Responses to the anonymous Referee #1

This a very interesting paper that brings new insights about the question of past and future runoff trends. The paper is well structured and easy to follow. The key points are to combine historical data with future projections from the CMIP5 exercise, and to use a recent statistical method, the Temporal Optimal Detection (TOD) method, with interesting capacities in case of spatially heterogeneous changes.

We thanks the reviewer for his positive comments

My main suggestions are the following:

1. Explain better the advantage of the TOD in case of spatially heterogeneous changes (cf p2126, L3-11). In particular how does this relate to the principles of the method (beginning of section 2.2 until p2125 L7) ? Can this method provide significance of trends at the continental scale ? If so, why not give this information too ?

This method can also provide the significance of trends at continental scale. This information is not included in the manuscript for the simple reason that the global trend was our own preoccupation. This figure is included her (Figure I) and we choose to incorporate such figure in the new version of the manuscript as Figure 4. As excepted (published before : Labat et al. 2004; Dai et al. 2009; Alkama et al. 2011; …), only observed Africa's runoff shows significant trend. The P-value spread between the CMIP5 models is very large especially over 20th century, and reduces at the end of 21st century becoming significant (less than 0.05) over northern high latitude regions. While, over the other regions (South Asia, South and Centre America), no change is detected by CMIP5 models till 2100. These informations are also added in the manuscript (see section III.1 and the end of the section III.3).

2. Still about the statistical method, the authors could explain better why high/low α are conservative/permissive, including a definition of these terms. It would also be useful to give the null hypothesis of the inference test.

The section devoted to the presentation of the method has been re-writen and split into two subsections as suggested by Reviewer 2. The null-hypothesis involved is now explicitly given. We add some explanation regarding how high/low α impacts the TOD results: "Note that if α is high, then the internal variability ε has a higher memory effect, and it will more likely produce trendy-like time-series. Consequently, to find a significant change will be harder."
The definition of "conservative" and "permissive" is suggested in the following sentence: "Figure 2(b) (c) and (d) suggests that the more suitable choice is α=0.2 (distribution close to uniform), while α=0.3 (resp. α=0) is too conservative (permissive), leading to less (more) than expected values under the 5% threshold." Indeed, at the 5% level, a test is conservative (permissive) if under H0, the H0 hypothesis is rejected with probability lower (higher) than 5%. It basically means that, while the nominal level of 5% is announced, the true level is lower. Consequently, the change needs to be really strong to be significantly detected. We hope the new version of the manuscript to be clearer in this respect.

3. The title and the paper are focused on global runoff changes, in opposition to basin scale or continental scale changes. But, the studied changes are not really global, as they are restricted

to 161 river basins in one case, and 687 when using reconstructed data. This should be stressed, for instance by giving the fraction of global terrestrial areas that are analyzed.

As described in the first paragraph of p 2126, only 31 % (43 %) of global land area excluding Antarctica are covered by the 161 (687) rivers basins which correspond to about 42 % (60 %) of global land discharge. For easier understanding, these informations are now added in the abstract.

4. I am surprised that the influence of direct human intervention is so quickly ruled out in section 2.1 (p2122 L 9 to p2123 L11). There are many recent evidences that water withdrawals by humans did change the water balance, at large scales, even on annual means, and especially in the most recent decades (e.g. Llovel et al., 2010, Wada et al., 2010, Sterling et al., 2013).

It is true that the humans did change the water balance, at large scales, even on annual means and this exactly what we said in the manuscript but it plays minor role compared to the climate signal : "Wisser et al. (2010), have quantified the impact of irrigation and reservoir operations over the 20th century. They concluded that the expansion of irrigation and the construction of reservoirs has significantly and gradually impacted hydrological components in individual river basins. Variations in the volume of water entering the oceans annually, however, are governed primarily by variations in the climate signal alone with human activities playing a minor role". Referring to Sterling et al. (2013) new sentence speaking about the impact of land use on evapotranspiration and runoff is now added in the manuscript (see lines xxx) but we choose to not speak about the two other papers mentioned by the reviewer because :

1) Llovel et al., (2010) demonstrated that global land water storage has been changed over 2002–2009, but this changes is it due to human activity ? Does this change impact the global runoff ? These two questions are not addressed by Llovel et al. 2010, and the time series 2002–2009 is too short to give any substantial conclusions about the impact of human activity on river runoff trends.

