Interactive comment on "Hydrological drought forecasting and skill assessment for the Limpopo river basin, Southern Africa" by P. Trambauer et al.

Anonymous Referee #1

Received and published: 14 October 2014

Reaction to the interactive comment by Anonymous Referee #1

We would like to thank this referee for his/her interesting comments and suggestions that contributed to improve our paper and to clarify specific points. Hereby we present the authors reply (AC) to the referee's comments (RC).

Overview

RC: The paper describes the potential for seasonal ensemble hydrological and drought forecasting over the Limpopo River basin in southern Africa. Three methods were applied to provide the seasonal meteorological forecasts that were used to force the PCR-GLOBWB hydrological model at a 0.05×0.05 degree spatial resolution and a daily time step. A historical simulation was carried out (1979-2010) using meteorological forcings from ERA-Interim that has a horizontal resolution of 0.7 degrees that were biased corrected using GPCP v2.1 monthly climatological data provided at a 2.5×2.5 degree resolution. This reference run was used to provide the hydrological initial conditions at each forecast time. Ensembles of seasonal forecast were made based on the ECMWF S4 system (15 ensemble based hindcasts), the Ensemble Streamflow Prediction (ESP) method and ESP conditioned on the ENSO index. Forecast skill is measured using standard skill metrics (ROC, BS).

Overall the underlying work is solidly carried out and the results informative as they show for the Limpopo basin that dynamical seasonal forecasting offers the most skill in predicting seasonal drought metrics (SPI, SRI), followed by ESP conditioned on ENSO. As expected, shorter lead times and smaller regions show less skill. The biggest shortcomings in the paper is that important, but practical, details are omitted, and, in general, the written English is sufficiently weak to make the paper difficult to read – incorrect verb tenses, conditional phrases in the middle of sentences (and not off-set by commas), split infinitives, etc. As for missing details, most are identified in the specific comments below but the following seems fundamental in helping readers understand the study.

1. There is a scale difference between the meteorological forcings (and forecasts) and the hydrological model. Were the forcings downscaled in time or space? If so, how was this done? If not, what are the consequences of this for both the study and for practical implementation?

    AC: The forcings (and forecasts) are not downscaled in time as these are provided with a daily time step, which is the temporal resolution of the hydrological model. Regarding space, the ERA-Interim data for the 32-year period from 1979 to 2010, corrected using the GPCP v2.1 dataset are converted to the same spatial resolution as the continental-scale version hydrological model. ERAI is archived on an irregular grid (reduced Gaussian) with an approximate resolution of 0.7° over the domain. The data is downscaled from the ERAI grid to the original 0.5˚ hydrological model grid using bilinear interpolation and assumed to be constant over the 0.5˚ grid cell. No further downscaling of the meteorological forcings was carried out. We have added this description in P9969, L9 to clarify.

2. The description of the weightings for the ESP_cond ensemble generation is poorly described.

    AC: This description was extended to clarify, and formulas on the computation of weights and probabilities were added.

3. The number of ensembles for the ECMWF_S4 hindcasts is 15 – is that how many were used in the FS_S4 runs (never well stated)? For ESP 30 ensembles were "created". Can the authors comment on the effect of the reduced S4 ensemble size on the study results? I think that if the paper is

augmented with additional information that can help the reader understand the details of the study, and be properly edited for English to make is more readable, then it is very appropriate for HESS and can be a valuable contribution to the seasonal forecasting literature.

AC: For the FS_S4 runs we used 15 ensemble members and for the FS_ESP (and FS_ESPcond) we used 30. For the FS_S4 we are limited by the number of ensemble members of the ECMWF_S4 hindcast. We made some trials to use randomly 15 ensemble members for the FS_ESP, in order to have the same number of ensemble members than for the FS_S4. The reduction in skill was very small. Other studies also show that the increase of the skill is not significative after 15 ensemble members (e.g (Weigel et al., 2007)). Moreover, as our purpose is not to compare in general the skill of the FS_S4 to that of the FS_ESP, but to assess the skill of these two seasonal forecasts in the Limpopo with the information available, we understood that there was no reason to reduce the number of ensemble members of the FS_ESP. However, we understand the limitations and recognize that the smaller ensemble size can be detrimental for FS_S4, when compared with FS_ESP.

The manuscript was edited for English.

Specific Comments

RC: P9963: There are statements in the introduction that require (or should be supported) by more references. As an example: “Climate change studies show evidence of an intensification of the global water cycle (Huntington, 2006), where extreme events including floods and droughts are expected to become more frequent.” Given the recent focus on the issue of historical and projected changes in floods and droughts, only citing the paper of Huntington seems a little brief. Also the statement (P9963/L 8-10) that seasonal forecasting hasn’t been applied widely in drought predictions fails to recognize that NCEP has a drought monitoring (multi-model) system as well as a seasonal hydrological forecasting system running at EMC, and the NMME is widely focused on seasonal prediction of meteorological drought (as a programmatic focus of the NMME experiment). So the introduction needs to more reflective of what’s really happening, and to cite the paper that are reporting this work.

