitudes are less consistent". (13781, lines 24-25);

We will provide the range of correlation coefficients to quantify the similarity of the observation-based SPI-series.

2. At other places the methodology/approaches/choices need some elaboration to be understandable or to justify. This holds for: (i) motivation why you use SPI-12 and not another time step (13777, lines 16-17),

   A very good point! We will repeat the analysis for SPI3, which may be more relevant e.g. for shorter European droughts, and include the results in the Supplementary Information. This will enable us to analyse the dependence of our conclusions on the SPI timescale.

   (ii) Section 2: you need to define somewhere what you mean with "drought magnitude". It is not the deficit volume, which some readers will confuse. In your study it seems to be the standard deviation relative to a threshold, also a standard deviation, e.g. -0.5 or -1.0,

   This is right, we will clarify this in the revised version.

   and (iii) motivate why RCP8.5 was used to extend the historic time series beyond 2005 (13780, lines 19-20) and not the other more modest GHG concentration scenarios;

   For these additional few years, RCP8.5 is closest to the actual GHG concentrations (compare Moss et al., 2010; Van Vuuren et al., 2011). Although we would not expect this choice to have a big influence on the results, we wanted to have the closest comparability with the observation-based SPIs.

3. Section 2.2.1: add equations for SPI-12 where your study adds to existing literature (e.g. McKee et al., 1993; Lloyd-Hughes and Saunders, 2002) on SPI calculation. Eventually, you use the annual averages of SPI12 values and occurrence frequencies of SPI12 below a threshold (captions Figs. 3 and 4). You also used "detrended annual values" (13794, caption Fig. 1). Please make clear how (reference). Equations can be rather simple, but make the paper more transparent and calculations reproducible;

   Thanks for the suggestion. In fact we do not add so much to the established ways of calculating SPI, just the estimation of the Gamma-distribution parameters via maximum likelihood is maybe different from other studies. We will see where equations make sense and include them in the revision, in any case trying to clarify these methodological details.

4. Section 2.2.1: add reference(s) and equations how you derive the monthly SMA Monthly ("...are calculated w.r.t. the 1979–2009 monthly means and standardised by the monthly 1979–2009 standard deviations"). Eventually, you used the annual SMA averages and occurrence frequencies of SMA values below a threshold (captions Figs. 5 and 6). Equations can be rather simple, but make the paper transparent and calculations reproducible;

   Here equations are certainly useful, thanks again!

5. Your study compared: (i) annual averages of running SPI12 values with annual SMA averages, and (ii) occurrence frequencies of running SPI12 and monthly SMA values below a threshold. Add motivation why it is allowed to compare SPI12, which essentially averages over 12 months with monthly SMAs (e.g. Figs.

4-6 and Figs. 7 and 8). The temporal scales (memory) are different, even if you calculate as a last step the average annual SMA;

This is a very good remark. Our motivation for choosing the 12-month SPI is that many of most severe recent droughts globally extend over at least one year. SMAs are a function of accumulated precipitation deficit or surplus, so certainly related to SPI12, although it is true that in regions like Europe, where droughts occur on shorter time frames, SPI12 may over-smooth the drought-relevant dynamics. We will include the analysis for SPI3 in the Supplementary Information and discuss any differences in the results between the two time scales.

6. The IPCC-SREX report (Seneviratne et al., 2012) concludes that no clear general conclusions on drought can be drawn, a.o. because of confusion about definitions/indicators, tools used, intercomparison of different periods. This paper also is at some places too generic (it says “drought”, but only addresses meteorological and soil water droughts and it does not cover hydrological drought, which is of most importance for water resources management, e.g. Stahl et al., 2010; 2012);

   Agreed, we will go through the entire text and make this distinction clearer.

