The following is the point-to-point response to all the comments. The page and line numbers in the following response refer to the revised manuscript with changes marked.

Response to Editor:

Please address the comments raised by both reviewers. The revised manuscript will be reconsidered after major revisions.

Reply: We have carefully addressed all the comments raised by the reviewers and revised the manuscript accordingly. These revisions include clarifying key methodological details, discussing major sources of uncertainty, improving figures and captions, and correcting editorial issues throughout the manuscript.

Response to Anonymous Referee #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: Thank you very much for the positive comments. We have substantially addressed all the comments and incorporated them into the revised version of the manuscript, especially regarding spatial data resampling, parameter sensitivity, and the clarity of certain visual representations. Please refer to the replies to the comments.

Additional line comments:

1. 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 R^2 or correlation coefficients. I recommend briefly adding key statistics in the abstract to enhance the quantitative persuasiveness of the findings.

Reply: We have revised the abstract by adding the key statistical values for the correlation analysis (see Lines 27-29).

"The SRC_{dam} increased linearly with check dam density and the share of area they controlled ($R^2 = 0.81$ and $R^2 = 0.78$, respectively; $P < 0.001$), whereas SAR_{dam} increased logarithmically with SY from upstream of the check dams ($R^2 = 0.62$; $P < 0.001$)."

2. 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 suggestion. We have further emphasized the advantages of the novel topological sorting algorithm proposed in this study (see Lines 87-92):

"The framework uses a novel topological sorting strategy to establish the upstream-to-downstream routing order of individual check dam from grid-based flow direction, flow accumulation, and check dam-point locations, allowing basin-wide sediment trapping to be calculated. This improves computational efficiency, avoids double-counting in overlapping check dam-controlled areas, and preserves grid-cell-level spatial patterns of erosion, sediment yield, and check dam-induced sediment reduction."

3. 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: This is a good suggestion. We have given the description for NDVI data harmonization method between the GIMMS-NDVI-3g and MOD13Q1 datasets (see Text S1), and added the validation accuracy of the corrected GIMMS-NDVI-3g against MOD13Q1 datasets during their overlapping period, which prove the use of the corrected GIMMS-NDVI-3g for deriving the *C* factor in RUSLE during the pre-MOD13Q1 period (see Lines 136-139).

