First we would like to thank the anonymous referee for the thoughtful and constructive review of our manuscript. Original comments by the reviewer are in normal text and our responses are in bold.

The authers use a variety of time series analysis techniques to investigate observed change in different aspects of the streamflow regime of four large river basins in the Midwestern US (in Minnesota, Wisconsin and Illinois), and to attempt to relate the observed changes to changes in precipitation magnitude and agricultural land use. The techniques and breadth of the analysis are strong, and the paper represents an interesting contribution to the on-going discussion of observed streamflow changes in the region.

I do have a couple of overall concerns. The potential impact of dams is mentioned in a few places in the manuscript, but there is no explicit mention of what streamflow gauges were used and which were affected by large dams and why these were included. I do not see how this type of analysis (looking at changes in high and low flow dynamics in river basins) can possibly be valid if using stream gauges from downstream of mid-to large size reservoirs. If gauges have been included that are affected by reservoir storage, then this represents a fundamental flaw and I think these gauges should be removed. If the gauges have not been impacted by dams, this should be clarified.

Response: This is an important point for us to clarify in the revised manuscript as our original discussion, which provided more detail than necessary, created a point of distraction for multiple reviewers. In reality, the dams have little to no impact on the streamflow gages we used for the analysis. In Section 3.3 (p. 11, line 10) we explain that multiple gages, explicitly listed in Table 1, for a single basin were used to calculate seven annual flow metrics, and only the basin outlet gage, identified as the downstream gauge in Table 1, was considered in other analyses. Our rationale for considering these gages is to determine whether or not we can detect the influence of tile drainage even at these large scales. For large watersheds we inevitably run into cases where some gauges have upstream dams. There are over 90,000 dams in the United States according to the National Inventory of Dams (United States Army Corps of Engineers, 2016). Dams may offset or dampen streamflows in response to increased precipitation and drainage in these basins. However, most of the dams in our study basins are small and were constructed in the late 1800's and early 1900's (Graf, 1999; Lian et al., 2012; United States Army Corps of Engineers, 2016). Therefore, the effects of these dams would have been established well before our study period. For example, in the Illinois River basin water budget we consider the downstream gauge at Valley City, IL (USGS gauge 05586100) during the period 1939-2011. All major dams on the Illinois River had been completed by 1939. Based on work by Lian et al. (2012), streamflow changes post 1938, specifically peak flows, have been influenced more by climate than dam operations (though they did not consider the effects of drain tile). The one exception might be the uppermost Illinois River Basin, which has been influenced by expansion of the Chicago metropolitan area. For that reason, we redid the IRB analysis using only the agricultural basins (inset in Figure 3a) and find that the trends are slightly amplified for the agricultural basins. We briefly discuss the possibility that this suburban development may have offset or damped the hydrologic response to increased precipitation and tile drainage (page 1, lines 29-30; page 8, lines 24-26; page 30, lines 9-14). We will clarify the language and better articulate our rationale for including gauges with upstream dams in the revised manuscript.

Overall, I would like to see more discussion and interpretation. In my opinion, the over-simplification of comparing annual precipitation and streamflow trends as a method of trend attribution neglects understanding of the non-linear nature of streamflow generation and the complexities of climate impacts on hydrology. We know that in many locations precipitation intensity is increasing, and we know that if the same amount of precipitation falls in a shorter time period, the runoff ratio is likely to be higher, but nowhere in this paper do the authors acknowledge this fundamental relationship.

It is true that the authors have done quite a bit more than a simple comparison of annual streamflow trends, but I maintain that it is theoretically possible to have no change in monthly or annual precipitation magnitude and still have an increase in streamflow, with no land management changes. (Although yes, in some cases the response may be amplified by the presence of drainage.) There was also no discussion of the influence of other climatic changes, such as changes in snow and frost depth that can have strong controls on the strengthening of the semi-annual period streamflow response seen in Minnesota.