7. In the paper you need to make a remark that soil moisture drought indicators derived from GCM' output are not more than indicative, because of the scale (coarse grid) and conceptualization of land-surface processes. These are hard to compare against field data that have a high spatial variability;

   Agreed. This point was also raised the Jamie Hannaford and we will mention this limitation throughout the paper, notably in the Conclusions.

8. I wonder how drying trends in soil moisture can be explained by increased runoff when using GCM output that shows increased precipitation or no change in precipitation (e.g. 13782, lines 16-19). Clearly, it can happen in reality and in more sophisticated hydrological models than GCMs, but the latter only have a storage-dependent runoff in their land surface scheme, which does generate lower runoff when storage (soil water) is lower. The only reason for drying trends in soil moisture in a GCM setting with increased precipitation is increased evaporation, which is correctly mentioned as one of the reasons in the paper.;

   This comment is also made by Jamie Hannaford and correct. We will rephrase with increased evapotranspiration as the main reason for drying soils. In addition, we will highlight the relative simplicity of the LSMs in GCMs.

9. HESS stands for “Hydrology and Earth System Sciences”. In the hydrological community we like to make a difference between real observations and observation- based datasets, like the three gridded precipitation datasets (CRU, CMAP and GPCC). The paper should not avoid the term “observation”, but use instead “observation- based” to make this clear to the reader of HESS (13777, lines 2-3);

   This is a good point, we will take it into account.

10. I do not believe that the Warm Spell Duration Index (WSDI) is really needed in the paper to pass the scientific message. The paper covers sufficient aspects to support the conclusions and recommendations. The WSDI only shows that the GCM formulation never substantially contributes to the total uncertainty, especially beyond the near future. So what? WSDI is only temperature dependent, while the

SPI depends on precipitation and SMA on precipitation and the latent heat flux (not purely determined by temperature). We know that the GCMs have reasonable skills in predicting the temperature, which usually outperforms prediction of precipitation and evaporation-dependent weather variables;

It is true that the WSDI doesn't add anything to the drought indicator analysis and it is also well known that temperature projections have a much smaller inter-GCM spread than anything else. However, we find this side-by-side illustration of the uncertainty contributions in the drought indicators useful, since it provides a benchmark against which the drought uncertainty can be compared. We would therefore rather keep it in the figure, however, complement the description and discussion to make the purpose of this comparison clearer. This is also along the lines of a comment by Jamie Hannaford, who appreciated the comparison but asked for a clarifying discussion of the differences in uncertainty between WSDI and drought indicators.

11. You mention that SMA cannot be compared against observation or an observation- based product (13782, lines 27-29 and 13783, lines 1-4). However, you could at least try to compare the maps with the work from, for instance, Sheffield and Wood (2008a; 2008b, 2011), who use an independent modeling approach (offline VIC model forced with GCMs or US reanalysis data);

    Agreed, we will add more discussion of related literature and will generally try to better put our results in the context of other studies.

12. Stahl et al. (2010) (Section 3.2 in their paper) report on recent critics on the determination of magnitudes and significance of trends in hydroclimatic time se-

ries. Elaborate why your approach (13780, lines 11-15) is acceptable given the concerns about the power of various tests in the presence of auto and cross-correlation;

    The significance inference of trends in hydroclimatic time series has been debated since long and the way we have tackled with this issue is not optimal, for sure. There are different ways of considering temporal auto-correlation, but for spatial correlations things become more difficult. In the revised version we try to circumvent this issue by choosing a rather demanding level of significance (0.5%), acknowledging that this does not solve the problem per se but makes us hope that the trends passing this significance test are actually robust, despite auto- and cross correlations. Interestingly the regions of consistent signals are hardly affected by this more demanding test, while the regions of contradicting signals (grey shading in Fig. 1) turn largely into consistent-no-signal areas (white shading).

