We have carefully addressed all the comments made by the two anonymous reviewers on our manuscript (EGUSPHERE-2026-1058) entitled "Impacts of Cascading Check Dams on Sediment Yield in the Middle Yellow River Basin: Insights from 50 Years of Grid-cell-level Simulation". The comments have helped us greatly improve the overall quality of the manuscript. The following is the point-point response to all the comments.

Response to Anonymous Referee #1:

1. General comments:

The manuscript "Impacts of Cascading Check Dams on Sediment Yield in the Middle Yellow River Basin: Insights from 50 Years of Grid-cell-level Simulation" presents an interesting attempt to integrate the RUSLE-IC-SDR framework with a dynamic sediment trapping efficiency (TE) module to evaluate the spatiotemporal impacts of cascading check dam networks on sediment yield. The topic is highly relevant to current challenges in large-scale soil erosion modeling and sustainable watershed management, particularly in the Middle Yellow River Basin. The authors' effort to quantify the sediment reduction contributions and storage capacity dynamics of over 47,000 check dams using a novel topologic routing algorithm provides practical insights for optimizing soil and water conservation strategies and regional spatial planning. However, the manuscript in its current form suffers from several methodological uncertainties—particularly regarding spatial data resampling, parameter sensitivity, and the clarity of certain visual representations—that must be addressed.

Reply: *We appreciate the reviewer's positive assessment of the topic and methodology. We will address the methodological uncertainties and presentation issues in the revised manuscript as detailed below.*

Additional line comments:

2. Comment:

Lines 24-27: The abstract mentions the results of the correlation analysis (such as the logarithmic growth relationship and $P < 0.001$), but does not mention specific R2 or correlation coefficients. I recommend briefly adding key statistics in the abstract to enhance the quantitative persuasiveness of the findings.

Reply: *We will revise the Abstract by adding the key statistical values for the correlation analysis. Specifically, we will report that* SRC_{dam} *was strongly related to check dam density and the proportion of dam-controlled area (*$R^2 = 0.81$ *and 0.78, respectively; both* $P < 0.001$*), whereas* SAR_{dam} *increased logarithmically with upstream sediment yield without dam trapping (*$R^2 = 0.62$, $P < 0.001$*).*

3. Comment:

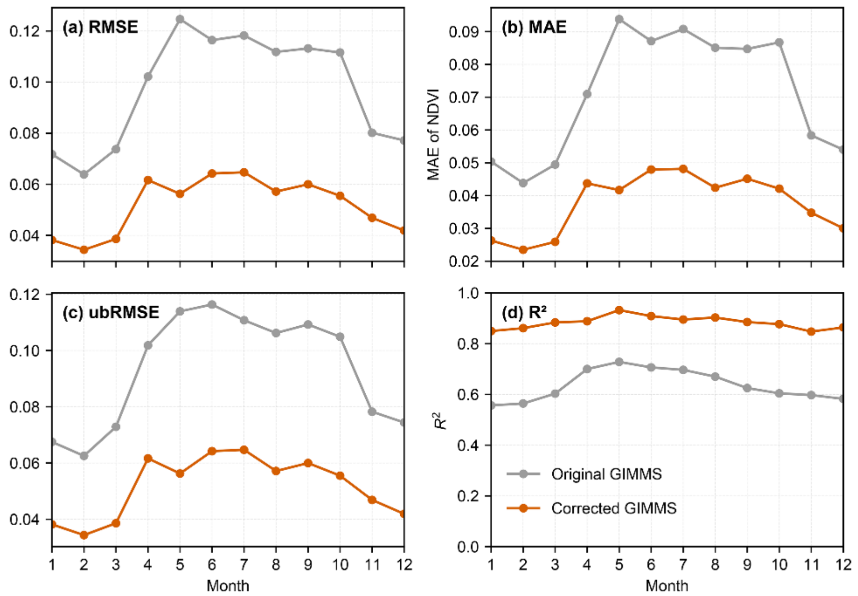
Lines 60-65: The introduction summarizes the shortcomings of existing research very well. It is recommended to further emphasize the specific advantages of the topological sorting algorithm proposed in this study in terms of "computational efficiency" and "avoiding double-counting" compared to traditional hydrological response unit or highly parameterized models.

