Dear reviewer,

Many thanks for your comments on our manuscript. Based on your very constructive comments, we have thoroughly revised the manuscript. Additional discussion and justifications have been added into the manuscript or into the Supplement. Please see below the detailed responses. Major changes have also been highlighted in the revised manuscript.

With best regards
Lishan Ran, on behalf of the coauthors

General comments:
Ran et al. reported new data on riverine carbon export in the arid-semiarid Wuding River watershed on the Chinese Loess Plateau. Considering that river systems in the East Asia, especially those in the arid-semiarid climates are under-represented in the global budget of riverine carbon fluxes, this study could provide valuable datasets. However, the paper can be improved further by explaining in detail how the errors were calculated in load estimates and CO2 evasion, offering detailed explanation in the methods (e.g. the river surface area), and providing discussion on the observed patterns with statistical significance testing results. Specific comments are below, which the authors may consider when revising the manuscript.
Reply: Thanks a lot for your constructive comments. Please find below our responses to your comments.

Specific comments:
Lines 48-: "substantial" is a relative term. Please provide a value or a range just like 1.8 Pg C year-1 in the previous sentence.
Reply: The annual OC burial due to sediment storage in global reservoirs and lakes ranges from 0.15 to 0.6 Pg C year^{-1}. This range has been added into the revised manuscript. (lines 47-48)

Lines 84-: "multi-annual" is an unspecific term. Please provide more information on how the mean of water discharge was calculated. For example, you can provide the period (e.g. 1980-2010?) Also, year-1 as a time unit would be appropriate for annual discharge. Is it 11.2 * 10ˆ8 m3 yr-1 (Ran et al., 2017)?
Reply: The calculated mean water discharge of 35 m^3 s^{-1} is based on the period 1956–2017. The mean annual water discharge during this period is 11.2×10^8 m^3 year^{-1}. These changes have been added into the text. (lines 84-85)

Lines 149-: Isn't the 14C half-life 5,730 years?
Reply: The half-life used in carbon dating calculations by the Beta Analytic Inc. is 5568 years, the value worked out by chemist Willard Libby, and not the more accurate value of 5730 years, which is known as the Cambridge half-life. Although it is less accurate, the Libby half-life was retained to avoid inconsistencies or errors when comparing carbon-14 test results that were produced before and after the Cambridge half-life was derived. Detailed description on the 14C half-life can be found at the Beta Analytic website: https://www.radiocarbon.com/PDF/AMS-Methodology.pdf.

Lines 156-158: Detailed explanation is needed on the validity of the methods on how the riverine carbon exports were calculated considering that the major findings of this paper are the new estimates of the riverine carbon loads. Detailed explanation is needed on how river flow was measured. The method of load estimation appears to be too simple and with many assumptions, not specifying errors associated with each step. There are several methods for load estimation (e.g. Sickman, J.O. et al., 2007, Water Resources Research, Effects of urbanization on ...) you may try these and compare the results because load calculation is crucial to draw conclusions. One way to calculate daily load of stream ions is to use the LOADEST software developed by USGS if daily water discharge data are available. The software also provides confidence intervals.

Reply: Many thanks for your comment. Estimating riverine carbon flux is a very important part of this study in which we attempt to investigate the fate of carbon after entering the drainage network from terrestrial ecosystems. Just as you have pointed out, there are a number of methods to estimate the annual fluxes of dissolved and particulate matter transported by rivers. Major methods currently used include linear interpolation and ratio estimators, regression-based methods historically employed by the USGS, and recent flexible techniques such as Weighted Regressions on Time, Discharge, and Season (WRTDS), etc. As you have also suggested, the most commonly used USGS software package for estimating constituent load using regression is known as LOADEST (Runkel et al., 2004. Load Estimator (LOADEST): A FORTRAN Program for Estimating Constituent Loads in Streams and Rivers. U. S. Geological Survey Techniques and Methods Book 4, Chapter A5). Lee et al. (2016) recently reviewed the potential for flux estimation bias across a broader range of estimation methods and concluded that the Beale's ratio estimator and WRTDS generally exhibit greater estimation accuracy and lower bias (Lee et al., 2016. An evaluation of methods for estimating decadal stream loads. *Journal of Hydrology*, 542, 185-203). Our annual carbon flux estimation in this study was based on the Beale's stratified ratio estimator. Since the riverine carbon concentrations were measured with "sparse" sampling frequency while flow and suspended sediment had a continuous daily measurement, this method could greatly reduce the bias introduced by relatively low sampling frequency, in particular the high flow events that are often under-sampled (Parks and Baker. 1997. Sources and transport of organic carbon in an Arizona river-reservoir system. *Water Research*, 31, 1751-1759). Indeed, we have already used the Beale's ratio estimator in our earlier estimation of carbon flux in the Yellow River with success (i.e., Ran et al., 2013. Spatial and seasonal variability of organic carbon transport in the Yellow River, China. *Journal of Hydrology*, 498, 76-88). And the Beale's ratio estimator has proven to be highly reliable and is recommended if the relationship between discharge and concentration is weak (e.g., Fulweiler and Nixon, 2005. *Biogeochemistry*, 74, 115-130; Awad et al., 2017. *Environmental Pollution*, 220, 788–796; Chen et al., 2014. *Journal of Geophysical Research: Biogeosciences*, 119, 95-109; Sun et al., 2017. *Hydrological Processes*, 31, 2062-2075). In comparison, we have also estimated the carbon flux by using the suggested LOADEST software package. The flux results show high consistency with each other, with a difference of less than 4.5%. We have added a detailed description of the estimate method (i.e., the Beale's ratio estimator) in the revised manuscript. Please refer to the highlighted changes in the text. (lines 161-180)

