Dear Editor,

thank you for your comments. Please find in attached the point by point answer to reviewers and the revised version of the paper. New sections are indicated in blue. Note that section 2 has been strongly modified and equations have been added.

The justification of the use SPI-1 is now discussed in the introduction.

The scores based on the drought conditions are added in section 3.6. Nevertheless 80 % of the duration of SPI-1 < -1 is 1 month, so the differences are not significant (see text for more details).

Finally, the comparison between numerical and statistical forecasts is a very important perspective of this work. Based on a complete, multi-model and multi-score analysis, this paper provides the first assessment of the predictability of droughts over the entire European continent.

Unfortunately it is not possible to directly compare this work with previous studies that used different periods, or analyse specific areas. Because this kind of comparison would require a long and detailed analysis, we consider this point as a logical follow-up of the research presented in this paper. The advantages of the use of statistical forecast, included Weather Types, will be analyzed and presented in the futur.

Please find in attached the point by point answer to reviewers and the revised version of the paper.

Reviewer #1

General comments This paper discusses the forecast of the ECMWF seasonal (SEAS) and their monthly ensemble forecasts to forecast drought which is measured by the standardized precipitation index (SPI1). The paper is long on statistics. There is not a lot of physical explanation. The paper is well written and should be published. Specific comments:

Reply -> We thank the reviewer for the encouraging and insightful comments, which helped us to make the paper more straightforward.

1. Drought usually indicates persistent lack of precipitation. In general, it means negative SPI (SPI< -1 as the indicator used by authors) for three months or longer. Therefore, the 3-months or 6- month SPIs are used to indicate drought instead of SPI1. Is there any reason that you use SPI1?

Reply -> This comment is a very important point. The use of the SPI-1 is motivated by several reasons:

a- The skill score of precipitation forecasts decreases drastically during the first month. So, the benefit to use a lead time of two months or more is not obvious (Dutra et al. 2013).

b- While in this study we intend to test the reliability of the forecast, an improvement could be achieved with the combination of different information types: monitoring by satellite or in-situ measurements that give an accurate characterization of ongoing drought conditions (e.g. during the last two months), combined with the forecasted SPI-1 that provides the best estimate of near future conditions. However, this will still not allow looking more than one month ahead and would bias the testing of the forecast skill, which was the intention of this paper.

c- The seasonal model (SEAS) is here compared to ENS, the up-to-date version of the ensemble system. Currently, the ENS, however, provides a forecast only up to 32 days once a week. So it is technically impossible to

compare these models for SPI-3 or SPI-6.

d- a one month forecast with a good reliability is considered to be a very valuable product for decision makers as it provides information on the probability of occurrence of a dry spell (in case of ongoing normal conditions) and of the probable persistence or end of a drought (in case of an ongoing precipitation deficit).

We have modified the introduction to include the above comments.

Also we consider it important to provide to stakeholders a trend of precipitation forecast one month in advance. As we have indicated in the introduction, we do not intend to detect the entire period of a drought but the objective is to assess the most robust product for drought forecasting. To help the stakeholders to make a decision, this work is an additional product to the drought monitoring and it will forecast a precipitation deficit that will occur in the next month over a region.

2. If you use SPI3 or SPI6, do your conclusions change?

Reply -> As we have not used any precipitation forecasted for more than 1 month using ENS, it is too speculative to answer as we explicitly intended to focus on SPI-1 (see text above). Nevertheless, based on the decrease of the skill scores with longer lead times, we expect lower skill scores for these lead times.

3. You use SPI1 so how well is your system to predict drought onset? (the first time in a time series that SPI1 is below -1 )

Reply -> In this study we have tested two different thresholds of drought detection: SPI-1 lower than -1 and lower than -1.5 (the time series was too short to analyze the case of extreme droughts with SPI lower than -2). So, the first month of forecast with SPI-1 lower -1 or -1.5 is defined as the beginning of a drought or a dry spell (defined as a short rainfall deficit).

4. Do forecasts have higher skill after a drought onset?

Reply -> This comment is very interesting. We have modified the paper to include a discussion on this:

'The importance of the drought duration has also been tested. The scores were calculated independently for drought onset (first SPI lower than thresholds), persistence (consecutive SPI lower than the threshold), or end of the drought (first SPI above the threshold after a drought). First, the duration of a large majority of SPI-1 lower than -1 (more than 80%) is one month (isolated value, dry spell). The scores display a slight increase of the score for persistent droughts (condition unchanged), for the median the POD score increases from 0.33 to 0.36. But the difference is not significant according to the t-test.'

5. It will be nice to show examples that SPI1 has high/low skill.

Reply -> The case studies selected are very different from each other and adding more examples is indeed desirable. However, the paper is already fairly long and we believe that the case studies presented illustrate the main aspects of the discussion.

Reviewer #2

Title: Early warning of drought in Europe using the monthly ensemble system from ECMWF Authors: C. Lavaysse, J. Vogt, and F. Pappenberger

Summary: In this study the authors compare the skill of the European Centre for Medium - range Weather Forecasts' (ECMWF's) extended range forecasts (lead time up to 32 days) and seasonal forecasts (lead time up to 12 months) in forecasting drought at one-month lead-time. The authors use the Standardized Precipitation Index (SPI) to identify drought events and

estimate drought severity. This is a very useful analysis. The methods used in this study are technically sound and appropriate. The conclusions are supported by the results. I would certainly recommend publication of this manuscript however after some minor to moderate changes that I believe can further improve this manuscript.