13. Global maps (Figs. 1 and 2), incl. the legends are too small. Furthermore the top of Figs. 4-6 (map with 12 selected hotspots) has to be given only once. In that case make it larger;

    We have contacted the Copernicus typesetters – in the present version Figs. 1 and 2 are small because the captions are so long. They proposed to wait until the final (two-column) layout to see if the Figs. become large enough. However, we will also rework the Figs. increasing the annotations. It's true that the hot spot map is redundant. We will incorporate it into the maps of Fig. 1.

14. SMA should not be used as an indicator of agricultural drought. It is a physical

indicator that should be called soil moisture drought. Agricultural drought includes more than only physical aspects. It also comprises economic factors (it is typically an impact indicator). You can say at the start and the end of the paper that SMA provides information for the assessment of agricultural drought;

We agree with the reviewer on this point, at the same time the term "agricultural drought" is very commonly used as synonym for soil moisture drought in the literature. We will clarify better this terminology in the revised article.

15. Section 2.3: future drought projections were used to identify hot spots (Figs. 1e and 1f). Rising GHG were considered (RCP8.5) only. Why not other GHG experiments, but more importantly you could also have used past trends, which likely are more reliable or last can be compared against observations (at least for precipitation);

    Generally the patterns agree well between the different RCPs, although the change is strongest in RCP8.5. For the selection of hot spots, we were only interested in the patterns. As you can see from the maps, even the regions with consistent trends show only rather weak changes. Changes in the past are even weaker (compare e.g. the time series in our Figs. 3 to 6), both in observations and simulations. Anyways, the selection is intended to be more representative than strict and exhaustive and Fig. 1 should help to understand why we chose the regions that we analyse. We will try to make this reasoning clearer in the revised manuscript.

16. Section 4.1: I believe it is sensible to make a remark that caution is needed to do any future drought projects using GCM simulation given the rather high uncer-

tainties, inconsistencies and low robustness in reconstructing the past (previous sections in this study);

    We totally agree and try to make this point stronger.

17. The longest common period (i.e. 1979-2009) was only based upon the observation- based data. I wonder why you did not consider the GCM simulations. The historical runs with observed GHG end in 2005. Then model experiments start. Why not 1979- 2005 as common period for the whole study? Then you do need to decide on a RCP.

    This is of course true, such a choice is always a trade-off between length of time series and other constraints such as the choice of an RCP. Given the fact that observed CO_2 emissions are closest to RCP8.5 while this doesn't say much since the RCPs hardly differ during these first four years and the choice thus doesn't matter a lot, we decided to make this choice and have an additional four years of data, which is a substantial advantage when the time series spans only 26 years otherwise.

Minor Points:

18. "However, high uncertainty should not be equated with low drought risk...". (13774, lines 26-28). Who is doing this? You expect a reference, but cannot be done in the abstract. Risk = probability on the hazard (topic of the paper) x vulnerability / exposure. Uncertainty in the hazard automatically leads to uncertainty in the risk (likely with a larger spread in the risk probabilities);

    This is certainly true for the scientific community, however, we feel that in other contexts uncertainty of a change is quickly misunderstood as ‘no change’. That’s why we wanted to emphasise this point.

19. “... in particular for several hot spot regions that are consistently projected to be more strongly affected by drought in future CMIP5 simulations (e.g. the Mediterranean, Central America/Mexico, the Amazon, North-East Brazil and South Africa).” (13775, lines 12-14). You expect a reference at the end of this sentence, or is it “(IPCC, 2012; Seneviratne et al.,2012)”?

    Thanks for pointing this out. Indeed, these projections are summarised in the SREX, which in turn draws upon e.g. Orlowsky and Seneviratne (2012). We will add these references.

20. The Introduction is too general about drought (13775, lines 1-27)). I believe that you have to state clearly that the paper is limited to meteorological and soil moisture droughts and that you do not cover hydrological drought (see Stahl et al., 2010; 2012), which could have derived from the runoff simulated by the GCMs. See point 6.;

    We agree, and this point has also been raised by Jamie Hannaford. We will clarify the notion of drought that we use in our study and provide a broader context discussing related studies such as the proposed papers by Stahl et al..

