Anonymous Referee #1

The investigation of drought predictability is an important topic. Especially in regions with low resilience and strong dependency on the magnitude and duration of the rainy season it may be helpful to forecast dry conditions as this might enable water management to mitigate the impacts.

Therefore I like the idea of this study to combine existing monitoring products with a state-of-the-art seasonal forecasting system to study the predictability of droughts in different climatic regions.

General comments:

The paper needs minor revisions. In my opinion the paper could be improved in terms of the clarity of the presentation (suggestions below). Furthermore it would be nice to see, as an example, how the forecasting system would have predicted the recent drought at the Horn of Africa in the Blue Nile catchment. Another (partly related) issue is the decrease of the number of rain gauge stations contributing to the monitoring. This may cause a reduction of skill of forecasts that consists of both, a monitoring and a forecasting component. Investigating skills of these forecasts in e.g. the first and second half of the observational period only may allow to speculate about the skill of such forecasts nowadays using the current stations.

We thank the reviewer for the positive comments and suggestions. The specific comments are addressed in detail below. Concerning the two general comments:

- The recent drought in the Horn of Africa was spatially restricted to Eastern Africa (Dutra et al. 2012), and a small impact was observed in the Blue Nile region. However, we agree with the reviewer that it would be interesting and informative to have an example of a drought event, showing the evolution of the forecasts. Therefore, we have selected the 1991/92 drought in the Limpopo region that can be observed in Figure 4 (and also Figures S1 and S2), and caused significant impacts on the population in different countries in South Africa (FAO 2004). A selection of SPI-12 seasonal forecasts issued from September 1991 to August 1992 were included in the supplementary material (Figure S3) and briefly described in the text.

-The decrease in the number of rain gauges is a main limitation for the verification and monitoring of precipitation in the different basins. This reduction is present in both GPCP (used for verification) and CAMS-OPI (used for monitoring), while ERAI, and S4 seasonal forecasts are not affected. The reviewer suggestion is very interesting, but it is not straightforward to address it because: (1) the decrease in the number of rain gauges is affecting the dataset that we use for verification; and (2) splitting the time series of precipitation in two periods would generate very short time series of data (16 years instead of 32 years - note even 32 years is already short) to be transformed to SPI, therefore increasing the uncertainty of the transformation.

We have calculated the temporal correlations between the GPCP SPI and the remaining datasets for the different SPI time-scales considering the first and second half of the record (see table in the end of the document, added to the manuscript as Table 3), but keeping the full record of precipitation for the SPI calculations. The correlations between GPCP and S4L0 (see table 3) do not change significantly. For ERAI there are only significant changes in the BN, with an increase of the correlation from the first to the second half of the period. This is likely associated with changes in the ERAI precipitation due to changes in the amount/quality of data entering the data assimilation system. The main changes occur in CAMS-OPI with a decrease in the correlations from the first to the second half of the period in the NG and BN, and ZB basins, which are the basins more affected by the drop in the number of rain gauges. These results partially address the reviewer comment, but are limited to the use of the full period for the SPI calculations. Thus, we feel that a proper evaluation of changes in the skill of the forecasts would require a deeper study, in particular the impact of using 16 years of data for the SPI transformation. This would be very informative, with further repercussions, since there are other precipitation datasets based on satellite data that have shorted time

periods (for example the Tropical Rainfall Measurement Mission -TRMM) that could be used for in monitoring. These new results were added to section 3.2 and conclusions.

