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

This paper reports hydrological analysis from four river basins, in the 20,000 – 70,000 sqr km range in area, in the upper Midwest. The study is designed to investigate the role of climate and land use and land cover (LULC) change on water balance. Streamflow data are used from outlet of relatively large basins. In each basin spatially and temporally averaged PRIMS precip data are used. For annual water balance and storage change discussions evapotranspiration (ET) from Livneh et al (2013) VIC model results are used.

Overall the paper is very long and lacks focus. At least one-third or more of the text can be cut. For example a relatively long paragraph is devoted to sediment and bank erosion and so on (pages 2, 3) which can be cut. Study site descriptions are particularly very long.

Response: While we acknowledge that the manuscript is long, we disagree with the referee that it lacks focus. The referee identifies that the focus of the paper is to investigate the roles of climate and land use and land cover change on streamflows in large Midwestern USA watersheds. All of our introduction, methods, results and discussion are focused on achieving that goal. While we agree with all three reviewers that there are places we can reduce text (perhaps by 1000-1500 words), we do not believe it is feasible to reduce the paper by one-third or more without diminishing the paper. Specific 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. Section 4.4 is currently mislabeled and should be labeled section 4.5 beginning on page 26, line 13. We apologize for this error. We agree that eliminating ancillary information will focus the paper better and ultimately result in a stronger paper.

I did not find the conclusions very strong in the end. While there is some clear changes in streamflow characteristics (e.g., flashiness) as impacted by both LULC and climate change, much of the hydrological impact of LULC, tile drainage, and climate change may be occurring at much smaller scales in this region. If streamflow data from subwatersheds were used the study would have sampled from a greater variability of climate and LULC change. I wonder why the authors studied this local problem at such a large scale. This manuscript could have been a lot stronger if a range of watershed sizes were studied and the role of spatial averaging were shown.

Response: We are aware that other hydrologic changes are occurring at small scales. Several recent studies have thoroughly documented those changes and we reference those studies in the manuscript (e.g. Danesh-Yazdi et al., 2016; Foufoula-Georgiou et al., 2015; Frans et al., 2013; Gerbert and Krug, 1996; Juckem et al., 2008; Novotny and Stefan, 2007; Schilling and Libra, 2003; Schottler et al., 2014; Zhang and Schilling, 2006). However, should we expect that the same effects we observe at the field or subwatershed scale should apply at significantly larger scales? Currently there is a paucity of studies conducted at this large scale. As we state on p. 2, lines 24-26 our goal here is to determine if the effects of tile drainage are evident even at these large scales. Often state and federal agencies and special task forces must make recommendations and/or manage watersheds of this size or larger (e.g. USEPA Mississippi River/Gulf of Mexico Hypoxia Task Force). Therefore we believe the scale of investigation is justified. We have included some temporal multi-scale analyses (daily to annual scales) in this paper, but believe that a spatial multi-scale study (field to large watershed) goes beyond the scope of our paper and could form an excellent next step in this line of inquiry.

I have some concerns using the Livneh et al (2013) ET product. In the application of the annual water balance equation, the ET product should come from a model that uses the influences of the changes in LULC and title drainage on water balance and ET calculations (e.g., Frans et al., 2013). If Livneh data is not driven by annual changes in LULC and tile drainage their ET product would not be suitable for making these inferences on LULC and tile drainage impact on changes in storage of the watershed. Because given P and Q are from data, the storage change calculations reported in Table 4 is directly result of ET estimates.

The paper also repeatedly mentions the fact that Livneh et al over estimates ET. This should be more clearly discussed. Where was this over estimation, under what LULC, and why a single number was used to correct ET estimates in all the river basins used in this work?

I don't think reducing 17% of Livneh et al ET for a more conservative budget makes a lot of sense unless justified at select locations from Ameriflux data or over available data in the watersheds. Alternative approach would be running VIC with the detailed LULC change data used in this study..? If the tile drainage influence on water balance is significant not parameterizing it in VIC may lead to higher ET.

Response: We agree with the referee that there are potential issues of using the L13 ET_a data even though the VIC model does not consider human (LULC) influence. By comparing these data to the AmeriFlux towers, two of which are influenced by corn-soy agriculture, we demonstrated that they are at least reasonable modern estimates to use in this study (page 14, lines 7-9). We are not aware of other modeled ET data that have the same spatial and temporal extent as the L13 data. Further, we are not aware of ET datasets that explicitly include the influence of tile drainage over the relevant space and time scales. For example, Frans et al., 2013 simulate tile drainage expansion over time by increasing the fraction of soil water that contributes to baseflow. This is done using a nonlinear baseflow release function in VIC, calibrated each decade from 1920 to 1980 for the 9389 km^2 Raccoon River watershed and then applied to the entire 443,000 km^2 Upper Mississippi River basin (UMRB). Such an approach relies on two implicit and unverifiable assumptions: 1) that agricultural drainage in the UMRB has increased in time in accordance with Raccoon River, and 2) that drainage increases the fraction of soil water that goes as baseflow. Rather than implicitly relying on these assumptions, we are simply trying to determine whether drainage effects are detectable at large spatial scales. If we observe a systematic decrease in the storage term over time, especially using the conservative assumptions that we have applied, then agricultural drainage is the likely cause for this shift since drainage effects are not included in estimates of ET_a from L13. Although we cannot fully rule out groundwater and/or dam operation effects on modern streamflows, there is no evidence to suggest change in regional groundwater recharge. Furthermore in basins with dams, such as the IRB we have observed dampened increases in streamflow despite widespread agriculture and similar increases in precipitation. Therefore we are confident with our conclusion that drainage is most likely amplifying streamflows in agricultural basins, especially those without major dams or urban centers. This point is further illustrated by the fact that in some seasons precipitation changes are minimal while streamflow changes are substantial.

