General Comments

The authors provide a description of a new hydrologic model framework, CREST-VEC, which couples CREST and vector-based routing framework, mizuRoute, and sub-surface runoff and lake schemes are also incorporated. Compared with the raster-based routing scheme implemented in the original CREST, the new framework shows a significant improvement in computational efficiency and reproducibility in river discharge. Therefore, the proposed model has an essential potential for a flood forecast and alert system, as the authors suggested, and the manuscript is informative for the readers of this journal. However, since two contexts (speed-up of the model and introduction of new physical processes) are mixed in the manuscript, it requires some revision of the manuscript and experimental configurations for a well-structured article.

Specific Comments

For the abstract and corresponding sections:
The authors are highly requested to discuss 1) vectorization of the routing scheme and the improvement in computational efficiency (and the ensemble simulation, as they suggested) and 2) the introduction of new physical processes, the change in reproducibility of river discharge, and the false alert problem separately. Therefore, the article should be divided into two parts for the sake of readability. However, if they are still to be published as one paper, some parts need to be reorganized. For example, the introduction should describe the significance of the subsurface flow and lake routing for the flood forecast.

The authors used computation time per step as a measure of computational efficiency. However, the number of computation steps varies depending on time-step width and constraints such as CFL conditions in some models (not sure if this is the case with your model). Therefore, it would be more appropriate to use the time it takes to compute a given period on a given spatial resolution (e.g., one year) rather than the computation time per step. The number of parallels and parallel efficiency are also important indicators for comparing computation algorithms.

In my understanding, the routing scheme calculates the time lag from runoff from the land surface to downstream. So, why does the bias change between CREST and CREST-VEC, as shown in Figure 4 and Table 1? Or is this bias calculated for peak flows? (One possible factor is the use of externally derived reservoir storage. However, such modifications should not be applied even if they improve the model accuracy since it is difficult to discuss the impacts of the model update when the water budget in the overall model is changed.)

1 Introduction

Mainly for line 69: Even if a numerical model can be run at a fine spatial resolution in a realistic amount of time, it does not necessarily mean that the physics assumed in the model hold at the resolution. For example, a 1-D river model with a 100-m resolution is difficult to apply directly to a wide river such as the Amazon River, even if it can be run. The authors' work is technically excellent, but the validity of the model physics should be discussed in the discussion section.

Line 84-85: Raster-based and vector-based models are reviewed in detail in the introduction section. However, information relevant to the research question listed here should be added. If previous studies on the impact of subsurface processes and lake considerations on a river model are presented, the reader will better understand the motivation for the study.

3.2 CONUS simulation

Line 263-289: This discussion using the basin attributes suggests a significant uncertainty in the land surface process, not the routing schemes. Remove it, or explicitly describe the contribution of the CREST-VEC (+subq+lake) to the discussion compared with the results when the same analysis is applied to the original CREST.

3.3 How likely are floods falsely detected?

Line 315-317: Previous research also reported the incorporation of a lake scheme mitigates the seasonal variability in the river discharge (Tokuda et al., *GMD*, 2021).

Technical Corrections

1 Introduction

Line 25: Add "." after "1".

Line 38: "the most time-consuming". The expression is too strong and should be loosened. It is because atmospheric models and (even within the framework of a land surface model) land surface models that account for 2-D groundwater flow are computationally expensive.

Line 57: 2011; Yamazaki et al., 2011;) remove the last ";".

2. Data and methods

2.1 Hydrography data

Line 92: Remove "MERIT-Hydro" before Lin et al. (2019), which corresponds to Yamazaki et al. (2019).

Line 94: Add the reference for the MERIT DEM (Yamazaki et al., 2017).

2.2 Forcing data

Line 118: "rainfall" implies snowfall is excluded?

2.5 CREST-VEC

Line 174: The reservoir operation is incorporated in the mizuRoute but does not the CREST-VEC consider it?

3. Results

3.1 Case study: Houston region

Line 210: Related to the major comment, why does the CREST-VEC overestimate the peak compared to the original CREST?

Line 219 (Figure 3): Make the orange lines visible.

3.3 How likely are floods falsely detected?

Line 321: The contribution of the routing scheme to the flood alert system is an essential topic for our society. However, additional analysis is needed since the current analysis is highly affected by the bias of the runoff, and we can not detect the impacts of the new routing and lake schemes from the results.

4. Discussion

4.1 Vector vs. Raster routing

Line 353-355: Which result (or previous research?) suggests this?

Line 359-361: What is the relationship between the one-to-many river network and the following sentences? What does “on the other hand” mean?

Line 363-368: What is the difference in the sub-grid routing scheme between raster- and vector-based approaches? The section title is “vector vs. raster”, but the sub-grid scheme represents the hydrodynamics within each grid, and it has nothing to do with the vector or raster approaches.

4.2 Room for improving large-scale hydrologic simulation

Line 370-396: The paragraphs are too long to discuss the uncertainty caused by the IRF scheme. Make them shorter. (Or, do you have any plan for the calibration with the latest, faster CREST-VEC model?)

4.3 How to operate flood forecasting with regulated flow?

Line 408 (The section title): Since the reservoir operation is not considered in this experiment, “with a lake scheme” is more appropriate instead of “with regulated flow”.

Line 411-412: Which process should be improved in the current lake scheme? The improvement plan should be discussed in the previous section (4.1).

Line 415-419: Does not this approach cause the “cry wold” effect?