

General comments

1. The results are clear and I do agree that it's an interesting topic which indeed needs more attention. The main part of the analysis is only applied for the control catchment and not for the treated catchment; based on the non-explicitly mentioned assumptions that only the control catchment behaved non-stationary during drought. The primary objectives are based on the concept that a catchment may not show non-stationary rainfall-runoff relationships during changes in climate. This need to be well introduced and discussed where appropriate in the manuscript.
 - A. Your hypothesis is based on the assumptions that only the control catchment behaved non-stationary during drought. Are you sure that this is the case and that it's fair to make this assumption?
 - B. The primary objectives are based on the concept that a catchment may not show non-stationary rainfall-runoff relationships during changes in climate. Can you add references stating that it's not "allowed" for a catchment to show non-stationary behaviour?
2. It's a flaw that the manuscript lacked detection of changes in the rainfall-runoff relationship for the treated catchment. Although it's obvious why there is a control and a treated catchment, I think it's important to apply the appropriate parts of the applied analysis for the treated catchment, also to have a better understanding of the underlying processes; to confirm or to deny the non-linear behaviour of the treated catchment.
 - A. If the treated catchment has not been reforested, did you expect the treated catchment to show similar behaviour as the control catchment?
 - B. Does the treated catchment already shows the same non-linearity as the control catchment?
3. You also may consider to add more information about the treatment (land use change), and the estimated changes in PET and rooting depth? The latter especially in introduction/area/discussion. As a reader I do not find the evidence that the processes at the treated catchment did not changed.
4. Figure 6 / Figure 8 / Area & Methods / results; It's not mentioned that before 1990 there was relatively more precipitation (Figure 2), which could have a major impact on your results because of the "memory" (storage) capacity of the catchment / soil's / geology.
5. In lines 428-431 you write "The common assumption of the three methods is that the interaction between climate variability and vegetation changes in very small and can be ignored" and that "The total changes of runoff are a linear combination of runoff changes caused by climate variability and vegetation changes". What is your opinion about this assumption? And what does your results imply? And what does other literature states about the assumption and the linear combination?
 - A. I suggest to strengthen your paper to dive deeper in these assumptions about linearity. You already started with this by mentioning "only the sensitivity-based method uses the change of rainfall and potential evapotranspiration to obtain the runoff change caused by climate variability" (lines 436-438);
 - B. Does the statement in lines 436-438 imply that the other two methods are not suitable to apply in your case?

Specific comments

6. Although it's possible that title need to be edited after revision of the manuscript, at the moment the title is not entirely clear to me, "non-stationarity" "invalidates the role of control catchment" (I suggest something like: Drought-induced non-stationarity in the rainfall-runoff relationship dismisses valid comparison with the control catchment at the Red Hill paired-catchment experimental site);
7. Lines 94-95; "The other reason is related to the non-stationarity rainfall-runoff relationship of the control catchment". Do you mean that only the control catchment showed a non-stationary relationship? So the afforested catchment did show a stationary rainfall-runoff relationship? Be clear, because this is important information to have not only a proper understanding of the rest of the paper, but this also affects your hypothesis and objectives.
8. In line with previous comments, I suggest to rewrite the conclusion and abstract after additional analysis and new input on the assumptions.

Technical corrections

9. Line 31: "experimental site, using experimental observations", I would suggest to change the second "experimental" to "field" (or in-situ, if that is what you're pointing at);
10. Lots of repetitive information or sentences which can be shortened throughout the manuscript;
11. In consequent reference of Fig. and Figure;
12. Line 42; perhaps add a more recent paper;
13. Line 58; "paired-catchment method is based on paired-catchment experimental observations", I would suggest to change "experimental observations" to "field or in-situ observations".
14. Line 59; perhaps add a reference to prove the "standard" approach;
15. Lines 62-67; references;
16. Although it's not the "area" section, I do think you make your case stronger if you already mention the land use history of both catchments, as well as land use change during the evaluated measurement period.
17. Line 105; what is the definition of much longer? Or even leave out "much".
18. Lines 125-128. In the past (Zhao et al., 2010) 16 years of observations where used, so now you're using the same dataset, with 10 years of additional data? So you also compare the present results with those of Zhao et al. (2010)? If this is the case, be clear about this in objective nr. 1.
19. Line 137; dominant soil texture?
20. Line 138: average slope? For both sites?
21. Lines 138-139; Sentence may be removed to next paragraph, it feels misleading because of the information which is "missing" in line 139, but actually described from line 143 on.
22. Lines 139-147; what is the variation in monthly rainfall? Seasonal? What Köppen climate?
23. Line 147-148; "potential evapotranspiration records....." yes, what? Ranges, values? Differences in AET between control and treated catchment?
24. Lines 149-151; I am curious to understand why you do show the prolonged drought for Kileys Run, but not for Red Hill;
25. Line 157; I don't want to be a nit-picker, but you say short length of the observed data record. I am not aware of many locations where they have a data record for this very nice (long) period of 25 years.

26. Line 248; you do introduce the parameters SC and C before giving any of the related equations.
27. Line 302-307 and Table 1; beta what? Add name of method;
28. Lines 308-310; unclear;
29. Figure 4.
 - A. The left figure is Kileys Run, but in the caption you mention Red Hill first, change either the caption or the figure order.
 - B. The x- and y-axis do have different ranges, this makes the graphs difficult to compare and to understand the meaning of the results.
30. Lines 310-314; add name of method (so the reader can go back to paragraph 3..... to understand the applied method);
31. Line 312; it's conflictingly that you express Q_{RH} first as Q_{KR} and after expressing Q_{RH} at as P_{RH} , as a reader I cannot compare Q_{KR} as P to see the differences in slope or offset based on P. Can you express Q_{kr} as P as well?
32. Lines 355-356; you made your point about the differences between the periods, but you may consider move this sentence to the discussion and use a physical understanding and references to explain the cause of drought;
33. Figure 6; you may consider add the daily flow duration curve for the entire measurement period (1990-2015) to indicate the differences;
34. Line 373; which method? Refer to paragraph....
35. Figure 7; The x-axis of the graphs could be better aligned;
36. Line 385; correlated with catchment runoff; you may consider adding the R and p value;
37. Line 404-405; "by using the method mentioned in section 3.3" why not mention the name of the method and refer to the section?
38. Figure 9; no data for 1999, 2000, 2006, and 2007? Not mentioned in the text/method section? Or did I misread something?
39. Figure 10; does it shows median values for a period or means, or?