Reply: *This is a good comment. We will revise the final paragraph of the Introduction to better highlight the advantages of the proposed topological sorting strategy. The revised text will clarify that this strategy establishes the upstream-to-downstream routing order of individual check dams from grid-based flow direction, flow accumulation, and dam-point locations. This improves computational efficiency, avoids double-counting in overlapping dam-controlled areas, and preserves grid-cell-level spatial patterns of erosion, sediment delivery, and dam-induced sediment reduction.*

4. Comment:

Lines 126-128: The authors spliced and downscaled two NDVI datasets with significantly different spatial resolutions (8 km for AVHRR and 250 m for MODIS). This is an important source of uncertainty in the model. Please briefly add the validation accuracy of the downscaling technique (such as RMSE) to prove the reliability of the C factor calculation between 1970 and 2000.

Reply: *We will revise the dataset description by adding the validation accuracy of the downscaled GIMMS NDVI against MOD13Q1 NDVI during their overlapping period. The revised text will report that the downscaled GIMMS NDVI showed good agreement with MOD13Q1 NDVI ($R^2 = 0.94$, RMSE = 0.053, and MAE = 0.038). Compared with the original GIMMS NDVI, RMSE and MAE decreased by 46.9% and 47.2%, respectively. We will also add a supplementary figure to show the monthly error metrics and spatial RMSE patterns before and after correction. These results support the use of the downscaled GIMMS NDVI for deriving the C factor during the pre-MODIS period.*



(e) Original GIMMS RMSE

(f) Corrected GIMMS RMSE

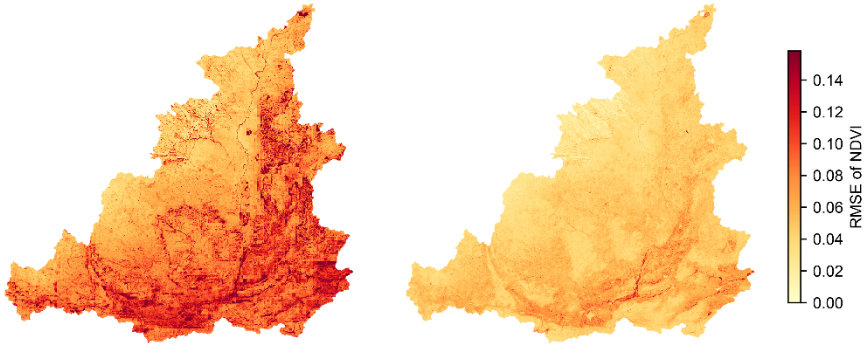


Figure S1. Accuracy evaluation of the corrected GIMMS NDVI using MOD13Q1 NDVI as the reference during the overlapping period. Panels (a–d) show the monthly variations in four accuracy metrics between GIMMS NDVI and MOD13Q1 NDVI: root mean square error (*RMSE*), mean absolute error (*MAE*), unbiased root mean square error (*ubRMSE*), and coefficient of determination (R^2). Panels (e–f) show the spatial distribution of pixel-wise *RMSE* for the original and corrected GIMMS NDVI, respectively.

5. Comment:

Lines 165-171: Spatial data was resampled to a 100 m grid. Since topographic factors (LS factor) are highly sensitive to DEM resolution, resampling from 30 m to 100 m may flatten the slope, thereby underestimating erosion. Please explain the trade-off for choosing the 100 m resolution (e.g., computational efficiency) and discuss its potential impact on the final erosion estimation in the discussion section.

Reply: *We agree that DEM resolution is an important source of uncertainty in LS-factor and sediment-yield estimation. We will clarify in the Methods section that the 100 m resolution was selected as a basin-scale compromise between preserving broad spatial heterogeneity, harmonizing multi-source datasets with different native resolutions, and maintaining* *computational feasibility for 50-year grid-cell-based simulations across the entire MYRB with more than 47,000 check dams.*

We also add a discussion explaining that coarser DEMs may smooth steep gully terrain and small channels, reduce local slope gradients, and alter flow-path continuity and slope-length estimates, thereby affecting the LS factor. The effect on sediment yield is not necessarily one-directional because changes in slope steepness, slope length, sediment delivery, and deposition may interact.

We further note that excessively fine DEMs may introduce data noise and substantially increase storage and computational costs. Although the 100 m resolution may limit the representation of fine-scale erosion and deposition patterns in small gullies and dam-controlled areas, the simulated sediment yield was calibrated and validated using observed records from 17 hydrological stations, which partly constrains the integrated effects of DEM-related uncertainty at the basin scale. Future work will be suggested to conduct multi-resolution sensitivity analyses for LS, IC, SDR, sediment trapping, and sediment-yield patterns in representative sub-basins.