21. “In this study we address past and projected future changes in droughts from a variety of perspectives.” (13775, lines 23-24). Should be “.. changes in meteorological and soil moisture droughts”. Do not confuse the readers;

    Will be adapted.

22. Section 2.1.1 CMIP5. Not all HESS readers are familiar with “RCP” (Representative Concentration Pathways), the resolution (Table 1), or T42 grid (13779, line 12). I suggest to add between brackets (approximate degrees, e.g. 2.80, 1.10);

    Good point, will be done.

23. In Table 1 (reference to it, 13776, lines 21-22) acronyms occur that are not explained yet (e.g. WSDI). CDD (Consecutive Dry Days is mentioned there) – not mentioned in the text. Relict of earlier draft?

    Thanks for pointing this out. We will explain the acronyms.

24. Elaborate why it is permitted to inter-compare trends in SPI12 (Fig. 2) with zonal mean precipitation anomalies (other indicator) over different periods (13780, lines 26-29, 13781, lines 1-3). See point 6;

    Will be done. We think that the trends in SPI and precipitation can be compared at least qualitatively, since the sign of in- or decreases is the same in both. However, it is correct that our comparison in the paper is incomplete and e.g. does not discuss the different periods. This will be added in the revised manuscript.

25. “Using that range as an estimate of internal climate variability (which seems reasonable, since ....” (13782, lines 2-3). I believe it more appropriate to derive the internal climate variability from the SPI12 range using the observation-based datasets (each hot spot area consists of many grids that can be used for the daily temporal distribution) rather than from models that have limitations;

    For sure, GCMs have strong limitations and we will rephrase this discussion, calling the GCM spread “GCM spread” instead of internal variability (which is more appropriate, see also the comments by Jamie Hannaford). However, we find it equally problematic to estimate the internal variability from the observation-based datasets, which have quality issues e.g. due to inconsistent measurements or interpolation errors. We will be more careful, e.g. stating that single drought events don't appear as exceptional compared to the observation-based and GCM simulated timeseries instead of saying that no event exceeds the natural variability range.

26. "The few spots of systematic SMA decreases in Fig. 2 are consistent with the drying regions identified in Fig. 6". Is it also the other way around?

    Not necessarily – Fig. 6 shows decreases without regarding their significance, in contrast to Fig. 2. That's why the areas with decreasing SMAs in Fig. 2 are a subset of those in Fig. 6, but not the other way round. We will make this clearer in the revised manuscript.

27. "Only for the Mediterranean the GCM simulations seem to indicate increased drought,..." (13782, lines 9-10) add "Only for the Mediterranean the GCM simulations seem to indicate increased drought frequency,..." and "...and SAF panels in Fig. 6 for increasing drought)" (13782, line 15) add "...and SAF panels in Fig. 6 for increasing drought occurrence)". Only " drought" is not precise (see point 6). Check other general phrasing in the paper;

    Point well taken, we will take care of a clean usage of the term 'drought'.

28. "....as Fig. 8b, d shows by the number of months per year in which the SPI12 and SMA drop below -1,..." (13783, lines 24-26). Motivate why you swap from -0.5

(Figs. 4- 6) to -1.0. Is there a reference; like the mild drought characterised by -0.5 (Lloyd- Hughes and Saunders, 2002)?

    Good point, we will add the reference, SPI< -1 is termed 'moderate drought'. The reason for the two different values is that in the observational period, SPI at this regional aggregation hardly goes below -1 thus no trend-in-frequency analysis was possible. On the other hand for the future, where stronger changes are projected, a threshold of -0.5 is overly mild. We will motivate this shift in the revised manuscript.