Lines 160-: Do you mean the POC concentration not "content"? It appears the term "content" is misused throughout the manuscript.

Reply: Based on your and other reviewers' comments, we have re-defined the concentration of POC throughout the manuscript. It should be the POC concentration (POC%) in the total suspended solids (dry weight). (lines 161-163)

Lines 170-: How large is the river width? If it is near or lower than 90 meters, how can you estimate river surface areas using the DEM data of 90-m resolution? In other words, aren't you using too coarse data to estimate river water surface areas?
More detailed explanation is required on how the water surface area is calculated since this is a critical term for CO2 evasion estimates.
Reply: Because the Wuding River catchment is located in an arid-semiarid climate zone, the rivers and streams of drainage network is generally narrower than their counterparts in tropical rivers due to lower water discharge. The widths of the rivers and streams vary from 1.8 (first order streams) to ~61 m (the mainstem channel) (see Table 1 in our earlier wor: Ran et al., 2017. *JGR-Biogeosciences*, 122, 1439-1455), significantly lower than the DEM resolution of 90 m. Therefore, we only used the DEM data to delineate the drainage network in terms of stream length (usually >2.5 km) and stream number. The delineated drainage network was also calibrated through ground truthing during our fieldwork. Because the width of all rivers is less than the resolution and it fluctuates between dry and wet seasons, we measured widths of all sampled rivers during our fieldwork and aggregated them based on stream order to calculate the water surface area. We have revised the description of the water surface area calculation in the revised manuscript. (lines 198-203)

Lines 183-198: The method is better than nothing for sure. However, it appears the used references are relatively old (1995 and 2000). Do you have newer references on heterotrophic respiration than those? How the errors associated with the approach are calculated?
Reply: To estimate the Wuding River catchment's net ecosystem production (NEP), we used the global soil CO_2 efflux database described by Raich and Potter (1995) and the heterotrophic soil respiration (Rh) estimated by Hanson et al. (2000). Based on the global soil respiration flux database (Raich and Potter, 1995), the S_R for this catchment is the range of 400-500 g C m^{-2} year^{-1}. Hence, we used 450±50 g C m^{-2} year^{-1} to represent its soil respiration. This rate is consistent with recent measurements under different vegetation types in this arid-semiarid region (e.g., Fu et al., 2013). Fu et al. (2013. Soil respiration as affected by vegetation types in a semiarid region of China. *Soil Science and Plant Nutrition*, 59, 715-726) measured total soil respiration in this arid-semiarid region. Their mean soil respiration rates under 4 different vegetation types are in the range of 1-1.4 µmol m^{-2} s^{-1}, which are equivalent to 380-530 g C m^{-2} year^{-1}. Thus, our estimate is reliable. Although the references are relatively old, using the ratios derived from Hanson et al. (2000) has been widely used to assess heterotrophic soil respiration in river catchments under different land cover types (e.g., Brunet et al., 2009. Terrestrial and fluvial carbon fluxes in a tropical watershed: Nyong basin, Cameroon. *Chemical Geology*, 3, 563-572; Lee et al., 2017. A high-resolution carbon balance in a small temperate catchment: Insights from the Schwabach River, Germany. *Applied Geochemistry*, 85, 86-96). In addition, the propagated errors are calculated and presented in the revised manuscript. We have also added new references to justify our arguments. (lines 316-326)

Lines 208-: Does the 'sediment' mean 'suspended sediment'? If so, please clarify it to prevent confusion. Do you mean the POC concentration not "content"?

