Responses to comments from Student Reviewer 1

Key:

SC1 – Comments from Student 1

AC – Author comments

General comments

SC1: In my opinion, the statements made by the authors about increased transpiration being the cause of decreased runoff found in WS77 are not supported by their results. The authors assumed solely from Jasechko et al. (2013) that transpiration must be the biggest driver in total evapotranspiration rates. However, Jasechko et al. (2013) states that every individual catchment does show a unique evaporation trend as they all have a unique climatology and hydrology. In addition to this, the method used by Jasechko et al. (2013) can be disputed as water from lakes is used to determine the isotopic composition. However, it is not exactly known what the sources are of this lake water. Therefore I think it is rigid to just assume, based on this article alone, that transpiration is the largest component of the total evapotranspiration of the studied catchments. And as said before, several other studies show that especially in a forested catchment the fraction of evaporation from interception can be of considerable importance to include in the hydrological analysis.

I would suggest to reconsider the terminology used and to also include the changes in interception evaporation as a possible explanation given its importance in temperate forests. Furthermore, it would be good to also cite other papers about the partitioning of evapotranspiration which are specifically focussing on forested catchment (e.g., Bryant et al., 2005, Klaassen et al., 1998).

AC: The reviewer has provided a constructive comment and suggestion to consider the role of canopy interception besides just transpiration as an influencer of streamflow. We agree with this. However, the portion of the interceptional loss in this humid coastal plain is substantially lower than the transpirational loss as shown by some of our own data. Based on canopy throughfall measurements made for 32 rainfall events from July 2003 to February 2004 on the control watershed (WS80), Harder (2004) found an average interception loss of 11% of the gross precipitation with a canopy storage capacity of only 0.7 mm for this pine and hardwood mixed forest. Both of these were lower than the 17.5% average canopy interception and 1.9 mm canopy storage capacity found for the managed loblolly pine (*Pinus taeda L.)* forests in humid coastal North Carolina (Amatya et al., 1996) in their five year study. In other earlier studies, Hoover (1953) obtained an interception loss of only 9.6% of the precipitation based on data collected from 67 storms on loblolly pine forest in Charleston, SC. Similarly, even for all pine forest Hewlett (1982; p 76) provided a formula to estimate annual interception as 0.10*precipitation (in cm) + 0.10*n (where n = number of storms). In general average annual rainfall at SEF is 127.0 cm and average number of sizable storm average is about 12. With that the average annual interception becomes 0.10*127 + 0.1*12 = 13.9 cm = 10.9% of 127 cm of average precip.

Hoover, M.D. 1953. Interception of rainfall in a young loblolly pine plantation. U.S. Forest Service, Southeast Exp. Station, Paper 21, 13 p.

Hewlett, J.D. 1982. Principles of Forest Hydrology. The Univ. of GA Press, Athens, GA30602, U.S.A. 183 p.

In another recent modeling study using DRAINMOD-FOREST on watershed (WS80), Tian et al. (2011) found the simulated canopy interception varying between only 8-17% with an average of 12.5% compared to 11% measured for the same WS80 forest for 2003-04. However, the average for the 5-year (2003-07) simulation study was only 10% with 69% of precipitation lost to transpiration and soil evaporation, which is almost negligible for the closed canopy.

In another long-term simulation study Dai et al. (2013) used combination of models including MIKESHE to simulate streamflow and ET of the larger watershed WS79 that contains both the WS77 and WS80. The authors used a range of LAI between 0.2 - 6.6, measured at multiple experimental plots on the watershed (WS80) using a LiCOR2000 canopy analyzer equipment, with an average of 2.8 with a canopy interception storage between 0.05 to 0.80. The simulated average annual ET was 1043 mm or 76% of the precipitation for the 1946-2008 period. Although the authors did not report the canopy interception part, it is assumed that it should also be within the ranges reported by Harder (2004) and Tian et al. (2011) for the given small values of LAI and canopy storage capacity.

Although there was no historical interception study conducted on these experimental forests these limited data from 2003-07 after the regeneration and coming back to the original base line levels clearly demonstrate that the total ET is dominated by transpirational losses in this resilient coastal pine and hardwood mixed forest system.

References:

Harder, S. V. 2004. Hydrology and Water Budget of a First-Order Coastal Plain Forested Watershed, South Carolina. M.S. Thesis, Master of Environmental Studies, College of Charleston, Charleston, SC.

Dai. Z., C.C. Trettin, and D.M. Amatya. 2013. Effects of Climate Variability on Forest Hydrology and Carbon Sequestration on the Santee Experimental Forest in Coastal South Carolina. USDA Forest Service South. Res. Station, Gen. Tech. Rep. SRS-172, 32p.