AC: We agree with the reviewer. The statement of intensification of the water cycles was supported with further literature (e.g. (IPCC, 2007; Hansen et al., 2012; Trenberth, 2012; Coumou and Rahmstorf, 2012). Regarding the sentence in P9963/L8-10 we added the word "as" before widely and we restructured the sentence to emphasize that operational predictions of streamflow are more commonly used for floods than for droughts. We have also added a description of available drought monitoring and forecasting systems in the introduction following the sentence in P9963/L8-10:

There are several Drought Early Warning Systems (DEWS) currently in existence in the world, though due to the complexity of drought these are arguably less developed than many flood early warning systems. Grasso (2009) reports that only three institutions provide information on the occurrence of major droughts at the global scale; FAO’s Global Information and Early Warning System on Food and Agriculture (GIEWS), the Humanitarian Early Warning Service (HEWS) operated by the World Food Programme (WFP), and the Benfield Hazard Research Centre at University College London.

In the United States the U.S. Drought Monitor (http://droughtmonitor.unl.edu/) was set up in collaboration between the US Department of Agriculture (USDA), NOAA, the Climate Prediction Centre, and the University of Nebraska. It provides insight to current drought conditions and impacts at the national and state level through an interactive map, presenting multiple drought indicators combined with field information and expert input. It also includes 6- to 10 day outlooks and monthly and seasonal forecasts of precipitation, temperature, soil moisture and streamflow. The National Weather Service's National Center for Environmental Prediction's (NCEP) also has a (multi-model) drought monitoring system, as well as a seasonal hydrological forecasting system running at the Environmental Modeling Center (Ek et al., 2010). Additionally, the North American Multi-Model Ensemble (NMME), which became an experimental real-time system in August 2011, is mainly focused on seasonal prediction of meteorological drought (Kirtman et al., 2013).

In Europe the European Commission Joint Research Centre (JRC) has established the European Drought Observatory (EDO, http://edo.jrc.ec.europa.eu/), which includes an interactive map viewer with drought-relevant information. It includes real-time maps of different drought indicators, including the Standardized Precipitation Index (SPI), snow and soil moisture anomaly, and vegetation productivity anomaly. These indicators are combined in an overall indicator that is used to provide warnings and alerts. A one week forecast of the expected soil moisture anomaly is also provided. The Beijing Climate Center (BCC) of the China Meteorological Administration (CMA) similarly monitors the development of drought across China, with maps on current drought conditions being updated daily on their website.

The FEWS Net for Eastern Africa, Afghanistan, and Central America reports on current famine conditions, including droughts, by providing monthly bulletins that are accessible on the FEWS Net webpage. However, a drought forecast is not provided. Other drought warning systems over Africa include the Botswana national early warning system (EWS) for drought (Morgan, 1985) and the Regional Integrated Multi-Hazard Early Warning System for Africa and Asia (RIMES). In the latter a drought early warning system is being adapted to identify climate and water supply trends in order to detect the probability and potential severity of drought (RIMES, 2014).

Advances regarding drought early warning systems in Africa in the last few years are remarkable. There is an increasing availability of drought monitoring and forecasting tools for decision making that can provide real time monitoring and forecasting of drought across the continent. The Land Surface Hydrology Group at Princeton University, USA, has recently established an African Flood and Drought Monitor (http://stream.princeton.edu/) with support from the International Hydrology Program of UNESCO. The system provides near real time monitoring of land surface hydrological conditions based on the Variable Infiltration Capacity (VIC) model. The monitor is updated every day at 2 days behind real time. The database provides the daily conditions of precipitation, temperature, wind speed, soil moisture, evaporation, radiation, and different components of runoff in the continent, as well as historic hydrological records in Eastern, Southern and Western sub-regions up to 10 antecedent years, and derived products such as current drought conditions. They also provide precipitation, temperature and SPI forecast (Sheffield et al., 2014). Recently Barbosa et al. (2013) developed a Pan-African map viewer for drought within the framework of the DEWFORA project, following the main features of the earlier developed EDO. The African Drought Observatory (ADO) is a web application hosted by JRC (http://edo.jrc.ec.europa.eu/ado/ado.html) that provides historical and near-real time monitoring information, as well as seasonal forecasts describing meteorological, agricultural and hydrological droughts (Barbosa et al., 2013).

RC: P9968/L18. Are the dates correct? September 2009 to December 2010? Should this be 1979-2010?

AC: Yes, the dates are correct. This sentence explains how we corrected the precipitation during the period that is not covered by the GPCP v2.1, given that GPCP v2.1 stops in September 2009.

RC: P9966/L10 “It is, however, unreliable, causing frequent droughts and floods also commonly occur in the rainy season.” Awkward sentence. Restructure the English

AC: This sentence was restructured to: "Moreover, rainfall is highly variable causing frequent droughts, though floods can also occur during the rainy season."

RC: P9970/L1 Additional information of when/how the ECMWF ensemble forecasts are generated, would help the readers. Are they all generated at the beginning of the month, or are they distributed throughout the month? As used, are they monthly averaged Tmax, Tmin and total precipitation, or daily forecasts out to 6 months? If hindcasts that were used are monthly mean forecasts, how do you downscale to a daily time step for the hydrological modeling? Or spatially downscale to the fine resolution of the hydrological model? If the ESP (ESP_cond) forecasts use 30 ensemble members, how many does the S4 procedure use given that the ECMWF hindcasts have only 15 members?

