This paper presents an interesting study that applies graph neural networks (GNNs) to model discharge over a river network. The methodology is sound, the paper is well written, and the structure is clear. This could be a very important contribution to the rainfall–runoff literature.

Recommendation: minor revision. Most of my comments are about clarity and reporting. However, a few items ***could become major*** if the justification is weak or if the implementation is not as intended: (a) whether the baseline LSTM comparison is fair (retrained vs "extracted"), (b) whether a 180-day input window is sufficient for the target processes and basin scale, and (c) whether the forcing selection (especially using soil moisture as an input) is appropriate and clearly framed/justified.

General comments:

1. The title can be more specific. For example: "A GNN routing module is all you need for **routing** LSTM rainfall–runoff models …" or "… for **accounting for routing** in deep learning-based rainfall–runoff models."
2. Please expand the literature review on existing GNN applications in hydrology in the Introduction (around lines 68–70). The current text mainly states that prior work does not use GNNs to explicitly model routing.
3. The model uses three forcings: precipitation, soil moisture, and air temperature (lines 100–102). Why is soil moisture treated as a meteorological forcing (or as a state / reanalysis product)? Please justify this choice, and explain why other variables (e.g., wind, radiation) are not included, as they are commonly used in other LSTM studies.
4. The introduction to the different GNN architectures should be explained in more detail, since this is a key selling point. Only listing names and references may not be sufficient for many readers. A short, verbal description of each architecture (or a simple summary in the Supporting Information) would help.
5. Baseline comparison: lines 193–194 mention a baseline LSTM. Is this LSTM re-trained / fine-tuned as a standalone model, or directly extracted from the LSTM-GNN framework? If it is extracted without proper re-training, this could bias the comparison, which *could* be a major problem of the paper. Please clarify.
6. Lines 130–131: why is a 180-day sliding window used? This may be short for capturing annual-scale dynamics of catchments. Please justify this choice and, if possible, add a sensitivity test (e.g., 365 days or longer).

Specific comments:

1. Line 161: please check the notation for concatenation of two vectors.

2. Line 169: "*which can be defined in different ways to investigate the impact of river network representation*", please add more details (what are the different definitions considered?).
3. Line 171: please clarify whether $A_{j,i}$ is non-zero whenever $A_{i,j}$ is non-zero (i.e., whether the river connections are directional or symmetric).
4. Table 1: please improve table formatting; the square-root notation is not clearly visible.
5. Figure 2: please improve aesthetics/readability; arrows and text overlap and are hard to read.
6. Figure 3: in the caption, "m3/s" should use a superscript: m^3/s.
7. Line 240: "NSELST-GAT-NSELSTM" — the mathematical notation is unclear. Please revise for readability.