29. Conclusions: "However, our analysis of the 12-month Standardised Precipitation Index (SPI12) indicates that the recent droughts are not exceptional in a climatological sense but are consistent with the range of internal climate variability estimated from the CMIP5 ensemble of GCM simulations" (13786, lines 9-13). Add "...indicates that the recent meteorological and soil moisture droughts are not exceptional...". Only " drought" is not precise (see point 6);

    Agreed. This statement will be further modified w.r.t. the notion of internal climate variability (see above, e.g. Point 25).

30. Conclusions: "Large internal variability and general uncertainty is also found..." (13786, line 25). Be a bit more precise on "general uncertainty" . Is it uncertainties due to GCM formulation and GHG concentration scenarios?

    Thanks for pointing this out. In fact, we refer to the total uncertainty here, and will rephrase it as such.

31. Conclusions: "Extreme drying scenarios are therefore about as likely as signifi-

cantly reduced drought risk". These cannot be compared directly (see point 18);

True. We will rephrase as '... significantly reduced drought.'

32. Acknowledgments: "We acknowledge partial funding from the EU 7th framework program through the DROUGHT-RSPI and EMBRACE projects" (13788, lines 15-16). Add grant number (mandatory, see annexes to EU Grant Agreement);

    Thanks! Will be included.

33. Fig. 2 (13795). Add legend/units under horizontal bar (is it % change?). What is the meaning of the colour of the symbols (o, + and -) in Fig. 2c?

    We will add 'No. of GCMs' to the colour bar and include the meaning of the symbol colours (blue for increase, read for decrease). Thanks for notifying this.

34. Fig. 3 (13796): ".....projections for the three GHG concentrations scenarios RCP2.6, RCP4.5 and RCP8.5 are combined". How, simply averaged?

    No, in the respective Figures for the submitted manuscript we had taken the time series from all GCMs and all RCPs together and have estimated median, inter-quartile and total range from this ensemble. However, in the revised version we will remove RCP4.5 and RCP2.6 from the analysis, keeping only RCP8.5 for consistency with the trend analysis in Fig. 2. This alters the ensemble spread of the years since 2006 only, but even these differences are barely visible and do not at all affect the conclusions from the analysis.

35. Figs. 4 and 6: (13797 13799): the period 1950-1960 is left out. Need to mention that you used a 10-yr forward moving average (Section 3.2);

    Agreed.

36. Fig. 7: "Wetting regions" (13800), add "Wetting regions in the 21st century" (makes numbers below x-axis better understandable). "Standardised Precipitation Indices (SPI12) change in "SPI12" You do not need to spell out, not done in previous graphs. "...in regions where SPI12 increases.". Median / average SPI12?

37. Fig. 8: see point 36;

    We will rephrase accordingly. We will refer to the Median for the SPI12 increases for easiest comparability in Fig. 7 and 8.

38. Fig. 9 (13802): Add y-axis with legend ("fractional uncertainty"?). "...from three different GHG concentrations scenarios (colours blue, green, red)". Add RCPs, like in Figs. 7 and 8.

    The y-axis will be added. The time series represent illustrative series from a GCM simulation, e.g. temperature or time series of WSDI. This is entirely random data for illustration purposes. But it's surely easier to stay in the context if we label the GHG concentration scenarios as representative of RCPs. Thanks for this!

References

R. Moss, J. Edmonds, K. Hibbard, M. Manning, S. Rose, D. van Vuuren, T. Carter, S. Emori, M. Kainuma, T. Kram, G. A. Meehl, J. F. B. Mitchell, N. Nakicenovic, K. Riahi, S. J. Smith, R. J. Stouffer, A. M. Thomson, W. J. P., and T. J. Wilbanks. The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282):747–756, 2010.

B. Orlowsky and S. I. Seneviratne. Global changes in extreme events: Regional and seasonal dimension. *Clim. Change*, 110:669–696, 2012.

D. P. Van Vuuren, J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. C. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, et al. The representative concentration pathways: an overview. *Clim. Change*, 109(1):5–31, 2011.