AC: The S4 ensemble seasonal forecasts are initialised on the 1st of each month and the ensemble is generated by perturbations in the initial conditions and by the use of stochastic physics in the atmosphere during the model integration (out to 6 months lead time) (Molteni et al., 2011). Precipitation inputs to the hydrological model are accumulated from the 6-hourly S4 model values, while evaporation was calculated using the daily maximum and minimum temperatures directly archived by the meteorological model.This was added in the text to clarify. They are used with a daily time step to force the hydrological model to forecast daily hydrological variables for the following six months. The daily time series outputs are then aggregated to monthly time series. As explained before, the meteorological forcings were not downscaled between the 0.5˚ and 0.05˚ grid cells, but assumed to be constant over the 0.5˚ grid cell. The FS_S4 system uses 15 ensemble members as it is limited by the ECMWF hindcast (see answer above).

RC: P9971/L18-19: By "multi-annual mean of precipitation" I assume you are referring to the precipitation climatology in both the base data set and hindcast data set. Using "climatological" distribution is more standard terminology.

AC: "multi-annual mean" was changed for "climatological long term mean" as suggested.

RC: P9971/L25-26. Are the problems with the ECMWF hindcasts/forecasts related to "other problems of the forecasts such as inter-annual variability, ensemble spread or daily variability" discussed? They need to be.

AC: We mentioned these problems as generic problems of any coupled atmosphere-ocean model (true for forecast ranges from medium-to-seasonal and also in climate projections). This is mentioned to highlight that we only performed a very simple bias correction that corrected the climatological long term mean, and does not address any further problems of the atmospheric forecasts. A detailed discussion of this would require the use of in-situ observations of precipitation and we think that such work is out of the scope of the current study.

We changed the sentence: "It does not address other problems of the forecasts such as inter-annual variability, ensemble spread or daily variability" to:

"It does not address other problems of the forecasts, common to all coupled atmosphere-ocean models, such as inter-annual variability, ensemble spread or daily variability"

RC: P9973. Section 2.2.3 is not well written. The procedures for determining the weights must be clarified. For example, the equations for the parameters ï ¸An and ˇ α need to be provided. On what basis were their values determined? Given the values, what are the weight assigned to the sorted years adjacent to the current year's ONI? Basically as written, readers will find it difficult to fully understand what was done in the study. Also, given a record of 30 years to generate a 30 member ensemble (by sampling with replacement) on average how many duplicate years occurred on average?

AC: This section was extended. Equations for the parameters are now provided. In the text, we already explain that "For the FS_ESPcond we chose to keep the parameters constant (λ = 2 and α =1) given that the optimal selection of parameters would vary for each sub-basin. Performing an in-depth selection of parameters for each sub basin is out of the scope of this study. Here we use λ = 2 and α =1, meaning that all ensemble members have a non-zero probability of being included in the ensemble, with that probability based on the distance between the ENSO indexes and the distance sensitive weighting parameter (linear for λ = 2)".

The number of duplicate years that occur on average on the ensemble depends on the probability assigned to each ensemble member based on the ONI indexes, which will vary for each forecast start date.

RC: Section 2.3.3. The uncertainty in the skill scores are computed by a bootstrap method where the ensembles are resampled (with replacement). This implies that you'd never get an ensemble member different than one already forecast (or in the case of ESP) previously observed meteorology. What is the impact of this on the uncertainty estimates? For ESP this could be assessed by using a longer data set than 30 years but only generating a 30 member ensemble.

AC: The reviewer is right. A limitation of this bootstrap procedure is that statistics computed from discrete bootstrap samples may differ from the ones based on continuous data, and this might lead to overestimation of the confidence. Although there are some limitations, this method is widely used in the literature (e.g. (Dutra et al., 2014; Friederichs and Thorarinsdottir, 2012; Wilks, 2011)) to estimate confidence intervals as it does not require a distributional assumption. For ESP, we do not have a longer data set than 30 years as we used the ERA-Interim in the period 1979-2010 as observed meteorology. An explanation of these limitations was added in the manuscript.

RC: P9976/L20. The "root stress" isn't defined, so readers won't know how this drought metric has been computed.

AC: The "root stress" (RS) is an indicator of the available (or the lack of) soil moisture in the root zone, which can be calculated for each grid cell. The RS varies from 0 to 1, where 0 indicates that the soil water availability in the root zone is at field capacity and 1 indicates that the soil water availability in the root zone is zero and the plant is under maximum water stress. This explanation was added in the manuscript.

RC: P9978/L1-2. Please explain more clearly the difference between "The mean runoff season and high runoff season".

AC: Instead of "mean", it should have said "main". This was corrected. By "main runoff season" we meant the season from December to May, which is the season with highest runoff. The "high runoff season" is the four months period, January to April, where the runoff is the highest during the year. These periods are presented in the new Figure 4 (see answer to reviewer #2).

P9978/L15+ I'm somewhat confused by how the forecast system is being evaluated. Earlier (pg 9977) there is a discussion that the reservoir level provides a decision metric as to curtailing irrigation so that would be the target for the ROCS and BS forecast metrics

AC: The forecast system is evaluated for SRI. However, as we also want to analyse the ability of the system to forecast distributed variables (RS for agricultural droughts) and water levels in the reservoirs (for irrigation curtailments), we also evaluated the skill of the forecast system in predicting these variables. We added an explanation on the manuscript in a new section 2.3.3 to clarify.