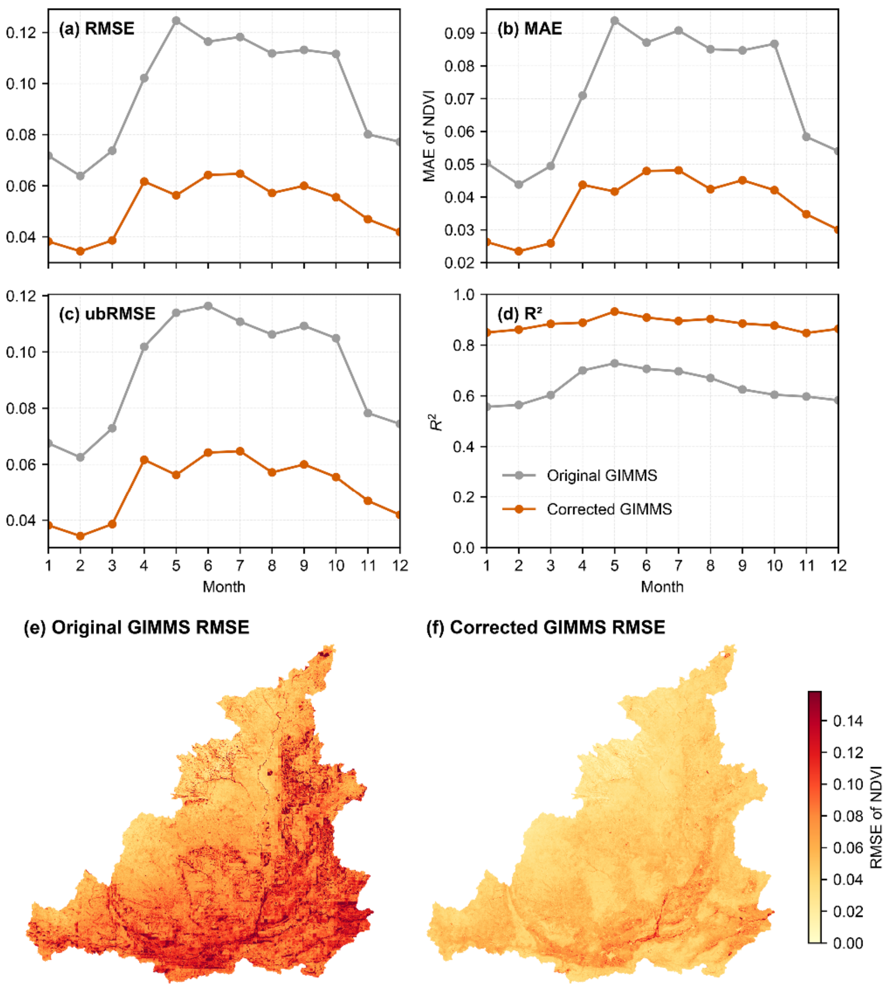

Figure S1. Accuracy evaluation of the corrected GIMMS-NDVI-3g using MOD13Q1 as the reference during the overlapping period (2000-2015). Panels (a–d) show the monthly variations in four accuracy metrics between GIMMS-NDVI-3g and MOD13Q1: root mean square error (*RMSE*), mean absolute error (*MAE*), unbiased 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-3g, respectively.

<u>4. Comment:</u>

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: This is a good comment. Yes, the DEM resolution can affect the estimation of topographic factors, especially the *LS* factor. In this study, to reduce topographic information loss from direct DEM coarsening, the *LS* factor was calculated at 30 m resolution and subsequently resampled to the common 100 m modelling grid (see Lines 189-191).

In the RUSLE modeling, the gridded variables were resampled to the 100 m grid before model calculation. The 100 m resolution was selected as a compromise between retaining basin-scale spatial heterogeneity and maintaining computational feasibility for 50-year grid-cell-based simulations across the MYRB (see Lines 203-206).

We have discussed the potential influence of DEM resolution on erosion estimation and future study scopes (see Lines 467-474):

"DEM resolution may introduce uncertainty into LS factor in RUSLE and SY estimates (Lu et al., 2020). Coarser DEMs can smooth steep gullies and small channels, whereas overly fine DEMs may increase data noise, storage requirements, and computational costs (Li et al., 2025). Therefore, DEM resolution should be selected according to study scale, terrain complexity, research objectives, and available computing resources (Bircher et al., 2019; Lu et al., 2020; Dile et al., 2024). Although the simulated SY was calibrated and validated using records from 17 hydrological stations, fine-scale erosion and deposition patterns may still be affected by DEM resolution. Future work should conduct multi-resolution sensitivity analyses for LS, IC, SDR, and SY in representative sub-basins."

5. 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: This is a good comment. As the Loess Plateau is dominated by fine-grained loess sediment, *D* = 0.046 was used as the baseline value following previous studies (see Lines 241-243).

To evaluate the influence of sediment texture on trapping estimates, we added a sensitivity analysis using different *D* values to simulated sediment trapped by check dams (see Fig. S11). We also discussed the potential uncertainty caused by using a fixed *D* under extreme rainstorm conditions (see Lines 474-483):

"The coefficient D is another uncertainty source because it reflects sediment texture and may vary with rainfall and sediment-source conditions in TE calculation (Verstraeten and Poesen, 2000). Our further analysis showed that larger D values increased long-term simulated sediment accumulation (Fig. S11). Moreover, the influence of D on TE depends on the ratio of remaining storage capacity to contributing catchment area (V/A_{cca}), with a comparatively limited effect for those with higher ratios. During extreme rainstorms, enhanced erosion may deliver more coarse particles to check dams (Zhu et al., 2024), increasing the effective D value and causing a fixed fine-sediment value to underestimate

TE, particularly for check dams with low V/A_{cca}. Extreme events may also rapidly reduce remaining storage capacity, accelerate TE decline (Chen et al., 2025), and increase overtopping or failure risks (Bai et al., 2020; Esteban Lucas-Borja et al., 2021). Future work should incorporate event scale particle size and hydrodynamic observations to develop time-varying TE parameterization."

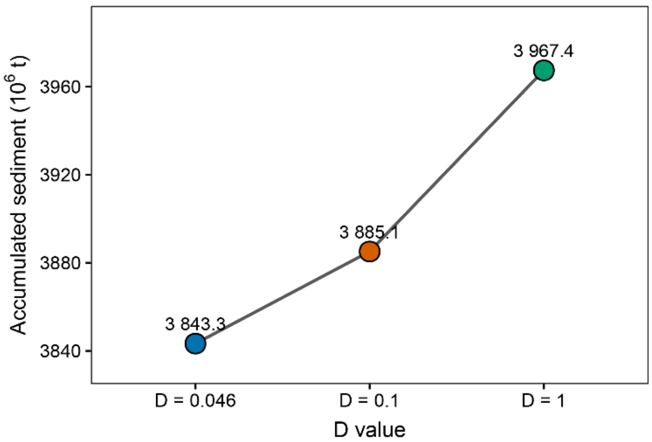

Figure S11. Accumulated sediment under different values of the empirical coefficient *D* in calculating sediment trapping efficiency.