6. Comment:

Lines 192-196: For the fine-grained sediment in the Loess region, the authors used an empirical coefficient of D = 0.046. Although literature is cited, it is recommended to discuss whether this parameter is time-variant under extreme rainstorm events on the Loess Plateau, and whether using a fixed parameter might overestimate or underestimate the trapping rate in specific years.

Reply: *We will clarify the physical meaning and selection basis of the empirical coefficient D, add a sensitivity analysis using different D values, and discuss the potential uncertainty caused by using a fixed D under extreme rainstorm conditions. We will also test D values of 0.08, 0.1, 0.2, 0.5, 0.8, and 1.0. The sensitivity analysis will be included in the Supplementary Materials.*

In the Discussion, we will explain that extreme rainstorms may alter sediment sources and particle-size composition. During extreme rainstorm events on the Loess Plateau, enhanced soil erosion may increase the proportion of coarse particles delivered to check dams, which would increase the effective D value and potentially make the use of a fixed fine-sediment value (D = 0.046) underestimate trapping efficiency in some years. Extreme events may also cause rapid sediment deposition, sharply reduce the remaining storage capacity, and lead to a faster decline in TE in subsequent years. In addition, intense rainstorms and flood pulses may increase the risk of check-dam overtopping or failure (Bai et al., 2020), which could significantly compromise flood control and sediment retention functions. Future work should incorporate event-based sediment particle-size observations and hydrodynamic information to develop a time-varying TE parameterization.

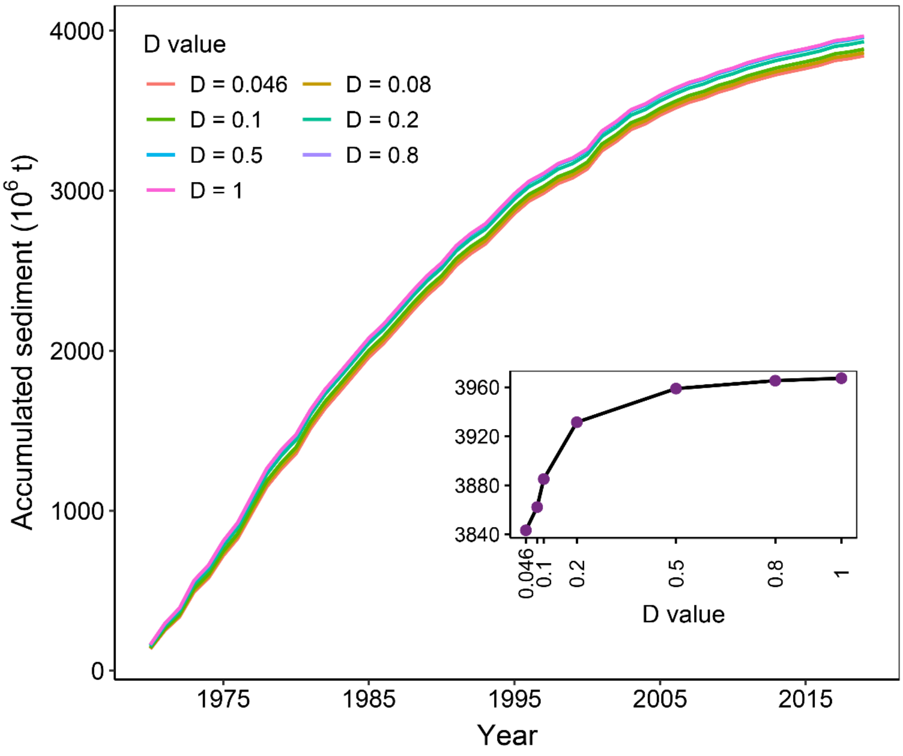


Figure S12. Sensitivity of simulated sediment accumulation in check dams to different values of the empirical coefficient D. The main panel shows cumulative sediment accumulation from 1970 to 2020 under different D values, and the inset compares the final cumulative sediment in 2020. D represents sediment texture and settling behavior in the trapping-efficiency equation; D = 0.046 was used as the baseline value for fine-grained loess sediment.

7. Comment:

Lines 358-362: The discussion mentions the contribution of the "Grain for Green" program to vegetation coverage and erosion reduction. However, in the methodology section (Lines 163-164), the P factor (soil conservation practice factor) is assigned empirical values based on land use types. Please clarify whether the P factor in the 5 time slices from 1970 to 2020 can fully capture the dynamic changes brought by the Grain for Green program.