Specific comments:

page 11094 line 4: change "mitigate their impacts" to "mitigate drought impacts"
Changed.

line 6: be consistent with the use of capital letters for "Standard Precipitation Index"
Changed.

line 9: change "then is" to "is then"
Changed.

line 22/23: rephrase sentence; the skill is not only reduced in 2 and 4 months lead time.
We did not understood this correction suggestion

What I am missing in the abstract is information about the fact that the study is based on monthly data and regarding the lead times until which forecasting skill beyond climatology is observed.
A reference to monthly precipitation from the monitoring and seasonal forecast was added; along with the maximum extend of 5 months lead time for the seasonal forecasts.

page 11095 line 4: change "have" to "has"
Corrected

line 5: Why "also"?
Sentence changed

line 7: It would be nice to provide an example (in the results section) how well the forecasting system would have predicted the recent drought in the Blue Nile catchment located close to the impacted area.
See the above reply to the general comments.

line 17: change "referred to time-scale" to "referred to as time-scale"
Changed

line 17: change "accordingly" to "according"
Changed.

line 19-27: Before you state that this study is focussed on meteorological droughts as it uses SPI, but here you try to relate the SPI of different time scales to soil moisture, runoff and groundwater. Evidence is needed to support this. Furthermore you could test how SPI-3 and SPI-6 compare to GPCP, additionally to SPI-12 as given in Table 3. You may find that runoff is better related with shorter time scales.
We added a citation to Vicente-Serrano et al (2012a) that exemplifies the different relations of the SPI time-scales with soil moisture and streamflow. The comparison with river discharge was removed from the manuscript (see general reply), however a reproduction of table 3, including the correlations between discharge and the SPI-3, 6 and 12 months is presented in the general reply.

page 11096 line 10/11: order studies chronologically and provide some information how their results compare with yours (e.g. in the conclusions section) line 15: "such" is not needed
The cited studies were given has examples of the potential benefits of use of seasonal forecasts for different sectorial studies, but not for drought. It was removed.

page 11097 line 2: change "in terms of" to "caused by" line 7: change "Africa" to "African" lines 9/10: rephrase sentence.
Changed, and sentence removed

page 11098 lines 18-21: I do not understand this.

The sentence was changed, and more information added. ERAI precipitation is a forecast product, i.e. generated by the NWP model. Therefore, different forecast lead times can be used to calculate the monthly means. In this work we have decided to use the forecasts +24 to +48 (i.e. the 2th day of forecast).

page 11099 line 5: what is CPC?
CPC: Climate prediction Centre (was defined previously in the text). Changed to "NOAA CPC".

page 11100 lines 1/2: change "(SPI) Mckee et al. (1993)" to "(SPI, McKee et al. 1993)"
Changed

lines 13/14: If there are statistical tests to assess the suitability of a selected distribution, you could use them and provide information on the suitability of your chosen gamma distribution.
This would be an interesting topic, but we think that such evaluation is out of the scope of the current study, and the use of the gamma distribution is widely accepted.

line 25: change "advance" to "advanced"
Changed

page 11101 line 14: change "alpha a multiplicative" to "alpha is a multiplicative" line 15: change "I the calendar" to "I is the calendar" lines 13 and 17: list these equations separated from the text as you did for equation (1)
Changed

page 11102 equation 1, upper row: the "max" is not needed line 15: change "this skill score" to "the related skill score"
Changed

page 11103 Why do you use so many different verification metrics as described in Section 2.3? What is the advantage of each as compared to the others?
Each of the verification scores focus on different characteristics of the forecasts quality. In section 2.3 one extra paragraph was included describing each score and its application.

page 11104 lines 7-9: rephrase sentence line 25: "(LP)" missing after "Limpopo" lines 26: The Limpopo and Zambesi basins are located in central south Africa rather than in the east
Changed to southern Africa.

page 11106: line 4: Why SPI-12 and not e.g. SPI-3?
This section was removed from the manuscript.

lines 19/20: Can you speculate why there is lower variability in S4L0?
The lower variability of S4L0 (compared with ERAI) can be primarily attributed to the long-range integrations of the coupled atmosphere-ocean model, and the tendency of the forecasts to predict climatological conditions on long lead times.

