Though the authors have made substantial revisions to the manuscript, I believe that my prior comment regarding reservoir aggregation in major issue hasn't been addressed. I think the authors should improve their Methodology, and further update the results.

Lumping reservoirs storage for water supply may be reasonable when only focusing on the total amount of water supply. However, it's quite unjustified to lump reservoirs storage when it comes to flood control. Neither the hydrological connections between reservoirs nor the spatial distribution of flood is considered, which will lead the overestimated local flood control capacity, and further severe loss. A simple example is detailed below:
There are three cascade reservoirs A, B, and C, from upstream to downstream. The storages for flood control of A, B, C are 100, 200, and 300, respectively. Lumping reservoir storage indicates that the flood control storage becomes 600. The 600-level flood can be regulated through reservoir operation when the flood comes from upstream side of reservoir A. However, as the 600-level flood comes from the river segment between reservoirs B and C, the flood will exceed the regulate capacity of the system, which cannot be identified by lumping reservoir storage. Therefore, it's highly-risked to lump reservoirs storage for flood control.

***Authors' response*:** We thank you for your kind review of our paper, your valuable comments earlier, and your comment on our revised manuscript. We appreciate that you have further concerns about the lumped system and have addressed them below.

We would like to illustrate our lumped system in the model through Fig. A below. In lumped system of Canada there are three treaty dams namely, Mica, Arrow and Duncan with their respective storages as 24.7, 10.3, and 1.77 km^3, which are 67, 28, and 5% of the total installed storage respectively. The stream through Mica and Arrow Dams is the primary stream order of Columbia river and the one through Duncan is a small tributary. These three dams are combined to calculate flood control and hydropower. We did consider all the hydrological connections and segments that constitute the inflows for each lumped system, as shown in Fig. A. Similarly, in lumped system of the U.S. there is only one storage dam i.e., Grand Coulee (GCL) Dam and rest of the treaty dams are non-storage / run-of-river hydropower plants as shown below in Fig. A. Therefore, the storage is not lumped for the U.S., and reservoirs are only lumped for the hydropower estimation, which is appropriate given the run-of-river design. The schematic of the streamflow and dams is also presented in Fig. 1 of the revised manuscript; and the nature of treaty dams is also illustrated in Table. 1 of the revised manuscript.

Furthermore, in our model we also considered hydrological connection through the international boundary to the location at The Dalles where flood damage usually occurs. We considered the tributaries along the international boundary that contribute to inflow in lumped system in U.S. (or the GCL dam) along with the outflow from lumped Canadian system. As illustrated by the comparison of observed and simulated streamflow for GCL dam (Fig. 5c and Fig. 6c of the revised manuscript), our computation replicates streamflow patterns appropriately for the paper question and scope. In addition, the hydropower benefits (i.e., the response variable) that depends on the streamflow (the independent variable) included in the model was able to accurately reproduce hydropower benefits for the U.S. (Fig. 5d and Fig. 6d of the revised manuscript). We are confident that hydrologically the lumped system is appropriate for the research question and scope.

The spatial distribution of the flooding is an important topic and is one of the major limitations of the lumped models. However, the spatial distribution of the flood is not relevant for the scope of this study because the historic flood damages occurred downstream of Columbia river at "The Dalles" and which was one of the key motivation for the CRT, thus in our study we considered flood control benefit by specifically considering flooding at The Dalles.

We have added the description of the lumped system throughout the section *3.2 Equations and parameters*, for Canadian and U.S. inflow (in lines 383 – 392 in revised manuscript), and for hydropower production is the lumped systems (in lines 551 – 556) of the revised manuscript. We realize that elaborating on the lumped systems might clear potential confusion and help readers. Thus, we added a section 4 "*Lumped systems in the model*" in the supplemental material with the description shown in Fig. A below. We also elaborated text in the section "*3.1 Socio-hydrological system dynamics model*" of the re-revised manuscript as discussed below in *Revision*.