2) Wada et al. (2010) demonstrated that the water withdrawals by humans induces the depletion of groundwater resources but its impact on river runoff is not addressed. The depletion of groundwater resources is mainly due to excessive irrigation. In our opinion, the irrigation could impact the evapotranspiration trends but cannot be considered as main driver of global runoff trends.

5. In section 3.2, why not comment about the way simulated trends fit with observed and reconstructed trends ? Also, the end of this section is no more about the comparison of simulations vs observation, but on simulated trends over the 21st century: this is not consistent with the section title, and this part is really (too) short, with no reminder of the bibliographical background on the matter.

In our knowledge, only two studies compares the simulated (using Ocean Atmosphere General Circulation Models OAGCMs) and observed runoff trends, and both of them (Milly et al. 2005; Nohara et al. 2008) are listed in this section. It is true that the section title is incoherent with the last paragraph. The new section title " Evolution of observed and simulated runoff " is more appropriate. New paragraph speaking about the simulated runoff mean, standard deviation and trends compared to the observation is added in the this section. Three new figures are also included (see Figures 6, 7 and 8).

6. In section 3.3, the differences found in “detection date” with 161 or 687 river basins raise a lot of questions. It would deserve a deeper discussion. For instance, could this also be explained by the fact that climate change is not the dominant driving factor in the past ?

It is true that the difference found in detection date with 161 or 687 river basins could raise a lot of questions and the recent direct human influence on global trends could be one of the explanations. But this hypothesis is less supported because we did not detect change even using 161 basins over the same period with the 687 basins (i.e. 1958-2004). The 161 river basins represent 42% of global discharge and the 687 basins represents about 60% of global discharge. The main difference (60%-42%) of discharge are mainly coming from the northern Canadian and South Asian rivers. Without doubt, the human influence are far from being the main driver of the Northern Canadian rivers. Whereas, the large land use and irrigation changes over South Asia could be one possible explanation of this difference. These informations are included in the new version (see section 3.3).

Minor questions/comments/suggestions:

Section 2.1 mentions about a third pre-processing (p2122, L4-5): has it been done by the authors, or does it come from Dai et al. (2009) ?
It comes from Dai et al. (2009). It is true that using the term pre-processing her could be misleading. This term is replaced by ‘product’ in the new version.

p2122 L9: in addition to beING
p2122 L28: upstream in one word; not without final e
ok, done.

p2124: the first sentence of section 2.2 is not very clear to me
This sentence has been rewritten in the newest version of the manuscript.

p2124 L20: I suspect that “as” should be changed to “a(s)”
p2126 L 22-23: what is explained for year 1980 is true for any year
ok, done.

p2126, M23-26: this sentence is misleading: the larger year-to-year variations are the ones of the P-value and not of runoff, and the main point is not about the larger variations in P-values, but on the larger P-values.
The change in the annual runoff is detected at the year “t” only if the P-value(t + i, $\forall$ i $\geq$ 0) is less than the significance threshold value and P-value (t-1) is greater.

This sentence is added in the text (see xxx).

P2128, L18: change “Even” to “Despite”
p2129 L9: add “model” after CMIP5
p2129 L11; replace “average” by “central”
ok, done

We thanks the reviewer for his careful reading which help us to improve our manuscript.