page 11108 line 28: change "outperform" to "ouperformed"
Changed

pag 11109 line 1: change "results showing" to "results, showing" line 7: also in the NG catchment the slope exceeds 1
Changed

lines 9-12: Why is skill of SPI-3 highest after 2 months and of SPI-6 hightest after 5 months?
These lead times (2 months for SPI-3, and 5 months for SPI-6) are the first forecast lead times with no monitoring information, i.e. only forecast data (the first 3 months for SPI-3 and the first 6 months for SPI-6). The skill score (ROCSS) in figure 10 compared the ROC of S4 with the climatological forecasts (CLM). In the NG, BN, and LP, these lead times have the highest ROCSS, i.e the S4 forecasts are skilful, while the CLM has no skill (ROC scores close to 0.5).

page 11110 line 18: change "although" to "despite" line 20: change "serve" to "serves"
Changed

page 11111 line 2: Is this statement based on Table 2? If yes, than the correlation of S4L0 with GPCP does not exceed the correlation between ERAI and CAMS-OPI with GPCP in the Zambezi catchment.
It was based in table 2, and Fig. 11. The sentence was rephrased. Our point is that in those situations, it is worth to test the skill of the first month of seasonal forecast, and it might return similar or even better results than other near-real time datasets.

Table 1: change "Basins" to "Basin" in the caption.
Changed

Table 2: rephrase caption
Changed

Figure 3: What is "TP" on the y-axis?
Total precipitation

Figure 4: 3rd line of caption: remove comma after "rows"
Changed

Figure 5: Label the axis.
Added "temporal correlation"

Figure 6: 1st line of caption: change "montlhy" to "monthly"
Changed

Figure 7: Label the x-axis with months as letters or words. First line of caption: change "(ACC) the" to "(ACC) of the" Fifth line of caption: change "beta,>" to ">"
Changed

Figure 8: The caption (and text) states that the false alarm rate is displayed versus the hit rate, but in fact it is the other way round.
Caption changed to: "hit rate (vertical axis) against false alarm rate (horizontal axis)"

Figure 9: Third line of caption: Diagonal line is dotted instead of dashed. Last line of caption: change "(lines)" to "(rows)" change "has" to "as"
Changed

Figure 10: After 3,6 and 12, "months" is missing in the caption.
Changed

Figure 11: y-axis label: change "(month)" to "(months)" Second line of caption: change "(lines)" to "(rows)", furthermore you mixed up "rows" and "columns" here
Changed

*Table (3 new in the revised manuscript) . Temporal correlation of the GPCP 3, 6 and 12-month SPI and ERAI, CAMS-OPI and S4L0 (each column) for the different basins. The correlations between [c1 c2] indicate the correlations in the first half of the record (c1 - 1979 to 1994) and second half of the record (c2 - 1995 to 2010).The * indicates that the difference between c1 and c2 is statistically significant at 99% (using a Fisher transformation).*

	ERAI			***CAMS-OPI***			***S4L0***		
	3	***6***	***12***	***3***	***6***	***12***	***3***	***6***	***12***
NG	*[0.62 0.58]*	*[0.50 0.47]*	*[0.42 0.36]*	*[0.72 0.59]*	***[0.77 0.59]****	***[0.87 0.59]****	*[0.60 0.51]*	*[0.54 0.43]*	*[0.46 0.40]*
BN	***[0.53 0.70]****	***[0.43 0.74]****	***[0.36 0.79]****	***[0.77 0.63]****	*[0.75 0.62]*	***[0.75 0.58]****	*[0.68 0.65]*	*[0.67 0.70]*	*[0.68 0.76]*
ZB	*[0.52 0.42]*	*[0.50 0.45]*	*[0.42 0.58]*	***[0.60 0.25]****	***[0.68 0.29]****	***[0.70 0.33]****	*[0.43 0.33]*	*[0.49 0.40]*	*[0.41 0.43]*
LP	*[0.84 0.83]*	*[0.88 0.84]*	***[0.93 0.86]****	*[0.88 0.85]*	*[0.90 0.87]*	*[0.92 0.86]*	*[0.42 0.59]*	*[0.53 0.61]*	*[0.70 0.65]*