Reply: Yes, here it means the POC concentration in suspended sediment. We have clarified this in the revised manuscript 'The POC% in suspended solids…'. (lines 161-163)

Lines 222- and throughout the manuscript: What is the "+/-"? Standard deviation? Or standard error?
Reply: The '±' denotes standard deviation (SD) throughout the manuscript. We have explicitly indicated this when it is used for the first time in the revised manuscript (i.e., in Figure 2). Many thanks.

Lines 225-: While [DOC] (3.3 mg/L) is larger than [POC] (0.61 mg/L), the DOC export is much lower (0.3*10^10 g (yr-1?)) than POC export (3.7*10^10 g (yr-1?)). Why is that so?
Reply: Here the DOC concentration (mg/L) is expressed as the DOC content per unit volume of water, and the POC is expressed as POC% in total suspended solids (TSS, dry weight). Although the DOC concentration is larger than the POC%, the annual water discharge (7.71×10^8 m^3/yr) at the catchment outlet Baijiachuan gauge is relatively low due to low precipitation and the concomitant annual TSS flux (610×10^{10} g/yr) is quite high owing to severe soil erosion. As a result, the annual DOC flux (g C/yr) is much lower than the POC flux.

Lines 228-233: The river water discharge and carbon loads can be highly dependent on precipitation. Was the year of field campaign categorized as wet, dry, or normal year compared to the long term mean (e.g. 1980-2017 precipitation)?
Reply: The multiannual precipitation for the Wuding River catchment is in the range of 300-500 mm during the period 1956-2010 with a mean precipitation of 430 mm/yr (available at http://www.yellowriver.gov.cn/; Li et al., 2007. *Hydrological Processes*, 21, 3485-3491). The precipitation in 2015 is about 410 mm, indicative of a normal year relative to the long-term mean precipitation. In comparison, the precipitation in 2017 is larger than 540 mm, significantly higher than the long-term mean precipitation (i.e., 26% higher). That is why we used the 2015 hydrological data to calculate the carbon flux. Another reason is because the three seasonal samplings were also performed in 2015. We have revised the hydrological information in the manuscript. (lines 256-261)

Lines 258–261: As the authors mentioned, the precipitation is high during summer. Thus, this assumption of no significant seasonal fluctuations may not be valid. Can you provide a range of stream surface area and CO2 evasion depending on season?
Reply: For CO_2 evasion from river waters, we separately estimated the total water surface of rivers in spring, summer, and autumn (please refer to Table S4 in Supplement for the estimated water surface area in the three seasons) and calculated the CO_2 evasion in these three seasons. The annual total CO_2 evasion was obtained by summing up the three seasonal CO_2 estimates. But for CO_2 evasion from reservoir waters, because these check dam-formed reservoirs are mostly constructed in steep gully channels and operated primarily for the purpose of sediment trapping and water storage, variation of the water surface area is much less significant than that of the rivers. Although there are also seasonal fluctuations, the magnitude should be quite minor compared with rivers. Thus, we assumed that there was no significant seasonal variation. (lines 289-293)

Lines 300–: Is the decreasing trend of DOC (Fig. 2) statistically significant? It appears the error bars are large. If this is not statistically significant, the following argument is vague. The decrease of DOC can be microbial- or photo-degradation to CO2, sorption to particulate matter, and dilution from increased water discharge of low [DOC]. The following discussion is speculative and could be strengthened by checking each factor.

Reply: Based on your comment, we have performed the significance test for DOC concentrations along the stream order. Because of the large error bars as shown in the figure, the decreasing trend of DOC is not statistically significant at the 95% confidence level. To reflect the downstream DOC concentration change, we aggregated the 1st-5th streams into 2 groups, including the headwater 1st-2nd streams and the higher order 3rd-5th streams, because it is usually believed that headwater low-order streams process organic carbon more rapidly and emit CO_2 at faster rates than downstream high-order streams (e.g., Butman and Raymond, *Nature Geoscience*, 4, 839-842; Crawford et al. 2013. *Journal of Geophysical Research: Biogeosciences*, 118, 482-494). Our results indicated that the DOC concentrations in the headwater 1st-2nd streams were on average 16-39% higher than that in the downstream 3rd-5th streams. Thus, the downstream DOC decline in the 1st-5th order streams likely suggests the mineralization of the bioavailable fraction of DOC along the river course (Figure 2), especially in spring and autumn. In addition, sorption and input of increased water with low DOC may partially dilute the DOC concentration as you commented. However, in view of the spatial homogeneity in terms of soil erosion rate, SOC content in soils, and hydrologic regime within each subcatchment and the spatially constant POC% from the headwater to the mainstem channel, the sorption and 'dilution effect' are expected to be minimal. Accordingly, we have revised the claims in the manuscript. (lines 230-245; 333-349)

Line 325–336 (and lines 385–394, and Fig. 6): I am confused. Do you mean the POC concentration not "content"? Why the "content" has the unit of concentration, %, not just grams? I think heavy rain during summer could generate high POC content but low POC concentration. Please clarify.