Reply: *We acknowledge that the P factor based on five land-use time slices cannot fully capture all annual changes associated with the Grain for Green Program. Ideally, annual maps of land use, terraces, and conservation practices would be needed to represent these changes more explicitly. However, such long-term annual datasets are not currently available for the entire MYRB over 1970–2020. Therefore, in our framework, the dynamic vegetation-cover changes induced by the Grain for Green Program were mainly represented by the NDVI-derived C factor, whereas the P factor was used to capture broader decadal-scale changes in land use, terraces, and slope-related conservation practices. We will clarify this treatment in the Methods section and acknowledge the related uncertainty in the Discussion, including the need for annual conservation-practice datasets in future studies.*

8. Comment:

Table 1: The table lists the spatial resolutions of the original data (e.g., 250 m for SoilGrids, 30 m for DEM). I suggest adding a column in the table or clearly stating in the table footnote the “uniform resolution used for model calculation (100 m)” to improve the transparency of the methodology.

Reply: *We will revise Table 1 by adding a footnote indicating that all spatial datasets were resampled to a uniform 100 m grid for model calculation.*

9. Comment:

Figure 1: The numeric labels (1-17) for each sub-basin do not have enough contrast against some background colors (e.g., Sub-basins 14 and 15). I recommend bolding the numbers or adding a white halo effect to improve readability.

Reply: *We will revise Figure 1 by enhancing the contrast of the numeric labels (1-17), using bold numbers with a white halo effect to ensure clearer visibility against the background colors.*

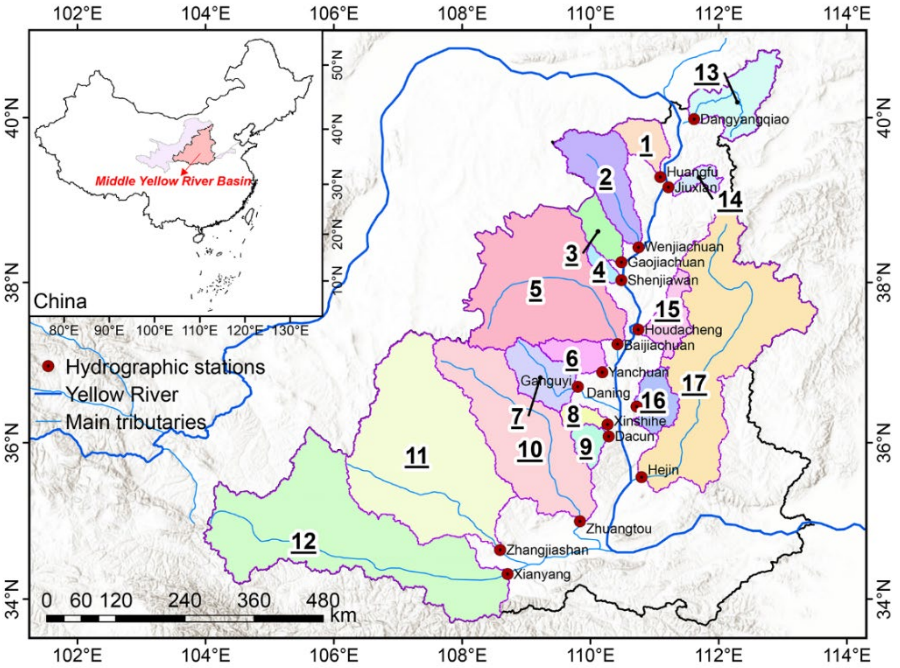


Figure 1. Distribution map of sub-basins and hydrological stations in the middle Yellow River Basin, with pink lines showing the sub-basins boundaries. The sub-basin IDs in the figure correspond to those listed in Table S1.

10. Comment:

Figure 2: Typo correction: A close inspection reveals a spelling error in the red text description of “Step 2”. “RUSLR-IC-SDR” should be corrected to “RUSLE-IC-SDR”. Additionally, please briefly label the data transfer format next to the flow arrows in the figure to allow readers to intuitively understand the computational logic of the grid processing.