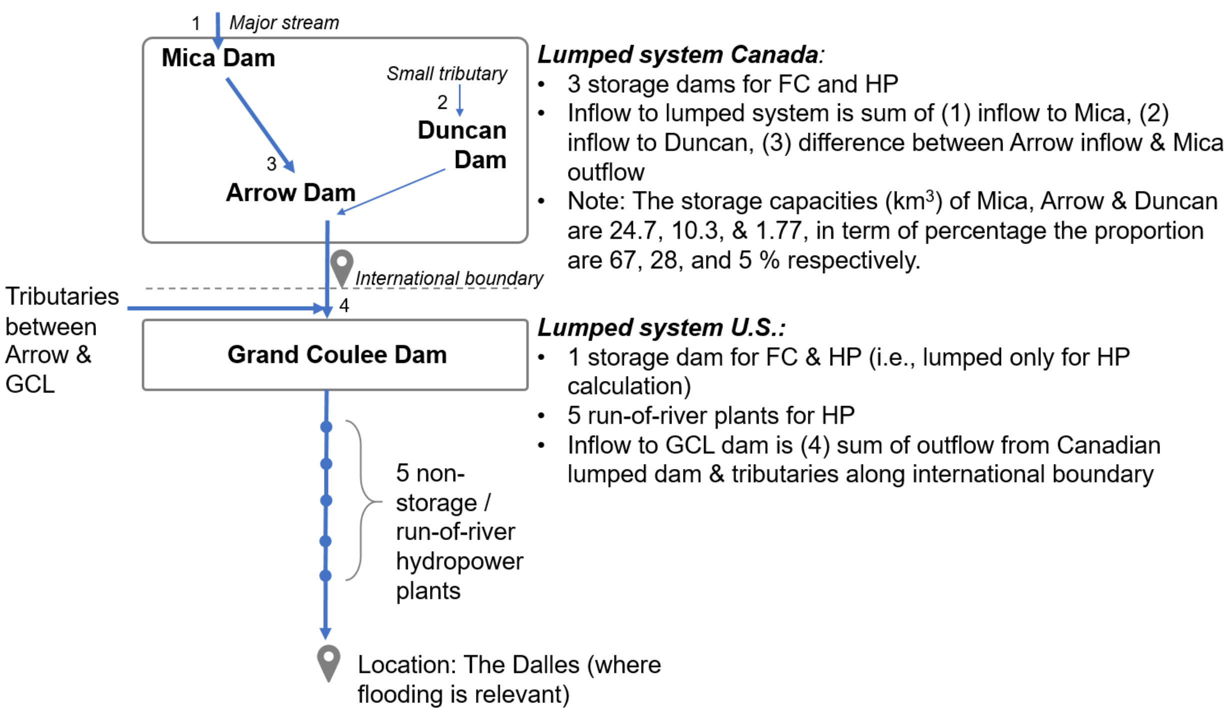

Figure A. Lumped systems in Canada and U.S.

Revision: We have added text (as shown by blue color) in third paragraph of the section "*3.1 Socio-hydrological system dynamics model*" as below:

The storage capacity of Canada (upstream) and the U.S. (downstream) are two important state (hydrological) variables which represent the aggregated storage of the treaty dams (Fig. 2), the operation of which is determined by the storage thresholds. The increase in a storage threshold results in an increase in the storage level. Three Canadian dams namely Mica, Duncan and Keenleyside are lumped into a single storage as all three dams are multifunctional for flood control and hydropower production. However, it should also be noted that Mica and Arrow Dams are the major dams in Canada contributing to flood control as those are along the primary stream order of Columbia River and Duncan Dam is in the small tributary (Fig. 1). In terms of storage volume Mica, Arrow and Duncan Dams are 24.7 km^3, 10.3 km^3, and 1.77 $km^{3,}$ or 67%, 28%, and 5% of total storage, respectively (Table 1). In the U.S., the Grand Coulee dam is the only multifunctional dam with useable storage for flood control. Given that the Grand Coulee is the only dam with storage in in the U.S. the system, we have only lumped the reservoirs for hydropower generation, not flood control. We used the lumped reservoir approach to simplify the system process required to investigate our research questions. The lumped approach is particularly appropriate because all the treaty dams work in coordination to achieve either of the hydropower benefits (by U.S. dams) or flood control (by Canadian dams). The schematic of the lumped system is also shown in Fig. S18, Section S4 of the supplemental material. In lumping the system, we have considered external input variables such as tributaries and added to the outflow from Canadian reservoir, or inflow to the U.S. reservoir. These dams along the Columbia River either have significant flood control capacity or significant hydropower production capacity (Table 1). Thus, the simplified reservoir operation described below in Sect. 3.2.1 was implemented in the lumped storages on each side of the border, which represent collective operation of all the treaty dams within each country. Other hydrological variables in the model (i.e., flows in the CL diagram) are inflow into Canadian storage, outflow from Canadian storage plus intermediate tributaries, inflow into the U.S. storage, and outflow from the U.S. storage. The higher the outflow from the dams, the lower the flood control as flood damages increase. A portion of the reservoir outflow passes through hydroelectric turbines, thus more outflow yields higher hydropower benefit.