Response: We agree with the referee that it is possible to observe changes in streamflow while having no change in monthly or annual precipitation magnitudes. High intensity, short duration events yield higher runoff ratios in poorly drained soils. Additionally warmer winter temperatures, earlier snowmelt, and more days when winter precipitation falls as rain instead of snow should most definitely affect and even increase winter baseflows, decrease the timing of ice break-up, and affect the magnitude of snowmelt floods. Several studies have documented such hydroclimate changes in the Midwestern USA (Feng and Hu, 2007; Groisman et al., 2001; Higgins and Kousky, 2012). Observed increases in runoff ratios could play a role, such that increased precipitation could lead to increased soil moisture, despite tile drainage impacts, and result in a nonlinear increase in runoff generation for similar precipitation events in the post-period. However, no theory exists to predict how big this effect could be on landscape scales. Furthermore there are very limited soil moisture data to determine whether or not soil moisture has in fact increased despite the immense amount of additional tile drainage that has been installed in the past few decades. Investigating this effect would be a good future step in this line of research. We have presented a range of streamflow metrics (including peak, low, and mean flows) at several scales (daily, monthly, and annual) in the manuscript, and "we acknowledge that the conversion of precipitation to streamflow occurs by a complex suite of physical processes", page 4 line 27. Using multiple lines of evidence from the analyses of individual basins and the basins taken together we stand by our conclusion that agricultural drainage activities have likely amplified the streamflow response to relatively small changes in total precipitation. We thank the anonymous referee for suggesting more discussion on the non-linear nature of streamflow generation and the complexities of climate impacts on hydrology. We will elaborate on these points in the discussion section of our revised manuscript.

I also would like to see more of a discussion of physical mechanisms involved. One of the largest streamflow trends displayed is that of an increase in summer low flows, and it is implied that this is due to an increase in the intensification of agricultural drainage. Field observations of subsurface drainage from around the Corn Belt consistently show that the subsurface drains stop flowing in the late summer, during the summer low flow period. This is true even in Minnesota, where the drains tend to be deeper than in other parts of the Corn Belt. If the drains are not contributing to streamflow in these months, can they be contributing to an increase in summer low flows? I can see that in areas where surface ditches are dug so deeply that they are intercepting a greater proportion of regional groundwater flow there is the potential for sustained baseflow to streams in the late summer, but I think this would only be true in a few isolated cases, with very substantial main stem ditches.

Response: While tiles do not flow continuously throughout the summer months in Minnesota, we maintain that tile drainage flows contribute substantially to streamflow during summer months. For example, we have runoff data from an edge-of-field site in Blue Earth County monitored with a

surface flume and at the outlet of the subsurface tile system from April 16, 2013 – September 9, 2013 (obtained from Minnesota Discovery Farms). In June 2013, 93% of runoff flowed through the tile system and only 7% left the field as surface runoff. In July, 100% of the runoff flowed through the tile system. In August 15% of the runoff flowed through the tile system. Certainly this varies from location to location and year to year, but we believe it provides sufficient evidence that the tile systems do indeed flow during summer months in our study area. It would be helpful if more of this type of data were available!

I have included some more specific comments below, tied to specific locations in the text:

1. Section 4.4.1 agricultural land use is not the only thing different between these river basins - climate is also very different. In particular, the seasonal timing of subsurface drainage tends to be very different between Minnesota and Illinois, due to the influence of soil frost.

Response: We acknowledge that land use is not the only difference between these river basins and discuss other differences (such as temperature, precipitation, soils and lithology, and presence of dams) in Section 2. It is true that the seasonal timing of subsurface drainage may be very different between Minnesota and Illinois. We know that the greatest total river runoff occurs in April for the RRB, MRB, and CRB, coinciding with spring snowmelt, and May for the IRB as snowmelt and soil frost exert less importance on total monthly flow than rainfall in the IRB. However, we simply do not have the information necessary to account for how similar or different subsurface drainage might be from month to month or season to season between basins due to a lack of available tile drainage data. This is an important point that we emphasize in the paper. We call for better documentation of agricultural drainage practices (page 2, lines 5-6; page 29, lines 19-21; page 30, lines 28-30) so that future studies may be able to better address this problem.

2. Page 2, line 1. define was is meant by streamflow for these percentage changes - average annual?

Response: Percentages refer to increases in the mean of the seven annual flow metrics (pre-1975 and post-1975) reported in Figure 4b. Some revising in the abstract and discussion will be done to clarify this point.