Reply: *We will correct the spelling error in Figure 2 by changing “RUSLR-IC-SDR” to “RUSLE-IC-SDR”. In addition, we will add brief labels next to the flow arrows to indicate the transferred data formats, such as .tif and .shp.*

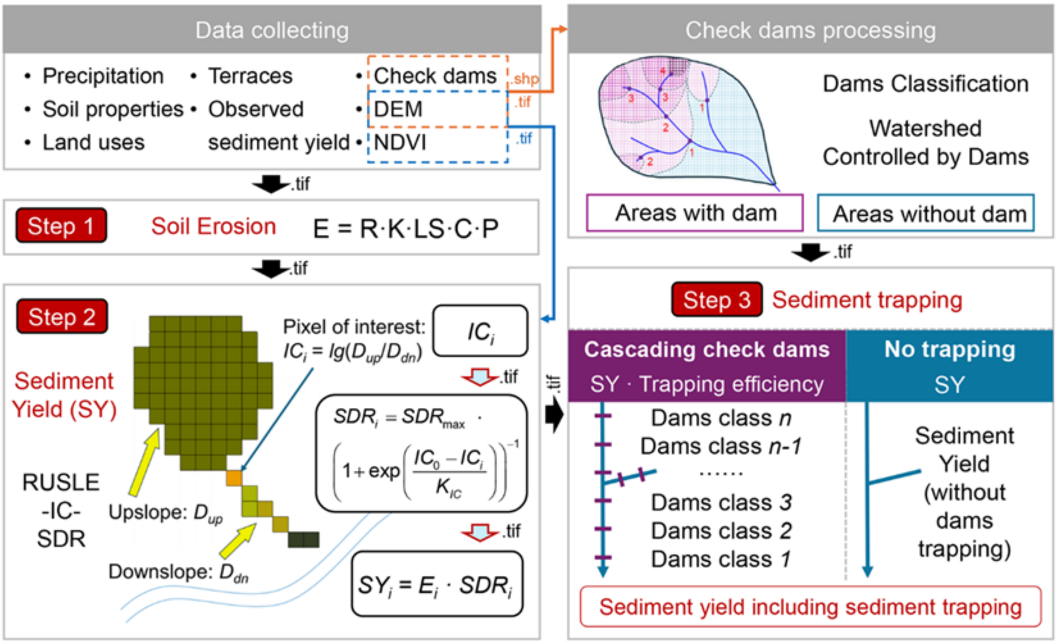


Figure 2. Framework of the integrated model for simulating sediment yield in the basin with cascading check dams. Step 1 denotes the soil erosion module by RUSLE. Step 2 indicates the sediment yield module by "RUSLE-IC-SDR" method. Step 3 represents the sediment trapping module by check dam. IC represents the index of connectivity, and SDR denotes the sediment delivery ratio.

<u>11. Comment:</u>

Figure 5: There are several black data points that significantly deviate from the fitted line in the scatter plot. Although the figure caption states that these points were excluded from the validation statistics, please add a sentence or two in the main text (Results section) explaining why these points were excluded (e.g., due to extreme flood events, missing data records, or anomalies caused by artificial water and sediment regulation).

Reply: *We will add an explanation in the Results section and provide additional diagnostic details in the Supplementary Materials. The revised text will state that three station-period records were identified as outliers or influential points using the Bonferroni-adjusted studentized residual test, Q-Q residual diagnostics, and Cook’s distance. These records showed large deviations between observed and simulated sediment yield, likely reflecting* *episodic flood events, local water-sediment regulation, or inconsistencies in long-term sediment records.*

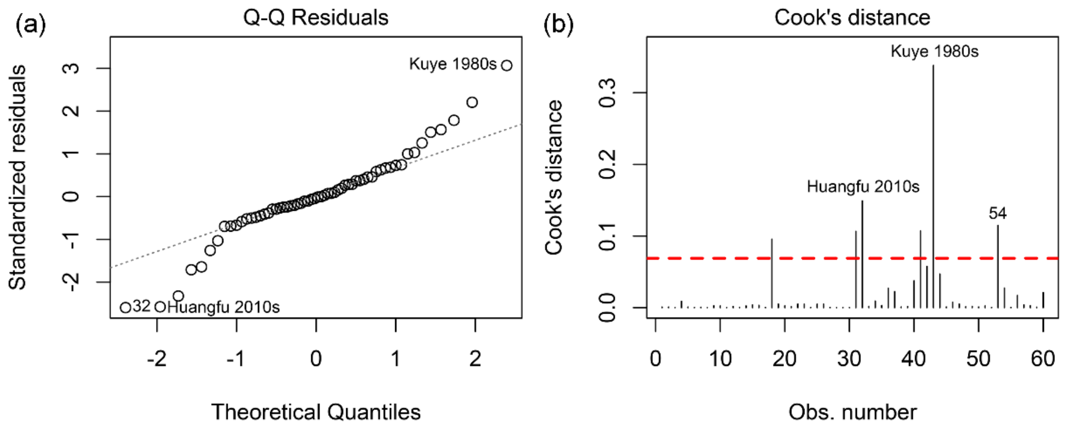


Figure S3. Diagnostic plots used to identify influential points in sediment yield validation. Influential points were identified using (a) Q-Q residual diagnostics and (b) Cook's distance.