(P9977/L17-18). On P9978/L15 it appears that river flows (SRI) and not reservoir levels will be the variable to be used in assessing the forecast system.

AC: Please see reply to previous comment.

RC: P9980/L19. It appears (from the wording) that the water level analysis assumes a virtual reservoir in each grid, which makes no sense. But figures 8 and 9 seem to indicate there are a number of reservoirs across the basin (the circles). If these are reservoirs, then it would be useful to indicate them on figure 2, where only Tzaneen dam is shown. The description of the reservoir forecast analysis needs to be clarified and written more clearly. The general discussion of reservoirs and some reference to the Tzaneen dam makes the section confusing and rather weak.

AC: These are actual reservoirs. As suggested, their location was added in Fig. 2. Moreover, we clarified the description of the reservoir forecast analysis and removed specific references to the Tzaneen dam.

RC: P9981/L27, P9982/L1-2 Figure 10 should also show the “verification” data (i.e. the simulated reference discharge that is your surrogate for the observed streamflow) that actually occurred during the forecast period.)

AC: Yes, actually the "verification" data is shown in the graph as "observed". This was explained and modified to clarify.

RC: P9982/L12. I think the weak response described here should be attributed to atmospheric noise” and not to “different climate forcings”.

AC: This point was not clear. There is not a weak response here, as there was actually no drought condition in 1997/98 and the S4 system correctly forecasted this. This sentence aims to explain that S4 could capture other climate predictors besides the SST. If our forecast was based on SST alone, 1997/98 would have been forecasted as a severe drought. This point was clarified in the manuscript.

RC: P9984/L26-28. Sentence as written is rather awkward. Perhaps “Maps of spatially distributed ROCS and BSS show, throughout the basin, skill higher than climatology (ROCS> 0.5, BSS> 0) for the FS_S4 forecast that agricultural droughts and water levels will to be lower than the 50 and 37.5 percentiles.”

AC: This sentence was modified to clarify: "The FS_S4 was also evaluated regarding its ability to predict agricultural droughts and curtailments in irrigation (water levels lower than the 50th and 37.5th percentiles). Maps of spatially distributed ROCS and BSS show that the skill of the FS_S4 forecast in predicting these conditions is higher than climatology (ROCS > 0.5, BSS > 0) throughout the basin."

RC: P9985/L11 Sentence “It is recommended that as a next step the forecast skill of the FS_S4 and FS_ESPcond is assessed in an actual forecasting mode for the following summer season.” Poor English; wrong verb tense. Try “As a next step, it is recommended that the forecast skill of the FS_S4 and FS_ESPcond be assessed in an actual forecasting mode for the following summer season.”

AC: This was modified as suggested.

RC: P9985/L14. What kind of data assimilation. What would be the source of the information?

AC: Examples could be: water levels in reservoirs and steams from data collected in the basin by water managers. This was added in the text and the idea was expanded: "As a next step, it is recommended that the forecast skill of the FS_S4 and FS_ESPcond be assessed in an actual forecasting mode for a following summer season. The seasonal meteorological forecast from S4 can be obtained in real-time for research purposes. To test a pre-operational system, the forecasting system ought to be statistically post-processed in order to remove biases in streamflow predictions. Moreover, the initial conditions for the forecasts could be better estimated through data assimilation of water levels in reservoirs and steams. This data could be obtained from the water managers of the basin."

RC: P9984/L13, L15. On line 13 it is suggested that S4 can be obtained in real-time, yet in line 15 it is stated that a limitation to S4 is its access. So is the earlier statement speculation? If so these sentences need to be edited for clarity for what is and is not available.

AC: We agree that this is confusing. ECMWF S4 is not publically available. It is produced operationally in real-time but only available to the ECMWF member-states and other institutions (public/private) with formal agreements. For research purposes, like this study, the re-forecasts may be acquired on request. When the purpose is to build an operational system, then the organization responsible for providing the operational drought forecasts should discuss with ECMWF member-states the terms and conditions for the access and use of the forecasts. This is one of the reasons why the simple FS_ESPcond could sound attractive, as there is no restriction in terms of data availability. This was clarified in the manuscript.

References added:

Barbosa, P., Naumann, G., Valentini, L., Vogt, J., Dutra, E., Magni, D., and De Jager, A.: A Pan-African map viewer for drought monitoring and forecasting, 14th Waternet Symposium, Dar es Salaam, Tanzania, 30 October to 1 November 2013,

Coumou, D., and Rahmstorf, S.: A decade of weather extremes, Nature Clim. Change, 2, 491-496, doi: 10.1038/nclimate1452, 2012.

Dutra, E., Pozzi, W., Wetterhall, F., Di Giuseppe, F., Magnusson, L., Naumann, G., Barbosa, P., Vogt, J., and Pappenberger, F.: Global meteorological drought - Part 2: Seasonal forecasts, Hydrol. Earth Syst. Sci. , 18, 2669-2678, doi:10.5194/hess-18-2669-2014, 2014.