6. 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: In this study, the contribution of the "Grain for Green" program to vegetation coverage and erosion reduction was reflected by both the *C* and *P* factors in the RUSLE model. The *C* factor was derived from monthly NDVI, which can capture temporal variations in vegetation cover and therefore reflects seasonal and interannual effects of vegetation cover on soil erosion. Thus, the vegetation-cover dynamics associated with ecological restoration, including the "Grain for Green" Program, can be represented by the *C* factor (see Lines 195-198).

Further, the *P* factor was updated using five land-use time slices (1975, 1990, 2000, 2020, and 2020) and available terrace maps, and therefore mainly represents decadal-scale changes in land use change induced by the "Grain for Green" Program (see Lines 199-200).

7. 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: This is a good suggestion. We have stated in the table footnote that the spatial resolutions listed in the table refer to the original datasets, and uniform resolution used for model calculation was on 100 m grid (see Lines 125-126).

8. 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 have edited Figure 1 to improve readability.

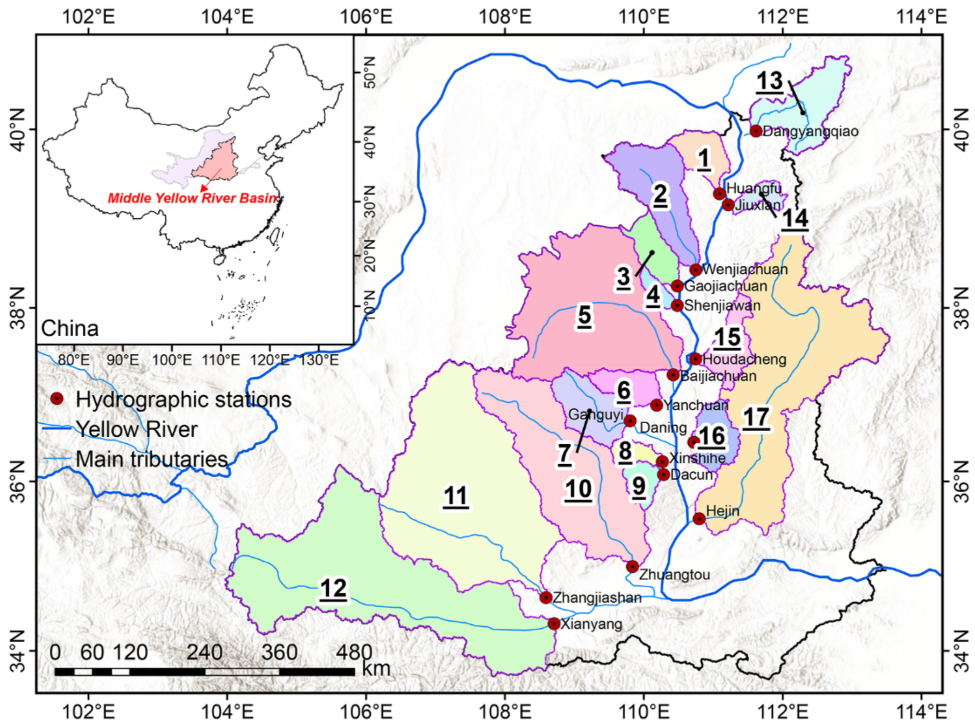

Figure 1. Distribution map of sub-basins and hydrological stations in the middle Yellow River Basin, with pink lines showing the sub-basins boundaries. The sub-basin IDs in the figure correspond to those listed in Table S1.

9. 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 have corrected the typo in Figure 2, and added brief labels next to the flow arrows to indicate the transferred data formats (e.g., .tif and .shp), which makes the grid-based processing logic more intuitive to understand.