Reply: Many thanks for your comment. It should have been POC concentration in the text. We have clarified the POC concentration in the total suspended solids (TSS, dry weight) in the units of POC% throughout the manuscript. By multiplying the annual TSS flux, we can calculate the annual POC flux.

Line 351–368: The pCO2 is a function of pH and alkalinity. The pCO2 is high when the water pH is low. The ground water of the area has the pH of >~8. Then, the calculated pCO2 is very low which is well described in the line 211. Then, how CO2 evasion can be high when pCO2 is low? Please clarify.

Reply: Just as you have pointed out, pCO_2 is a function of pH and alkalinity, and it can be calculated from the latter two variables. The observed pH in the study catchment ranged from 7.68 to 9.29 and the pH in groundwater is generally slightly higher than 8.0. Even so, for the sandy subcatchment reservoirs, the pH of the groundwater is still lower than that of the river water into the reservoirs (e.g., 8.7-9.3). With extremely high alkalinity (DIC) concentrations, despite the relatively high pH of around 8.0, the calculated pCO_2 is well above the atmospheric equilibrium (i.e., ~390 µatm), and facilitates the observed CO_2 evasion. We have revised the manuscript to make the claim more clear and accurate. (lines 398-405)

Lines 400: Which part of the Figure S1 supports this sentence?
Reply: We have added the information on carbon export during typical floods in the Supplement (Figure S1).

Lines 430-481: Very interesting findings.
Reply: Many thanks for your comments. We collected carbon isotope samples of the emitted CO_2 from river waters and attempted to explore its potential sources in association with carbonate dissolution and respiration of recent organic matter.

Tables: What is the "+/-"? Standard deviation? Or standard error?
Reply: The '±' denotes standard deviation (SD) and the description has been added in the revised manuscript.

Table 1: Please provide information on how many reservoirs were used to draw the table.
Reply: There are currently 337 reservoirs in operation within the Wuding River catchment (please see the figure below). This information has been added into the caption, and this map has also been included in the Supplement (Figure S2).
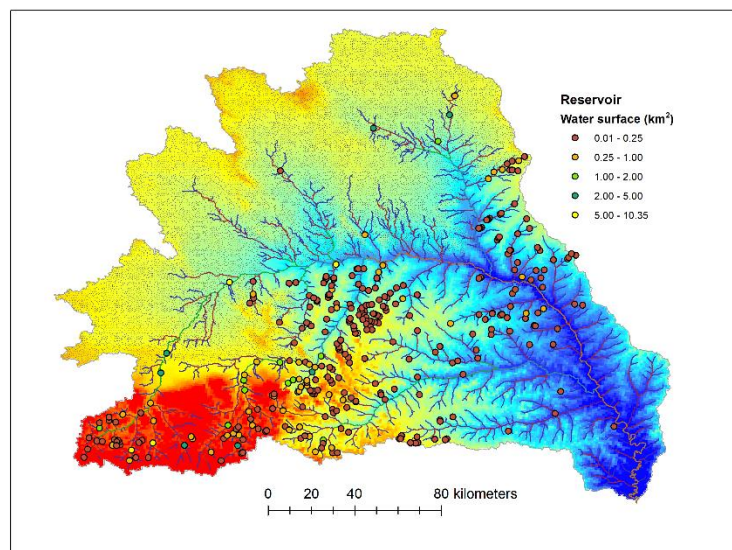

Figure: Spatial location of the 337 reservoirs within the Wuding River catchment.

Figure captions need to provide more detailed description of the figures including explanation on legends.
Reply: We have significantly improved the figure captions based on your comments and detailed information has been added. Please refer to highlighted changes in the revised manuscript.

Figure 1: It is hard to differentiate the colors of the stream order, especially with the background altitude colors. Please revise the figure so that each symbol can be seen clearly.
Reply: We have carefully adjusted this figure in terms of color scheme, marker size, label size, etc., and have added the subcatchment boundaries to make the figure much easier to read.