Ek, M., Xia, Y., and the NLDAS team: NCEP/EMC NLDAS Support for Drought Monitoring and Seasonal Prediction, US National Oceanic and Atmospheric Administration, Climate Test Bed Joint Seminar Series, NASA, Goddard Visitor Center, Greenbelt, Maryland, http://www.nws.noaa.gov/ost/climate/STIP/FY10CTBSeminars/mek_041410.pdf, access 01/12/2014, 2010.

Friederichs, P., and Thorarinsdottir, T. L.: Forecast verification for extreme value distributions with an application to probabilistic peak wind prediction, Environmetrics, 23, 579-594, 2012.

Grasso, V. F.: Early Warning Systems: State-of-Art Analysis and Future Directions - Draft Report, United Nations Environment Programme (UNEP), 2009.

Hansen, J., Sato, M., and Ruedy, R.: Perception of climate change, Proceedings of the National Academy of Sciences, 109, E2415-E2423, doi:10.1073/pnas.1205276109, 2012.

IPCC: Summary for Policymakers. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Kirtman, B. P., Min, D., Infanti, J. M., Kinter, J. L., Paolino, D. A., Zhang, Q., van den Dool, H., Saha, S., Mendez, M. P., Becker, E., Peng, P., Tripp, P., Huang, J., DeWitt, D. G., Tippett, M. K., Barnston, A. G., Li, S., Rosati, A., Schubert, S. D., Rienecker, M., Suarez, M., Li, Z. E., Marshak, J., Lim, Y.-K., Tribbia, J., Pegion, K., Merryfield, W. J., Denis, B., and Wood, E. F.: The North American Multimodel Ensemble: Phase-1 Seasonal-to-Interannual Prediction; Phase-2 toward Developing Intraseasonal Prediction, Bulletin of the American Meteorological Society, 95, 585-601, doi: 10.1175/bams-d-12-00050.1, 2013.

Molteni, F., Stockdale, T., Balmaseda, M. A., BALSAMO, G., Buizza, R., Ferranti, L., Magnunson, L., Mogensen, K., Palmer, T., and Vitart, F.: The new ECMWF seasonal forecast system (System 4), ECMWF Tech. Memo., 656, 49 pp., 2011.

Morgan, R.: The development and applications of a Drought Early Warning System in Botswana, Disasters, 9, 44-50, 10.1111/j.1467-7717.1985.tb00909.x, 1985.

RIMES, Regional Integrated Multi-Hazard Early Warning System for Africa and Asia, Water Related Hazard - Drought: http://www.rimes.int/wrh/drought, access: 11 December 2014, 2014.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., and Ogallo, L.: A drought monitoring and forecasting system for sub-Sahara African water resources and food security, Bulletin of the American Meteorological Society, 95, 861-882, doi: 10.1175/bams-d-12-00124.1, 2014.

Trenberth, K.: Framing the way to relate climate extremes to climate change, Climatic Change, 115, 283-290, doi: 10.1007/s10584-012-0441-5, 2012.

Weigel, A. P., Liniger, M. A., and Appenzeller, C.: The Discrete Brier and Ranked Probability Skill Scores, Monthly Weather Review, 135, 118-124, doi: 10.1175/mwr3280.1, 2007.

Wilks, D. S.: Statistical methods in the atmospheric sciences - 3rd edition, International Geophysics series, vol 100, 676 pp., 2011.

Interactive comment on "Hydrological drought forecasting and skill assessment for the Limpopo river basin, Southern Africa" by P. Trambauer et al.

Anonymous Referee #2

Received and published: 21 November 2014

Reaction to the interactive comment by Anonymous Referee #2

We would like to thank this referee for his/her interesting comments and suggestions that contributed to improve our paper and to clarify specific points. Hereby we present the authors reply (AC) to the referee's comments (RC).

RC: The paper presents hydrological seasonal prediction experiment using three different driving meteorological inputs (dynamic ECMWF seasonal forecast, ESP and ENSO conditioned ESP) for Limpopo River basin. The study in general follows correctly a common methodology applied for such studies, however does not provide detail information about some important steps of the whole process (see below). The text should be more inclusive to provide reader with all important information on methodology without simply referring to other existing studies; e.g. the results of NS criterion of hydrological model calibration should be stated (authors only refer to another study of the same team - P9978 "In these stations the performance of the hydrological model is found to be satisfactory based on evaluation measures and ranges proposed by Moriasi et al. (2007). These results are presented by Trambauer et al. (2014)."). In general, a text is understandable but some sentences are difficult to read and need overall grammar revision ("It is, however, unreliable, causing frequent droughts and floods also commonly occur in the rainy season."). In a whole, if revised for English and completed by missing detailed information I consider this study a valuable contribution to extended hydrological prediction system literature.

AC: As suggested, we included the results of the model performance. We added a Table (see below) with the performance measure for the selected basins, basin area, mean annual observed runoff, and observed runoff coefficient (RCobs).