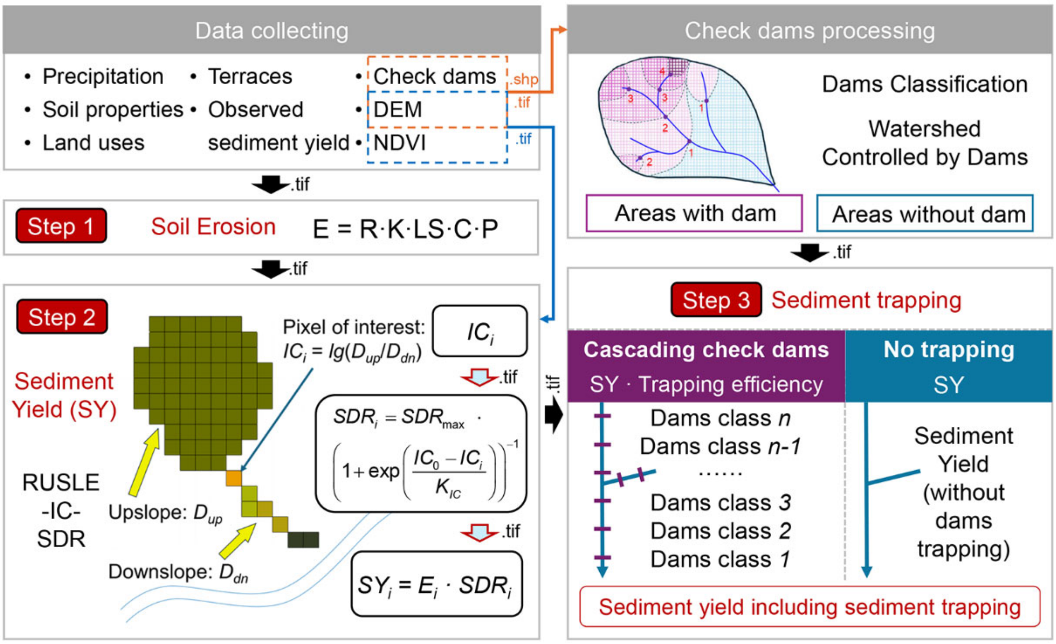

Figure 2. Framework of the integrated model for simulating sediment yield in the basin with cascading check dams. Step 1 denotes the soil erosion module by RUSLE. Step 2 indicates the sediment yield module by "RUSLE-IC-SDR" method. Step 3 represents the sediment trapping module by check dam. IC represents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

10. Comment:

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: In this study, the simulated *SY* was compared with observations from 17 hydrological stations to evaluate the overall performance of the model. Several *SY* outliers were identified using the Bonferroni-adjusted studentized residual test, Q-Q residual diagnostics, and Cook's distance (see Lines 294-297). The identified outliers may be associated with artificial water and sediment regulation, and they were excluded from the validation statistics (see Lines 331-332).

11. 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: This is a good suggestion. We have revised the caption of Figure 11 to provide a clearer definition and reading guide for the x-axis (Lines 454-457):

"The x-axis shows normalized relative benefits of SRC_{dam} and SAR_{dam}, with values ranging from 0 to 1. Red outlines indicate sub-basins with relatively high SAR_{dam} but low SRC_{dam}, suggesting insufficient check dams ("Insufficient"); blue outlines indicate relatively high SRC_{dam} but low SAR_{dam}, suggesting potentially excessive check dam construction ("Overmuch"); grey bars indicate other sub-basins."

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 given the definition of the GIMMS-NDVI-3g data (*Global Inventory Modeling and Mapping Studies third-generation NDVI dataset*) in the first occurrence (see Lines 128-130).

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: This is a good suggestion. We have included the calculation methods for RUSLE factors in the main text (see Lines 172-202).

3. Comment:

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

Reply: We have added the physical meaning of the index of connectivity (IC) in the revised manuscript (see Lines 225-229):

"Physically, IC expresses the potential structural connection between a sediment source cell and the downstream channel. A higher IC value indicates stronger upslope sediment-supply potential and/or lower downslope impedance, meaning that eroded sediment is more likely to be delivered to the stream network. A lower IC value indicates weaker connectivity, longer or rougher transport pathways, and a higher probability of deposition before reaching the channel (Shi et al., 2025)."

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: This is a confusing statement. In this study, the erosion rate was firstly calculated on a monthly scale using the RUSLE model, and then summed up to obtain the annual erosion rate (see Lines 166-167). All sediment routing and trapping calculations were conducted on an annual basis. For each year from 1970 to 2020, sediment trapping efficiency of check dams was calculated using the remaining storage capacity at the beginning of that year; trapped sediment was then subtracted from the remaining capacity and carried forward to the next year (see Lines 276-279).

5. Comment:

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

Reply: In Figure 5, the shaded areas indicate the 95% confidence intervals of the fitted regression lines (Lines 332-333).

6. Comment:

Line 383, Page 19: What is GGP?

Reply: GGP refers to the "Grain for Green" Program (conversion of cropland on slope to forestland and grassland) (see Lines 109-110). We have used the full name, rather the abbreviation to avoid confusion (see Line 436).

7. Comment:

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

Reply: In this study, each sub-basin was represented by 100 m grid cells, which served as the basic computational units for estimating soil erosion, sediment yield, and dam-induced sediment trapping. The sub-basins were used for model calibration and validation, and aggregation of grid-cell-level results (see Lines 284-286).

8. Comment:

See annotated manuscript for secondary "editorial" errors!

Reply: Thank for very much for detailed editing. We have carefully reviewed the annotated manuscript and addressed the secondary editorial issues throughout the revised version.