Tian, Shiying, M. A. Youssef, and D.M. Amatya. 2011. Unpublished Data. A report to USDA Forest Service, Center for Forested Wetlands Research, Cordesville, SC submitted by Dr. M.A. Youssef at Department of Biological & Agricultural Engineering, North Carolina State University, Raleigh, NC 27695 (Contact author: stian@ncsu.edu) -

Amatya, D.M., R.W. Skaggs and J.D. Gregory. 1996. Effects of Controlled Drainage on the Hydrology of a Drained Pine Plantation in the North Carolina Coastal Plains. J. of Hydrology, 181(1996), 211-232.

Tian, S., M.A. Youssef, R.W. Skaggs, D.M. Amatya, and G.M. Chescheir. 2012. DRAINMOD-FOREST: Integrated modeling of Hydrology, Soil Carbon, and Nitrogen Dynamics, and Plant Growth for Drained Forest. Journal of Environmental Quality 41:764-782 (2012).

SC1: A second general comment concerns the technique of moving sum of recursive residuals (MOSUM) used by the authors to determine changes in runoff. This technique is not often used in hydrological studies and the authors should therefore elaborate more on how the method has been applied, why they have chosen it and what the possible constraints and advantages of this method are (e.g. that a moving-estimates test is more sensitive to parameters that are temporarily unstable than for instance the CUSUM test because in the latter, the cumulated sums are used which become less sensitive to parameter changes as the number of residuals becomes larger (Chu et al., 1995)).

AC: Section 3.5 has been re-written to provide more background on MOSUM test. Additional information includes literature on use of MOSUM to detect temporal changes in land use and other eco-hydrologic variables. We have also included the general equations for determining recursive residuals and the moving sums of recursive residuals (MOSUM). We have included an extra sentence as suggested by the reviewer to justify the choice of MOSUM over CUSUM. It relays the same information as the reviewer's suggestion. We have also included explanations on the differences on how structural changes, number of structural breaks (break points or break dates), and the location of the structural breaks are detected.

SC1: In addition to that, to me it remains unclear how the authors dealt with the periods of missing data in figures 4 and 5 which result from using the 'strucchange' package in R implemented by Zeileis et al. (2012). It is mentioned that non-linear axes are used to deal with the data gaps, but this seems not to be the right terminology. The axis itself is linear, but the data within is not linear due to the missing data. To clarify the figure I would suggest implementing arrows or vertical lines at the place of the two gaps of missing data. To make it even more clear the x-axes can be renamed to for instance: number of day with observations. Another more elegant option would be to add another panel which shows the original data with its gaps in order for the reader to understand the relation to the data shown. In this way the interpretation of the figures would be easier.

AC: The implementation of MOSUM using the "strucchange" R package automatically ignores missing data and rescales the time axis form 0 to 1 using the number of available observations. The above explanation is included in the revised manuscript. Therefore, one has to keep track of the dates corresponding to each observation number.

The use of the term non-linear is excluded

Based on the suggestion of the reviewer, the time axis is renamed to "Number of months with observations (scaled from 0 - 1)"

SC1: Furthermore, in section 3.5 it was mentioned that analysing the points on structural change, only a single breakpoint was assumed (p.11527, l.16). I would suggest to add to this explanation that one breakpoint, can generate more than one break dates (as I now had to find out myself using the package 'strucchange' in R), especially because now in section 3.5 it appears for the reader that breakpoint and break date are the same (p.11527, l.11).

AC: One can set a maximum number of breaks for the time series using the "breaks" parameter in "strucchange". For our analysis we set it at 3 because of the flip and flop eras. The mention of one in the original manuscript was an oversight and we thank the reviewer for highlighting it.

The reviewer's observation of breakpoint and break dates being the same are true . The breakpoint is the observation number while the break date is the rescaled value corresponding to the observation number for data with missing values or actual date for data with no missing values.

SC1: All in all, I think both figures 4 and 5 need a better explanation on how the break dates are obtained because the explanation given in section 3.5 and the caption beneath figure 4 are confusing as they refer to two different 95% confidence intervals; 'the long horizontal dotted lines', but also the 'small horizontal lines that cross each break date'. Now the break dates especially of the second panel in figure 5 not seem to fit the description given in the caption of figure 4.

AC: The paragraph below has been added to section 3.5:

The MOSUM test for change detection follows a three step procedure. The first step checks for existence of structural change based on the assumption that the variability of the moving sums of recursive residuals under structural stability follow a Brownian motion (a random walk) with an expected mean of zero. If the MOSUM cross the 95 % confidence boundary, then structural change is detected. For details on the technical basis and the asymptotic function of the 95 % confidence boundary, readers are referred to Zeileis et al. (2012). When structural change is detected in the first step, then steps two and three determine the number and location of the change points (break points or break dates).The break points and corresponding 95 % confidence intervals are estimated based on methods developed by Bai (1994, 1997) and Bai and Perron (1998) and implemented by Zeileis et al. (2012). The second step determines the number of break points by minimizing the Bayesian Information Criterion. However, one can predefine the maximum number of breakpoints for a given time series. For this analysis, this number was set to three to account for flip and flop periods. The third step iteratively determines the location of the break points by minimizing the regression sum of squares.