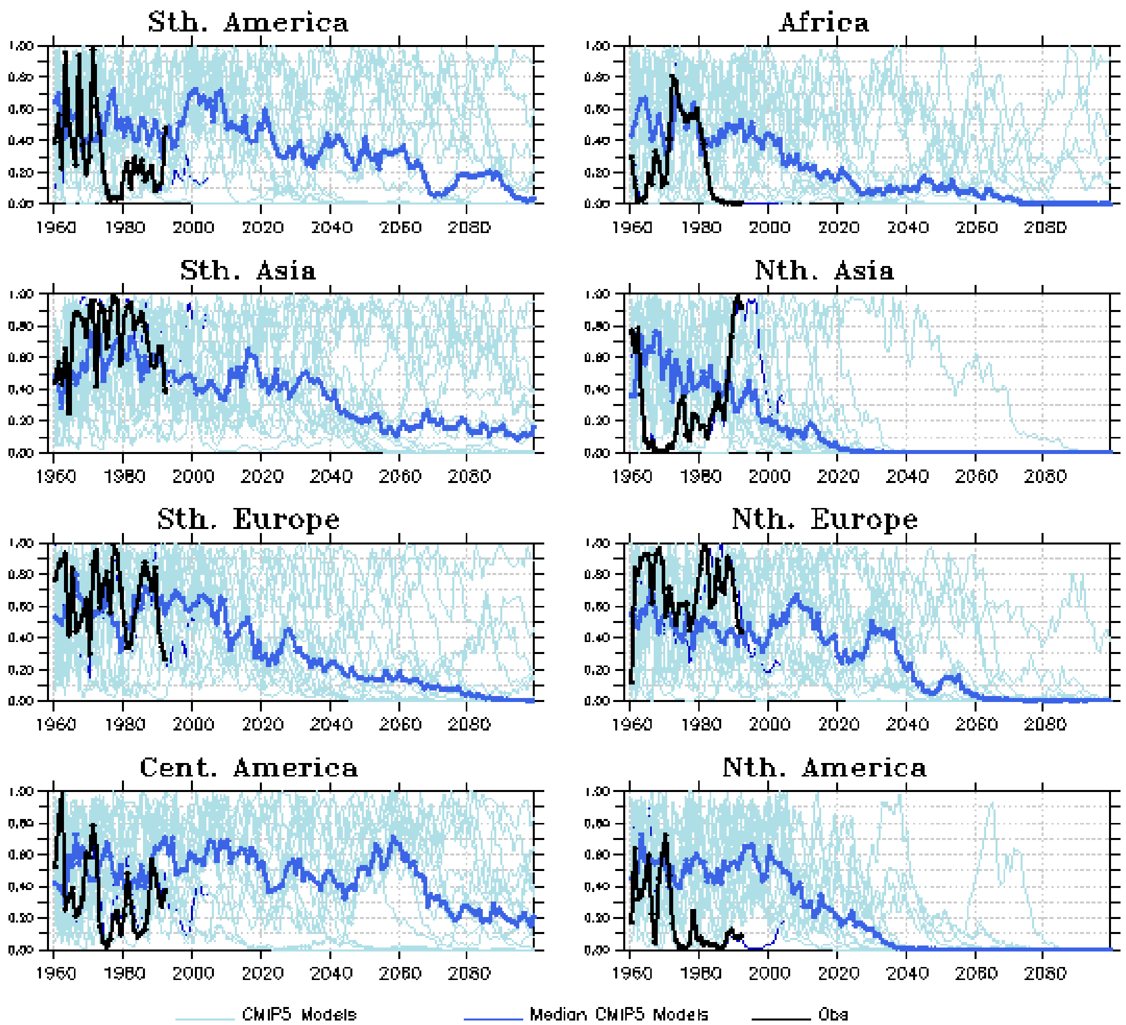

Fig I. Temporal P-value of observed (black) and simulated (light blue) runoff over the 8 regions merging 161 river basins using α at 0.2. The median of the 14 is in blue.

