Associate Editor Comments:

Nearly all comments have been carefully addressed. It is good that it is now clear from the introduction that this study provides a proof of concept, several comments from the reviewer have herewith been addressed.

1. Please extend your explanation of why you focus on 2 instead of 1 models. It is perfectly fine and even interesting, but the added sentence ' As the NWM and NHM have different constructions, their sensitivity to catchment drivers is likely to differ' does not provide sufficient explanation for a reader who does not know the two models and the most important differences between them. The response to reviewers is more extended and valuable.

Apologies for the oversight, we have now rewritten this intro paragraph and included what exactly differs and why these differences are crucial for our approach. Key notes and changes are underlined.

> *"The National Water Model (NWM) and the National Hydrologic Model (NHM) are two process-oriented, continental-scale hydrologic models used in operational decision-making (Towler et al., 2023). The NWM framework applies the Weather Research and Forecasting Hydrologic model (WRF-Hydro) formulation, which simulates infiltration, evaporation, transpiration, overland flow, shallow subsurface flow, baseflow, channel routing, <u>and passive reservoir routing, but not active reservoir management</u> (Cosgrove et al., 2024). The NHM framework applies the Precipitation-Runoff Modeling System (PRMS) formulation, which represents evaporation, transpiration, runoff, infiltration, interflow, groundwater flow, and channel routing, <u>but not reservoir operations, water withdrawals, or stream releases</u> (Regan et al., 2019). <u>See Text S1 for more details on each model</u>. <u>A key distinction is that the NWM targets high spatial (~250 m) and temporal (hourly) resolution flood forecasting. In contrast, the NHM assesses long-term water availability at hydrologic-response-unit scales (~100 km^2, driven by daily forcing) (Towler et al., 2023).</u> Both models exhibit spatially variable streamflow skill across US catchments (Tijerina et al., 2021), with the strength of prediction varying as a function of catchment-scale climate, land use, and physiography. <u>Collectively, differences in resolution, process formulation, and treatment of human regulation make the NWM–NHM pair an ideal testbed for structural sensitivity analysis: drivers influential in both frameworks likely denote overarching hydrologic controls, whereas divergent sensitivities flag processes that are represented differently (or omitted) in either approach.</u>"*

2. Reviewer 2 requested more specification on the NWM and NHM in the Methods section or supplementary materials, this has not been addressed, while most readers will not be familiar with the two models on which data this study is centered. I strongly argue that a brief description or overview table of the 2 models is required, this can briefly be done in the main text and extended in the supplementary information. The reader should not be obliged to read the other papers on the models to get to know more details. For the analysis it is also relevant to know more details on the difference between the models than the text that is currently provided in the introduction.

Thank you for bringing this point to our attention, we fully agree that providing more information on the models behind this study would aid the readers in interpreting our results. We have provided greater overview information on the two models in the introduction as described above, and we have added a text section in the supplement providing more detail on both of the models.

We have added this text at the start of the methods section:

> *"Text S1 summarizes the models that produced the data used in this study."*

We have added a new supplement (Text S1):

> ***Text S1: An overview of the National Water Model (NWM) maintained by the Office of Water Prediction of NOAA and the National Hydrological Model (NHM) maintained by the U.S. Geological Survey. These two national-scale models, while both aiming to simulate hydrological processes across the continental U.S. (CONUS), differ significantly in their underlying modeling frameworks, primary operational objectives, spatial discretization, input datasets, and the specific hydrological processes they explicitly represent. Please see Towler et al. (2022) for additional details on each model****.*
>
> ***National Water Model (NWM) Version 2.1***
>
> *The National Center for Atmospheric Research (NCAR) developed WRF-Hydro, an open-source hydrologic model that serves as the foundation for the National Oceanic and Atmospheric Administration's (NOAA) National Water Model (NWM). NWM simulates and forecasts key water components (e.g., evapotranspiration, snow, soil moisture, streamflow) in real-time across the continental U.S., Hawaii, Puerto Rico, and the U.S. Virgin Islands. NWM version 2.1 utilizes 1 km atmospheric data from NOAA's Analysis of Record for Calibration (AORC) and employs the Noah-MP land surface model to compute energy and water states on a 1 km grid. Hydrologic routing occurs on a 250 m resolution terrain grid, utilizing WRF-Hydro's baseflow parameterization and the Muskingum–Cunge river routing scheme on an adapted NHDPlus version-2 river network. The model features a level-pool scheme for 5,783 lakes and reservoirs, although it lacks active reservoir management. While operational data assimilation is included, it is not applied in the retrospective simulations. Calibration of 14 parameters occurred from water years 2008 to 2013, validated against data from 2014 to 2016 across 1,378 gaged basins.*
>
> ***National Hydrological Model (NHM) Version 1.0***
>
> *The U.S. Geological Survey (USGS) developed the National Hydrologic Model (NHM, version 1.0) based on the Precipitation–Runoff Modeling System (PRMS), a modular system often employed for water resource assessment and scenario analysis. NHM simulates water flow and storage processes, including snowpack, soil, and stream networks, using daily discharge simulations. The NHMv1.0 results used here come from a calibration workflow focused on observed streamflow and the Muskingum–Mann routing option. Climate inputs consist of 1 km resolution daily precipitation and temperature data from Daymet. The model's spatial structure is defined by geospatial fabric version 1.0, which for PRMS typically delineates Hydrologic Response Units (HRUs). Calibration employs a stepwise approach to optimize parameters for water budgets and streamflow, first aligning hydrologic responses to baseline observations and then timing streamflow against data from 7,265 headwater watersheds. Final calibration occurs at 1,417 stream gage locations. The calibration period spans odd water years from 1981 to 2010, with*

validation using even years. NHM does not simulate reservoir operations or water withdrawals; it outputs daily streamflow for analysis.