Reply -> We thank the reviewer for the encouraging and insightful comments, which helped us to make the paper more straightforward.

Please see my comments below.

(1) This manuscript can benefit a lot by a careful copy editing for several typos (mostly grammar related). I think it will improve the readability of the manuscript.

Reply -> We amended the document and included corrections. Please note that the paper has been professionally copy edited and we apologize for any mistakes.

(2) I understand that the focus of this study is the drought forecast at lead-time of 1 month however I have to wonder, for practical purposes, how useful it is to know about drought severity in the next month. What are the stakeholders that can benefit from the forecasts? I can certainly see the benefit of this during an ongoing drought event but how can one use the drought forecast over the next month to make decision on drought onset or drought propagation since typically drought that persists over a long period of time (varying from a few months to years) are the ones that the decision makers would be concerned about. I also understand (and am sympathetic to the fact) that the skill of seasonal forecasts, beyond a lead time of one month, is generally limited which may have influenced the authors decision to focus on one month lead forecasts nonetheless the implications of the choice of lead time do need to be discussed. Please consider doing so.

Reply -> The use of the SPI-1 is motivated by several reasons :

a- The skill score of precipitation forecasts decreases drastically during the first month. So, the benefit to use a lead time of two months or more is not obvious (Dutra et al. 2013).

b- While in this study we intend to test the reliability of the forecast, an improvement could be achieved with the combination of different information types: monitoring by satellite or in-situ measurements that give an accurate characterization of ongoing drought conditions (e.g. during the last two months), combined with the forecasted SPI-1 that provides the best estimate of near future conditions. However, this will still not allow looking more than one month ahead and would bias the testing of the forecast skill, which was the intention of this paper.

c- The seasonal model (SEAS) is here compared to ENS, the up-to-date version of the ensemble system. Currently, the ENS, however, provides a forecast only up to 32 days once a week. So it is technically impossible to compare these models for SPI-3 or SPI-6.

d- a one month forecast with a good reliability is considered to be a very valuable product for decision makers as it provides information on the probability of occurrence of a dry spell (in case of ongoing normal conditions) and of the probable persistence or end of a drought (in case of an ongoing precipitation deficit).

We have modified the introduction to include the above comments.

Also we consider it important to provide to stakeholders a trend of precipitation forecast one month in advance. As we have indicated in the introduction, we do not intend to detect the entire period of a drought but the objective is to assess the most robust product for drought forecasting. To help the stakeholders to make a decision, this work is an additional

product to the drought monitoring and it will forecast a precipitation deficit that will occur in the next month over a region.

(3)The authors use several metric scores for the evaluation of ECMWF's forecasts, which is a real strength of this study however I think those metric scores can be better explained. I would suggest dividing the section 2.4 into subsections for each metric scores and explaining them separately. Please also provide the corresponding equations where applicable.

Reply -> We have improved the explanation of the scores and added references to illustrate that they are part of the standard scientific literature.

Minor comments:

(1)Page 1975 Lines 5-9: In this paragraph the different categories of drought are mentioned. I found the sequence of drought categories a bit odd. In general, meteorological drought is mentioned before agricultural drought followed by hydrological drought. The reason for which of course is that this the sequence in which drought events generally propagate. Please consider revising this paragraph.

Reply -> We have modified the order of the drought definitions as suggested.

(2) Page 1975 Lines 16-17: Do you mean a specific region or is this statement generally valid across the globe?

Reply -> Stochastic or neural networks can be used across the globe.

(3)Page 1977 Lines 21-22. Probably no need to mention section 1 here because it precedes this sentence?

Reply -> Removed as suggested.

(4) Page 1979 Line 14: I think the authors mean the real-time forecasts here which have 50 ensembles members. Please mention that in this sentence.

Reply -> Sorry, we do not understand this comment. The extended ENS up to 32 days and the seasonal ensemble forecasts both have 50 members plus the unperturbed member. Both are real-time forecasts.

(5)Page 1984 Lines 1-2 and Figure 2. If I understand correctly Fig. 2 shows correlation between observed and forecasts time series across all seasons. How do you think the fact that forecasts capture the seasonal variability (dry vs wet season) might be inflating the correlation score here?

Reply -> In Figure 2 and in this paragraph we use the SPI. The standardization is done independently for each forecast. This method removes the seasonal cycle by comparing the precipitation amount in relation to the 20 years of climatology. So here the correlation corresponds to the ACC (anomaly correlation coefficient). We have clarified this in the manuscript.

(6) Page 1984 Line 18-20: “This result….”. Please clarify this sentence. I am not sure what you mean by this.

Reply -> The sentence:

'This result could be due to the spatial and temporal characteristics of drought events that are better simulated in a global model one month ahead.'

has been modified as follows:

'This result could be explained by the usually large spatial and temporal scales of drought events that are better predictable by a global model even one month ahead.'