Based on the above procedural implementation of MOSUM, the position where the MOSUM cross the 95% confidence boundary is not always the location of the breakpoints. Also, when the MOSUM returns within the 95% confidenceboundary, it does not mean the relationship has regained the previous structural stability. Finally, the strength of the linear relationship, the size of the moving window, and the number of predetermined breakpoints influence the location of the breakpoints.

The two different 95% CI allude to two different tests. The first test is for existence of structural change. Therefore, the long horizontal dotted lines represent the boundary for this test. If the first test confirms existence of a structural change, then the second test is about breakpoints (the number and location of breakpoints-previously defined as steps two and three). For each breakpoint, there is a corresponding 95% CI because they are estimated using empirical fluctuation test framework. These are the small horizontal lines crossing each breakpoint.

SC1: Another general comment concerns the interpretation of section 4.3 about seasonal streamflow trends and its accompanying figures 7 and 8. This is in my opinion only briefly discussed in the discussion section and therefore I would recommend some further elaboration on the significance of difference in magnitude found, and if this significance does hold for all months or that there is a seasonal effect, are necessary to comment on, as this is done in all other parts of the results. The authors claim in the discussion section that changes in the transpiration rate must be the cause of the differences found between streamflow of WS77 and WS80 throughout most of the year. However, alternative explanations are possible. The effect of soil moisture can be important especially when focusing on seasonal trends (Johnson and Kovner, 1956). In addition, especially when interpreting figure 7, a note should be made about the interception term as the relative importance of transpiration and interception do change as the forest grows older (Naranjo et al., 2012) during the Flop era in comparison to the Flip era. In the current version of the manuscript, section 4.3 is not related to the rest of the results and its significance is not clear. But it could be an interesting section as Brown et al. (2005) found that there are found impacts on seasonal yield and flow regime due to vegetation change, but that these impacts are unique per catchment. Therefore it would be interesting to be able to determine if there is a significant seasonal effect for these paired watersheds.

AC: After reading this reviewers notes and those by other reviewers, we too are of the opinion that short term or seasonal trends in flow do not add to the manuscript. We are of the opinion that this section detracts from the larger objective to characterize decadal-long changes in relative hydrologic character between watersheds. We are therefore proposing a removal of the section related to monthly streamflow trends. Please also see our response related to this issue in our comments to Anonymous Refereees.

SC1: Another comment refers to how this work relates to other studies on land use effects on evapotranspiration, for instance the recent work by Naranjo et al. (2012) who studied the change in discharge and actual evapotranspiration both relative to the values found before the disturbance using K-curves. This paper shows another way to determine structural changes even at longer temporal scales. It could be interesting to see if a same relation can be found when creating such a K-curve with the data the authors used, and to note how it will evolve on the longer time scales (100 years) and then to compare the first few years with the results from the LOESS analysis done.

AC: We thank the reviewer for this suggestion. We will certainly add the s Naranjo et al. (2012) study to the manuscript as background literature. However with all due respect to our reviewer, recreating K-curves simply to test methodology is beyond the scope of this study.

SC1: The last general comment refers to the introduction where the authors mention the limitations of the paired watershed approach, but they do not justify why this approach is valid for their application. Did they have no other options regarding the data they have got, or is it an appropriate method for the aims they want to achieve?

AC: We thank the reviewer for pointing this out. An anonymous reviewer also pointed out the fact that we suggest the paired watershed approach is only described in negative terms. We did not mean to imply that the paired-watershed approach is not a useful concept – we erred in pointing out its limitations with providing the strengths of the approach. We will rephrase the introduction to include both strengths and weakness of the paired watershed approach, highlighting how our study takes advantage of the strengths of paired watersheds to overcome some of its weaknesses.

Specific Comments from SC1

p.11521, line 16; which is believed to be related to soil water and depression storage. (AC: thank you)

p.11522, line 8; most of the temperate knowledge is from a 1938 hurricane... (AC: thank you)

p.11524, line 2; WS77, the treatment watershed, (first comma should be added). (AC: thank you)

p.11530, lines 8-9; ...altered the reported break dates. The lines above suggest that the break dates of the MOSUM analyses of watershed-specific rainfall-runoff relationships is meant (figure 5), but it is not stated clear. It could be interpreted in more than one way. (AC: please see statement of revision of this paragraph and that related to MOSUM analysis)

p.11533, line 4; 4.3 mm2 ha−1 should be 4.3 m2 ha−1. (AC: thank you) p.11533, line 12; Table 9 should be Fig 9. (AC: thank you)

Table 1; In the column 'Years' the end date of the Flip era should be 2003 and the end date of the Flop era 2011. (AC: thank you)

Figure 1; The locations of the weather stations are not clearly visible. (AC: thank you – we will redo figure 1)

Figure 6; In the description of the second box seasason should be season. (AC: thank you)

Figures 7 and 8; When printing in black-white, they are not readable anymore. A suggestion is to use different line types. (AC: thank you – figures 7 and 8 will be removed for other reasons)