Table 1 Model evaluation measures for runoff for selected stations, ordered by basin size

Station number	Sub-basin area (km^2)	Mean annual observed runoff (m^3/s)	RCobs (%)	R^2	NSE	RSR
24	342,000	96.9	1.7	0.92	0.90	0.32
1	201,001	39.5	1.2	0.69	0.57	0.65
18	98,240	12.2	0.7	0.68	0.62	0.62
20	12,286	14.8	5.3	0.70	0.65	0.59

The mentioned sentence ("It is, however, unreliable...") was modified to clarify to: "Moreover, rainfall is highly variable causing frequent droughts, though floods can also occur during the rainy season"

The manuscript was edited for English

Specific comments:

RC: P 9963-9965 Introduction does not provide literature review of existing studies (or operational implementations) on seasonal hydrological prediction systems.

AC: We have added a review of existing drought early warning systems in the revised manuscript, which reads:

"There are several Drought Early Warning Systems (DEWS) currently in existence in the world, though due to the complexity of drought these are arguably less developed than many flood early warning systems. Grasso (2009) reports that only three institutions provide information on the occurrence of major droughts at the global scale; FAO's Global Information and Early Warning System on Food and Agriculture (GIEWS), the Humanitarian Early Warning Service (HEWS) operated by the World Food Programme (WFP), and the Benfield Hazard Research Centre at University College London.

In the United States the U.S. Drought Monitor (http://droughtmonitor.unl.edu/) was set up in collaboration between the US Department of Agriculture (USDA), NOAA, the Climate Prediction Centre, and the University of Nebraska. It provides insight to current drought conditions and impacts at the national and state level through an interactive map, presenting multiple drought indicators combined with field information and expert input. It also includes 6-to 10 day outlooks and monthly and seasonal forecasts of precipitation, temperature, soil moisture and streamflow. The National Weather Service's National Center for Environmental Prediction's (NCEP) also has a (multi-model) drought monitoring system, as well as a seasonal hydrological forecasting system running at the Environmental Modeling Center (Ek et al., 2010). Additionally, the North American Multi-Model Ensemble (NMME), which became an experimental real-time system in August 2011, is mainly focused on seasonal prediction of meteorological drought (Kirtman et al., 2013)

In Europe the European Commission Joint Research Centre (JRC) has established the European Drought Observatory (EDO, http://edo.jrc.ec.europa.eu/), which includes an interactive map viewer with drought-relevant information. It includes real-time maps of different drought indicators, including the Standardized Precipitation Index (SPI), snow and soil moisture anomaly, and vegetation productivity anomaly. These indicators are combined in an overall indicator that is used to provide warnings and alerts. A one week forecast of the expected soil moisture anomaly is also provided. The Beijing Climate Center (BCC) of the China Meteorological Administration (CMA) similarly monitors the development of drought across China, with maps on current drought conditions being updated daily on their website.

The FEWS Net for Eastern Africa, Afghanistan, and Central America reports on current famine conditions, including droughts, by providing monthly bulletins that are accessible on the FEWS Net webpage. However, a drought forecast is not provided. Other drought warning systems over Africa include the Botswana national early warning system (EWS) for drought (Morgan, 1985) and the Regional Integrated Multi-Hazard Early Warning System for Africa and Asia (RIMES). In the latter a drought early warning system is being adapted to identify climate and water supply trends in order to detect the probability and potential severity of drought (RIMES, 2014).

Advances regarding drought early warning systems in Africa in the last few years are remarkable. There is an increasing availability of drought monitoring and forecasting tools for decision making that can provide real time monitoring and forecasting of drought across the continent. The Land Surface Hydrology Group at Princeton University, USA, has recently established an African Flood and Drought Monitor (http://stream.princeton.edu/) with support from the International Hydrology Program of UNESCO. The system provides near real time monitoring of land surface hydrological conditions based on the Variable Infiltration Capacity (VIC) model. The monitor is updated every day at 2 days behind real time. The database provides the daily conditions of precipitation, temperature, wind speed, soil moisture, evaporation, radiation, and different components of runoff in the continent, as well as historic hydrological records in Eastern, Southern and Western sub-regions up to 10 antecedent years, and derived products such as current drought conditions. They also provide precipitation, temperature and SPI forecast (Sheffield et al., 2014). Recently Barbosa et al. (2013) developed a Pan-African map viewer for drought within the framework of the DEWFORA project, following the main features of the earlier developed EDO. The African Drought Observatory (ADO) is a web application hosted by JRC (http://edo.jrc.ec.europa.eu/ado/ado.html) that provides historical and near-real time monitoring information, as well as seasonal forecasts describing meteorological, agricultural and hydrological droughts (Barbosa et al., 2013)."

RC: P 9966 There were four stations evaluated in the study. Two of them representing smaller area have provided generally less satisfactory results. A bit surprisingly the best result has not been gained for the closing profile of a study basin (P 9979/5-10). This is not discussed and authors do not attempt to explain it. With respect to this I miss more detailed information about sub-basins (area, general geographical conditions etc.). Such information might be interesting (and supporting) for interpretation and discussion of results.

AC: A table with extra information on the basins considered (area, mean annual runoff, and observed runoff coefficient) was added, see Table 1 above. Regarding the lower skill of the largest basin, we included a possible explanation in P 9979/8:

"The lower skills for the station with largest contributing area for FS_S4 may be due to the shift from an arid to a more tropical climate in the downstream part of the basin, which means that the persistence of initial conditions would be expected to be lower."

RC: P 9968 A method of deriving of precipitation data is not sufficiently described and validated (a critical impact of meteorological input is obvious from gridded pattern of fig. 7). A way how monthly precipitation data are converted to daily time series for model simulations remain unexplained.