12. Comment:

Figure 11: The concept of the x-axis "Relative benefit for SRCdam, Relative benefit for SARdam" is not fully explained in the figure caption. Please provide a brief quantitative definition or a reading guide in the caption (e.g., explaining the physical meaning of a bar leaning to the left with a red outline) to help readers quickly grasp the core conclusion without having to refer strictly to the supplementary material (Text S2).

Reply: *We will revise the caption of Figure 11 to provide a clearer definition and reading guide for the x-axis. The revised caption will explain that the x-axis represents normalized relative benefits of* SRC_{dam} *and* SAR_{dam}*, with leftward bars indicating* SRC_{dam} *benefit and rightward bars indicating* SAR_{dam} *benefit. We will also clarify that red outlines denote sub-basins with relatively high* SAR_{dam} *but low* SRC_{dam}*, suggesting insufficient check dams or high storage pressure, whereas blue outlines indicate relatively high* SRC_{dam} *but low* SAR_{dam}*, suggesting potentially excessive check-dam construction.*

Response to Anonymous Referee #2:

Comments - Questions (to be clarified in the text):

1. Comment:

Line 123, Page 5: What are GIMMS data?

Reply: *We will clarify the meaning of GIMMS at its first occurrence in the manuscript. The revised text will define it as the Global Inventory Modeling and Mapping Studies third-*

generation NDVI dataset (GIMMS-NDVI3g), produced from the Advanced Very High Resolution Radiometer (AVHRR) sensor.

2. Comment:

Line 164, Page 7: The calculation methods for RUSLE factors should be included in the main text, because RUSLE factors are not calculated usually according to the guidelines of the corresponding agricultural handbook (Renard et al., 1997).

Reply: *We agree that the calculation methods for the RUSLE factors should be more transparent in the main text. We will revise Section 2.3.1 by moving the key calculation methods for the R, K, LS, C, and P factors from the Supplementary Materials into the main manuscript.*

3. Comment:

Equation (2), Page 8: What does the index of connectivity express from physical point of view?

Reply: *We will clarify the physical meaning of the index of connectivity (IC) in the manuscript. The revised text will explain that IC expresses the potential structural connection between a sediment source cell and the downstream channel. It reflects the balance between upslope sediment-supply potential and downslope transport impedance. Higher IC values indicate stronger potential sediment delivery downstream, whereas lower IC values indicate weaker connectivity and lower sediment-transfer potential.*

4. Comment:

Line 231, Page 10: "Sediment accumulation in each check dam was recorded annually.....". Question: Were the calculations made also on an annual basis?

Reply: *Yes. We will clarify the temporal calculation sequence. Soil erosion was first calculated monthly and summed to annual values; SDR, SY, TE, and sediment trapping were then computed annually, with remaining dam capacity updated year by year.*

5. Comment:

Figure 5: What do the shaded areas on the regression lines indicate?

Reply: *We will revise the caption of Figure 5 to clarify that the shaded areas indicate the 95% confidence intervals of the fitted regression lines.*

6. Comment:

Line 383, Page 19: What is GGP?

Reply: *GGP refers to the Grain for Green Program. This abbreviation will be checked and defined at its first occurrence in the manuscript as "Grain for Green Program (GGP)" to avoid ambiguity.*

7. Comment:

General question: Is a sub-basin divided into cells?

Reply: *Yes. We will clarify that each sub-basin was represented by 100 m grid cells in the spatial modelling framework. Soil erosion, IC, SDR, SY, and dam-induced sediment trapping were calculated at the grid-cell level, whereas observed sediment-yield data at hydrological stations were used for sub-basin-level calibration and validation.*

8. Comment:

See annotated manuscript for secondary "editorial" errors!

Reply: *We will carefully review the annotated manuscript and address the editorial issues.*