3. Section 2, general. I agree with previous comments that the paper is very long. In this section some of the detail regarding physiography and sediment generation could be removed, instead increasing discussion of differences in hydrologic or drainage regime in these basins.

Response: In response to all referees' comments regarding the length of the paper, we will reduce the text in the revised manuscript, perhaps by 1000-1500 words, but we feel that shortening the paper more than that would eliminate essential content. Sections we plan to reduce include the context/introduction, methods and study areas, presentation of annual changes (section 4.4.1) and water budget (section 4.4) results, and redundancies in the discussion section. The section on sediment generation provides an important impetus for the study, but we agree that it can be shortened.

4. Page 10, line 8; page 15, line 23. Some discussion of differences in both the extent and physical impact of surface and subsurface drainage would be useful - surface ditches are generally deeper, and the impact of surface drainage on things like peak flows has been much more clearly established.

Response: Unfortunately we cannot entirely follow the page/line references provided by the referee. We feel we have sufficiently discussed the extent and physical impact of drainage in Sections 1.2 and 4.1 and prefer not to add additional text unless there is something specific the reviewer feels is missing.

5. Page 11, line 15. High flow days and extreme flow days are never defined. How are these calculated?

Response: Complete definitions of the seven different flow metrics can be found in Novotny and Stefan (2007), as cited in the paper. The number of high and extreme flow days refers to the number of days in a given year when mean daily flows are one and two standard deviations above the mean, respectively. We will consider adding that simple definition in the revised manuscript so readers do not need to refer back to the original paper.

6. Page 11, line 18. I don't understand how the multiple gauges were combined into one metric time series, or why. Late in the manuscript it seems to indicate that some of these gauges are affected by dams (page 18, line 23; page 30, line 23). I do not see how the inclusion of gauges affected by dams can be justified in this analysis.

Response: Multiple gages are combined into one metric by normalizing annual values at each gauge by the mean (1950-2010). This allows for gauges of different drainage areas and dramatically different values (peak flows versus low flows) to be plotted consistently together to facilitate comparisons between basins. We address the point about gauges affected by dams earlier in this reply and will clarify why they are not problematic for our analysis in the revised manuscript.

7. Page 13, line 22. The Livneh simulations used static vegetation, so there is no crop transition in this dataset, given that you have attributed part of the transition in Q to changes in ET seasonality is this a problem?

Response: The referee is correct that Livneh et al. 2013 (L13) used static vegetation. Referee #1 further pointed out that they did not account for artificial drainage in their simulations of ET. We address these concerns in our reply to Referee #1. In short, we have taken advantage of the fact that crop change and agricultural drainage were not included in the ET model. This is what allows us to test, external to the ET predictions, whether or not a LULC effect exists. There is no evidence of regional groundwater change and the effects of dams on streamflow are well known and will be discussed more explicitly in our revised manuscript. Note: we have not attributed the transition of Q changes to changes in ET seasonality on page 13. Here we are clarifying how we interpreted the water budget results in Section 4. We will clarify this point in the revised manuscript.

8. Page 17, line 8. I think 10 - 21 years is more than slightly different.

Response: Page 17, line 5 – the word "slightly" will be removed.

9. Page 18, line 18. Was there a different statistical test done on this period (post-1995)? What date ranges were used?

Response: In Section 4.3 we report the results of t-tests performed on the seven different flow metrics to compare means pre- and post- 1975 for all basins and pre- and post- LCT for the MRB and IRB, defined in Table 3. For the RRB, the post- LCT only includes 10 years of data (small sample size) and a

rapid increase in soybeans was observed beginning around 1995. Therefore we used pre- and post-1995 rather than pre- and post- LCT for the RRB.

10. Page 18, line 23. How were cyclicity and synchronicity defined/quantified?

Response: We have identified an increase in synchronicity and cyclicity of flow metrics in the MRB and RRB simply by visual inspection of Figure 4a. Synchronicity is meant to refer to all metrics exhibiting the same trend at the same time, and cyclicity is meant to refer to the ~11 year cycle that becomes more apparent, especially in the MRB, after about 1980. It is not a critical point in the paper, but we believe it is a sufficiently interesting qualitative observation to point out that may stimulate future research.