AC: The precipitation data (both reanalysis and re-forecasts) are obtained at a daily time step. No conversion from monthly to daily time series was needed. In P 9968, L7 we added "at a daily time step" after "with the ERA-Interim forcing meteorological data". The monthly time-scale is used only for the correction of the monthly means of ERA-Interim and for the climatological mean correction of S4 forecasts. In this study we did not focus on the detailed evaluations of this data, or further refinement, as our intention was to use available and documented datasets to force the hydrological model. The point that the reviewer raises is indeed very important, but such evaluation would take the paper to another focus on the evaluations of precipitation forecasts on the region.

RC: P 9969 Information about initializations dates and lengths of simulations is quite confused in this section and in results description. I would propose to include a figure with overview of forecast periods during the year and its relation to Limpopo river flow regime.

AC: A figure was added as suggested. The text was clarified and now reads: "In the hindcast, the first forecast of each season is issued in August and includes the seasonal (6-months) forecast from August to January. The forecast is updated at the beginning of each month from September to February. The last forecast of the season is issued in February, covering the period from February to July (see Fig. 4). All simulations are done at a daily time step."

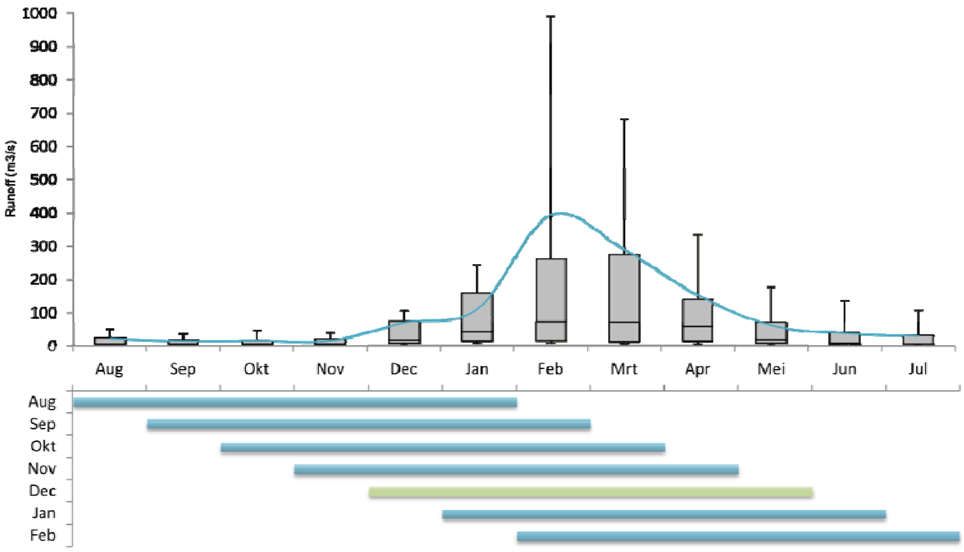

Fig. 4 Upper plot: Limpopo river flow regime for Station 24 at Chokwe. The blue line represents the average observed runoff, and the whiskers of the boxplots represent the 10th percentile and the 90th percentile. The lighter and darker shaded areas represent the main runoff period and high runoff period, respectively. Lower plot: Initialization dates and length of forecasts during the year. The forecast issued in December is highlighted as the one that captures the main runoff season.

RC: P 9971/18 Does "multi-annual mean" mean the same as "mean" long term climatology?

AC: Yes, "multi-annual mean" was changed for "climatological long term mean" to clarify.

RC: P 9971 According to a described precipitation bias correction the monthly mean correction factor is “linearly interpolated from monthly values to daily assuming it corresponds to day 15 of the particular month”. This might suggest that interpolation has been done the way illustrated in fig. 1. Daily time series (a) is bias corrected on monthly basis (b). Let’s suppose the correction factor (alpha) is 2.39 for a given month and 0,6 for a preceding and following months. If this number is applied to correct daily rainfall uniformly during the given month the corrected monthly precipitation total would equal to 255 mm (8,5 mm per day) while if interpolated according to description of authors the monthly corrected total would be 214 mm (7,1 mm per day) only (c).I believe that was not the case, only a method description should be more precise.

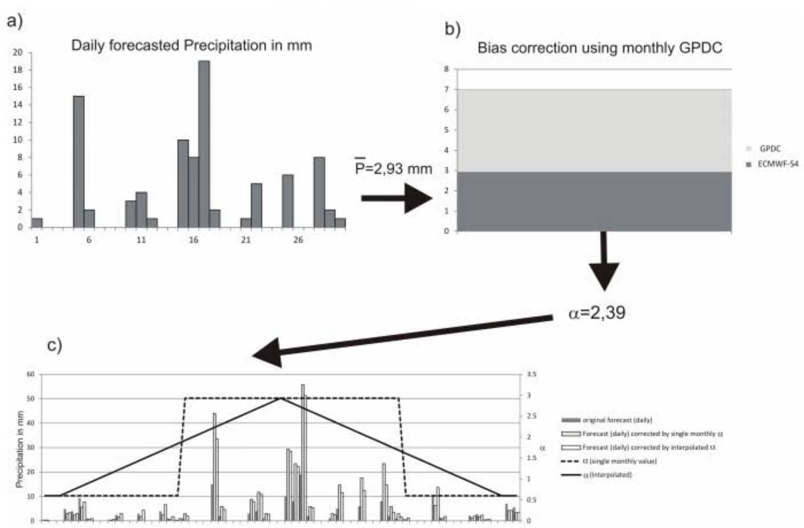

Fig. 1. Fig. 1 bias correction process (see text for explanation)