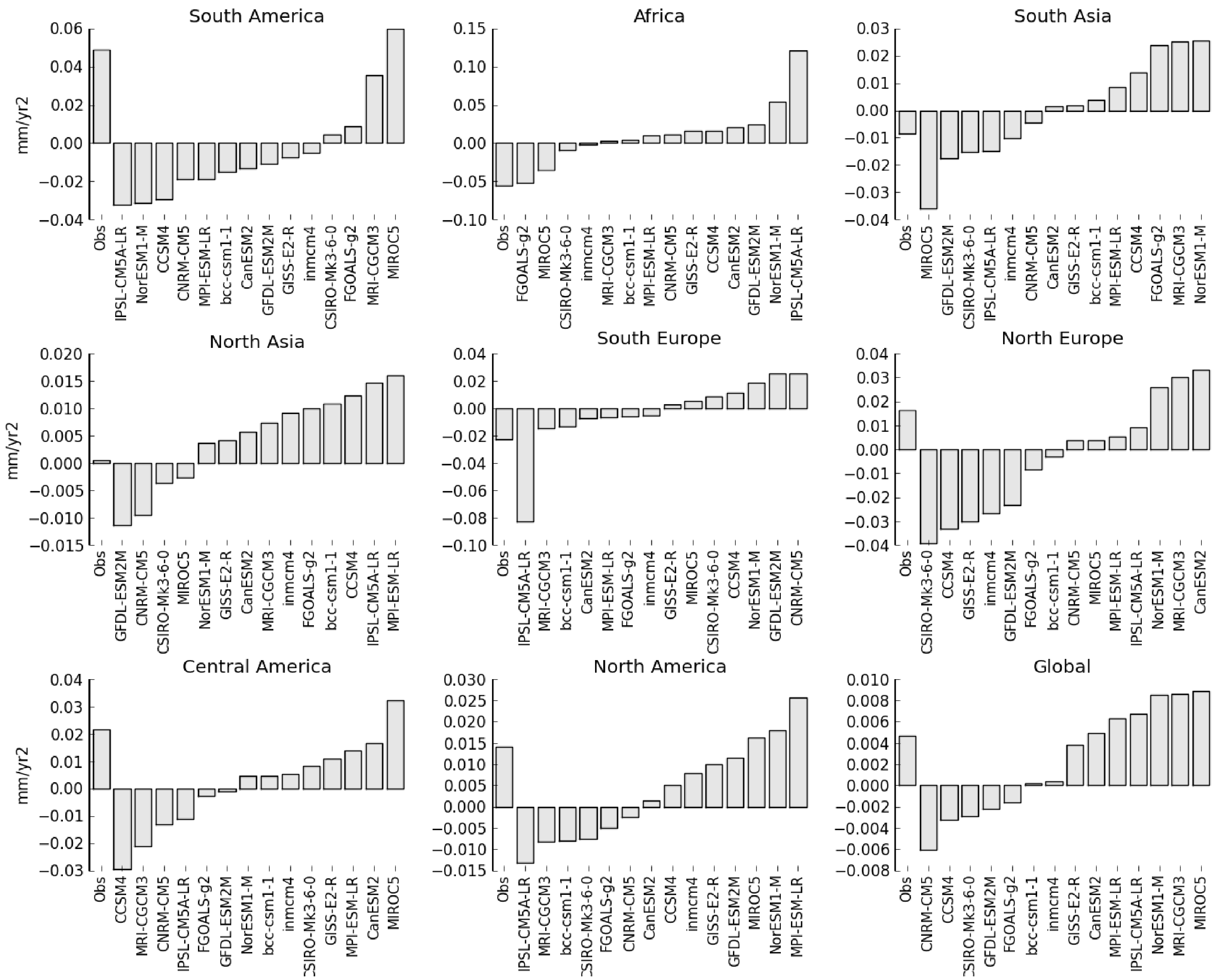

Fig II. Observed and simulated runoff trends over 1958-1992