11. Page 21, Figure 5. Could the increase in power associated with sub-annual periods indicate an increase in winter thawing/ winter flows, which may be exacerbated by drainage?

Response: Winter thawing/winter flows may be one mechanism that could explain an increase in power associated with sub-annual periods. However, we do not believe we can attribute such energy changes to a single mechanism from the continuous wavelet transform plots. We will consider expanding the explanation of multiple factors that could explain the patterns observed in Figure 5 in the revised manuscript.

12. Page 23, Figure 7. This is somewhat of a general comment. I agree that the lack of change in the Chippewa River basin relative to the other basins is striking, but it must be acknowledged that this is not a true control, the climate and physiography are different, with the lowest observed precipitation change for this watershed. Figure 7 also illustrates that the kernel density of monthly streamflow for the CRB exhibits a distinct shape from that of the other basins for the "pre-" period. It seems that it might look more like the "post" period for the other watersheds?

Response: We agree with the referee that the Chippewa is not a true control and believe we have sufficiently acknowledged physiographic and climate differences. It would be unreasonable to expect to have a true experimental control for evaluating hydrologic change in such large basins. However, that does not diminish the urgency for information regarding if/how artificial drainage may be affecting hydrology at these large scales. We will consider adding a statement in our results/discussion based on the referee's suggestion.

13. Section 3.6 and 4.4, hydrologic budgets, general. I think at its best, the hydrologic budget analysis can identify if climate is responsible for trends in mean annual stream- flow, not all of the streamflow metrics presented, and I would like to see this clarified. The annual change in soil moisture may be completely different from the short-term changes in watershed storage responsible for peak runoff generation. Overall, I think the uncertainty in the ET, and comparison to Ameriflux has rendered this part of the analysis inconclusive, and this might be a good section to target for removal.

Response: The referee is correct in that the hydrologic budget does not, and was never intended to, account for all the flow metrics considered in this paper. Nor does it explicitly account for soil moisture, groundwater, or anthropogenic withdrawals for consumption or in rare cases irrigation. We simply do not have the data for the periods of record necessary to account for such factors. We have attempted to acknowledge this uncertainty, along with the uncertainty in the ET calculations throughout the paper. We report three main findings in this paper on page 30, lines 20-25, and based

on these findings we suggest that artificial drainage has played a role in the hydrology of large river basins, lines 25-28. As this paper and others have demonstrated, artificial drainage is a ubiquitous practice in much of the Midwestern USA, and is known to have effects at the field and small to intermediate watershed scales. Based on the weight of evidence from multiple lines of information presented in the paper we stand behind our conclusion that agricultural drainage has likely amplified the streamflow response to increased precipitation at large scales. We disagree with the reviewer that the water budget section could be targeted for removal, but will clarify more specifically in the revised manuscript how the hydrologic budget advances our understanding of hydrologic change.

14. Page 26. The hydrologic budget analysis starts with an assumption that ET is stationary. I can't recall that ET trend results were ever reported. I think that the best use of the Livneh dataset is to evaluate any climatological trends in ET, which you have not emphasized here.

Response: We did not assume that ET was stationary in the hydrologic budget. However, we have reason to believe ET should not have changed much and if anything should have increased between the pre- and post- period due to climate and crop changes. Though conversion of perennial grasses and small grains to soybeans can reduce ET, especially early in the growing season (page 3, lines 28-31), we expect that increases in crop productivity and yields have likely increased ET during the study period. Schottler et al. (2014) reported less than 10% changes in ET related to crop conversion and climate for watersheds in southern Minnesota studied over a similar period. If ET goes up, all else constant, the runoff ratio should decrease, which makes our estimates of changes in storage even more conservative. The reviewer is correct that we did not report ET trend results. L13 ET, which represents ET changes due to climate only and does not consider crop or drainage changes, shows relatively small increases in ET 1%-5% between the pre-and post- periods for the study watersheds. We will consider briefly reporting ET trends in the revised manuscript.