AC: The method described by the reviewer as “interpolated” was the one we used. The reviewer is right that it will not match the mean. However, the correction factor is only a long-term climatological value for each calendar month and forecast lead time, i.e does not change from year to year. The decision of using a linear interpolation instead of a step function, as suggested by the reviewer, was to avoid “jumps” in the precipitation amounts, as we can see in the Fig.1. We did not evaluate the impact of using a different method for the bias correction, and we only applied a very simple correction. In our opinion, a more sophisticated bias correction would only be justified if we had good quality in-situ observations of precipitation, and this was not the case for the region.

RC: P 9974 and 9976 Resampling procedures have to be described in more detail.

AC: The resampling procedures were described in more detail. The explanations in P9974 and 9976 were expanded, respectively:

- "For each forecast start date, we construct an ensemble meteorological forecast of 30 members to be consistent with FS_ESP. The selection of the members is based on a resampling with replacement procedure given the probability assigned to each member. From the 30 possible ensemble members to be included, those with an ONI index closer to that of the forecast year, have a higher probability of being included in the ensemble. This means that some ensemble members are included more than once, and some are not included at all. "
- "The uncertainty of the ROCS is estimated by applying a bootstrap resampling with replacement procedure.
  For the FS_S4 and FS_ESP forecasts, we randomly replace (allowing repetition) the original forecast and verification pair to produce a new sample of the same size as our original sample. We then calculate the ROCS from the new sample. We repeat this procedure to create 1000 new samples from which we generate an empirical distribution of the ROCS. The 90% confidence interval is estimated from the 5th and 95th percentiles of this empirical distribution."

RC: P 9976/20 Root stress indicator is used, however it is not defined. Is the Root stress the same as a modeled soil water deficit?

AC: The "root stress" (RS) is an indicator of the available (or the lack of) soil moisture in the root zone, which can be calculated for each grid cell. The RS varies from 0 to 1, where 0 indicates that the soil water availability in the root zone is at field capacity and 1 indicates that the soil water availability in the root zone is zero and the plant is under maximum water stress. This explanation was added in the manuscript.

RC: P 9978/1 The "mean runoff season" and "high runoff season" need to be defined.

AC: Instead of "mean", it should have said "main". By "main runoff season" we meant the season from December to May, which is the season with highest runoff. The "high runoff season" is the four months period, January to April, where the runoff is the highest during the year. These periods are presented in the new Figure 4.

RC: P 9984 Authors conclude that initial hydrological conditions (IHC) contributes to predictability up to 2 to 4 months but do not discuss these findings with Shukla et al. (2013) who have found shorter impacts of IHC.

AC: This statement is here clarified. We did not find that the initial conditions dominate the predictability up to 2 or 4 months, but that they contribute to the predictability. The higher skill of FS_S4 over that of FS_ESP during the wet season for every lead time suggests that the Meteorological forecast (MF) might dominate the hydrological drought predictability for every lead time, as reported by Shukla et al. (2013). However, we cannot make the same kind of conclusions as Shukla et al. (2013) as we did not apply the reverse ESP procedure, which derives its skill solely from the perfect forecast. In our case FS_S4 derive its skill both from the IHC and meteorological forecast (MF). Our statement refers to some contribution of the IHC to the skill, given that when we examine the results of FS_ESP (which derives its skill only from the knowledge of IHC) we see some skill only up to 2, and in some cases 4 months. After that there is no skill at all from the IHC.

In the manuscript, P9984/17-25, it already states: "The higher skill of the FS_S4 and FS_ESPcond compared to that of the FS_ESP for every lead time is in line with the study of Shukla et al. (2013) who show that for the region of the Limpopo river basin the meteorological forecast dominates the hydrological predictability for the wet season for almost every lead time considered. Only for the 1-month lead time forecasts issued in October they found a higher influence of the hydrological initial conditions to some extent. Moreover, Yossef et al. (2013) indicate that for semi-arid regions the initial conditions do not contribute much to the skill given the high sensitivity of the runoff coefficient to rainfall variability."

RC: P 9982-9985 Station 24 (closing gauge of the basin) is in general predicted with less skill than upstream station 1 but physical explanation is not discussed.

AC: It is true that the skill is in general better for station 24 than for station 1. The lower skills for the station with largest contributing area for FS_S4 may be due to the shift from an arid to a more tropical climate in the downstream part of the basin, which means that the persistence of initial conditions would be expected to be lower. Also, given that this is mostly the case for the FS_S4 and not so much for the FS_ESP and FS_ESPcond, we can speculate the ECMWF S4 seasonal forecast might have a better skill for the northern (more arid) part of the basin (area corresponding to sub-basin draining to station1), than for the southern part of the basin. This is also reflected in the spatial analysis of skill presented in Figure 9.