There is a rational and conservative basis for applying the 17% reduction in JJA (page 13, lines 27-31; page 14, lines 1-10). Additionally, it is during summer months that ET is most likely limited by soil water availability. However as pointed out by another referee, this adjustment does not change the interpretations of the water budget. In response to all comments we plan to include a brief discussion of assumptions and limitations and point to the fact that these do not alter the analysis and main results of our study.

Table 4.. I'm skeptical about these water balance values. Several issues needs to be justified:

1. the changes are reported as cm/year.. These don't seem to be right and I suspect the authors mean cm per century. I see up to 90 cm/year values as precip change..

2. most increase in precip is from june to september in the region. this would increase ET compared to post 1970s. wouldn't reducing ET 17% take away this additional ET from wetter conditions.

3. If drainage and LULC change was not represented in the ET model, how can you conclude that storage changes may be attributed to artificial drainage.. Of course one could make all sorts of arguments but at the end of the day these would remain as speculations. In addition, the storage change values are so small that given the uncertainty in P and Q measurements and interpolation errors in space and time as well as the uncertainties in ET estimates would easily lead to 10% or more uncertainty in the storage change estimatesâˇA ˇTe.g., Dingman second edition page 15, estimation of ET uncertainty. I would either remove this part or talk more about the role of uncertainties in the estimation of storage change.. Plus at the annual and mean annul scales changes that occur in space and time could cancel out each other and reporting these numbers at the scale of these large river basins may not mean much for management purposes.. Regional modeling papers that varied LULC and drainage conditions over space and time have discussed the effects of averaging in water balance in this region.

Response: (1)There is a misunderstanding about Table 4 due to an error in the table caption. We apologize for this error. The word "change" should be removed from the caption as the table represents the observed average annual precipitation (P), flow (Q),…
(2) The hydrologic effects of increased precipitation in the fall are considered in estimates of ET post 1970's.
(3) The fact that drainage is not included in the ET model is what allows us to test, external to the ET predictions, whether or not a drainage 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. We feel we have given sufficient attention to uncertainty by comparing multiple precipitation and ET datasets in the manuscript and in much greater detail in the supplement. We acknowledge that given the uncertainty in all estimates of ET, the relative difference (or direction of change) between the pre- and post-period storage term might be of more value than magnitude of the storage terms individually (page 28, lines 14-20). Furthermore, we stand behind our conclusions, which are based on inductive versus deductive (e.g. Frans et al. 2013) reasoning. Please provide citations to the regional modeling papers referenced. We will gladly review them, but as we argued above, we do not believe there exists an ET dataset that accounts for artificial drainage in a meaningful way at the space and time scales that are relevant for our study. This point will be better emphasized in our revised manuscript.

Conclusion reads as though the paper was largely written to study the effects of tile drainage.

Overall, there is a tone of a relatively large term paper. Sections are not well connected, and the narrative does not flow well, interrupted with details of study sites and text and tangential topics. There is awkward wording throughout the paper eg.,– "trends observed are relatively correct" what does this mean are there also relatively incorrect trends. I suppose this should do something about the statistical significance of trends.

Response 5: The comment made by the referee, “trends observed are relatively correct” is taken out of context.

References:

Danesh-Yazdi, M., Foufoula-Georgiou, E., Karwan, D. L. and Botter, G.: Inferring changes in water cycle dynamics of intensively managed landscapes via the theory of time-variant travel time distributions, Water Resour. Res., 52(10), 7593–7614, doi:10.1002/2016WR019091, 2016.
Foufoula-Georgiou, E., Takbiri, Z., Czuba, J. A. and Schwenk, J.: The change of nature and the nature of change in agricultural landscapes: Hydrologic regime shifts modulate ecological transitions, Water Resour. Res., 51(8), 6649–6671, doi:10.1002/2015WR017637, 2015.
Frans, C., Istanbulluoglu, E., Mishra, V., Munoz-Arriola, F. and Lettenmaier, D. P.: Are climatic or land cover changes the dominant cause of runoff trends in the Upper Mississippi River Basin?, Geophys. Res. Lett., 40(6), 1104–1110, doi:10.1002/grl.50262, 2013.
Gerbert, W. A. and Krug, W. R.: Streamflow trends in Wisconsin’s Driftless Area, J. o fhte Am. Water Resour. Assoc., 32(4), 733–744, doi:10.1111/j.1752-1688.1996.tb03470.x, 1996.
Juckem, P. F., Hunt, R. J., Anderson, M. P. and Robertson, D. M.: Effects of climate and land management change on streamflow in the driftless area of Wisconsin, J. Hydrol., 355(1–4), 123–130, doi:10.1016/j.jhydrol.2008.03.010, 2008.
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.
Schilling, K. E. and Libra, R. D.: Increased baseflow in Iowa over the second half of the 20th century, J. Am. Water Resour. Assoc., 39(4), 851–860, doi:10.1111/j.1752-1688.2003.tb04410.x, 2003.
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.
Zhang, Y. K. and Schilling, K. E.: Increasing streamflow and baseflow in Mississippi River since the 1940 s: Effect of land use change, J. Hydrol., 324(1–4), 412–422, doi:10.1016/j.jhydrol.2005.09.033, 2006.