15. Page 26, line 27. I think it would be helpful to present the cumulative change in storage calculated for the pre- and post- periods. I am a little unclear over what periods were used here, but I was trying to see if it is even feasible for such changes to result from drainage – a net reduction in storage in the MRB of -2.7 cm/yr, over a 40 year period, so 108 cm of water lost from storage. With drain depths on the order of 3-4 ft, and a porosity of maybe .47 there is only about 57 cm of water in the soil above the drain to be removed, and with only 45% of the watershed drained subsurface drainage could only explain maybe 26 cm of this decrease?

Response: It is useful to think through these calculations at this basic level and we appreciate that the reviewer has taken the time to consider whether or not the numbers add up. One issue with the reviewer's calculation is the assumption that the cumulative storage volume can only be filled and depleted once. In reality, the soil moisture storage can be filled and depleted multiple times each year. The storage is not permanently lost, it is only reduced on an event, seasonal, or annual basis. Draining an additional 2.7 cm/yr does not seem at all unreasonable given that it is about 10% of what could be drained for any given event (~26 cm). In any case, we will consider adding a brief statement on cumulative change in storage in the discussion.

16. Page 27, line 1. I question whether the 5% bias in the Livneh ET relative to the Ameriflux site would systematically affect the time series. I don't know what vegetation was simulated for the Ameriflux site, but looking at the scale of the entire watershed, the static vegetation of the Livneh dataset would capture the dominant corn/soybeah rotation circa 1990, so going back in time it would like underestimate the ET associated with the greater cover of perennial vegetation.

Response: The vegetation type for the Ameriflux sites is reported in Table 2. L13 does not account for drainage changes or crop changes in calculations of ET. Thus differences between L13 and the Ameriflux satiations, which have similar modern vegetation cover, represent uncertainty in the modern simulated ET. This uncertainty may systematically affect the time series. Our rationale in comparing L13 ET to the Ameriflux stations and other ET datasets (reported in the supplement), is to demonstrate that they are at least reasonable modern estimates to use in this study (page 14, lines 7-9). Changes in ET due to crop changes and drainage are not explicitly considered, and thus captured in the storage term. We feel we have further addressed the referee's concerns in the response to Referee #1, Referee Pitlick and in the above replies.

References Cited:

Feng, S. and Hu, Q.: Changes in winter snowfall/precipitation ratio in the contiguous United States, J. Geophys. Res., 112(D15), D15109, doi:10.1029/2007JD008397, 2007.

Graf, W.: Dam nation: A geographic census of American dams and their large-scale hydrologic impacts, Water Resour. Res., 35(4), 1305–1311, 1999.

Groisman, P. Y., Knight, R. W. and Karl, T. R.: Heavy precipitation and high streamflow in the contiguous United States: Trends in the twentieth century, Bull. Am. Meteorol. Soc., 82(2), 219–246, doi:10.1175/1520-0477(2001)082<0219:HPAHSI>2.3.CO;2, 2001.

Higgins, R. W. and Kousky, V. E.: Changes in Observed Daily Precipitation over the United States Between 1950-1979 and 1980-2009, J. Hydrometeorol., 120822104745000, doi:10.1175/JHM-D-12-062.1, 2012.

Lian, Y., You, J.-Y., Sparks, R. and Demissie, M.: Impact of Human Activities to Hydrologic Alterations on the Illinois River, J. Hydrol. Eng., 17(4), 537–546, doi:10.1061/(ASCE)HE, 2012.

Livneh, B., Rosenberg, E. A., Lin, C., Nijssen, B., Mishra, V., Andreadis, K. M., Maurer, E. P. and Lettenmaier, D. P.: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions, J. Clim., 26, 9384–9392, 2013.

Novotny, E. V. and Stefan, H. G.: Stream flow in Minnesota: Indicator of climate change, J. Hydrol., 334(3–4), 319–333, doi:10.1016/j.jhydrol.2006.10.011, 2007.

Schottler, S. P., Ulrich, J., Belmont, P., Moore, R., Lauer, J. W., Engstrom, D. R. and Almendinger, J. E.: Twentieth century agricultural drainage creates more erosive rivers, Hydrol. Process., 28(4), 1951–1961, doi:10.1002/hyp.9738, 2014.

United States Army Corps of Engineers: National Inventory of Dams, [online] Available from: http://nid.usace.army.mil (Accessed 16 January 2017), 2016.