This study constructs numerical models to simulate the hydrological processes in the Middle Heihe River Basin in response to glacier loss, permafrost degradation, and temperature increase. The topic is interesting and the study area is an important area. However, there are some problems in the manuscript, which are listed as following.

Major comments:

- Commonly, specific yield is smaller than porosity. In the manuscript, the authors used specific yield as an analogy for porosity. In addition, the authors simplified specific yield values in the model. Specific yield data from 17 unique values calibrated in Tian et al. (2015a) were simplified to three intervals of 0.1, 0.2, and 0.3. The authors stated that this simplification was used to lessen computational demand. The reviewer doubts the reasonability of such simplification. This oversimplification may cause the model far from real condition and the results may be inaccurate. As supercomputer and parallel computation is so common, computation demand may not be a problem.

- The vertical thickness of the model is 472 m, which may be too shallow for a groundwater model. Many regional groundwater models have vertical thickness of 3-5 km. Can the 472 m thickness capture the main groundwater flow system in the study area?

- For a regional groundwater model with relatively large thickness, the decrease of hydraulic conductivity K and specific storage Ss with depth should be taken into account. Did the authors consider the decrease of K and Ss with depth in their model?

- The constant flux boundary condition along the border between the Upper and Middle
Heihe may not be reasonable. The flux from the Upper Heihe is variable, and the flux is different in different seasons and among different years. The authors should clarify why a constant flux boundary is reasonable.

- The authors stated that "about 75% of water coming into the middle basin domain is from streamflow, 20% from precipitation, and 5% from the groundwater boundary condition." Where are the percentages from? Are there any evidence for these percentages?

- A uniform water table depth of 20 m was used in the model. Why the authors use a uniform water table depth? Is there any support material, publications, or evidence for such a water table depth? For such an area, water table depth should be different in different regions.

- For the Combined scenario, the authors used 15% reduction in thawing season flow and 50% reduction in baseflow. Are these reduction percentages have any supporting data, references, or other evidences? Or are they chosen arbitrarily?

- For comparison of the observed and modeled flow at the HRB2 gage, it can be seen from Figure 3 that the modeled values are significantly smaller than the observed values. The fit between observed and modeled data should be greatly improved.

- The authors use Spearman's rho as the standard to determine correlation. Why not use the Nash–Sutcliffe efficiency coefficient to evaluate the fit between observed and modeled data? The Nash–Sutcliffe efficiency coefficient is a widely accepted standard for this purpose.

Minor comments:

1) Add a space between numbers and its units, throughout the manuscript.

2) Line 20: 2C should be 2°C, there are other places in the manuscript

3) Line 40: Zongxing et al. should be Li et al., Li is the family name. Please correct this throughout the manuscript.

4) Line 85: "ground and surface water" should be "groundwater and surface water"

5) Lines 216 and 231: Abbreviation "2011WY" should be defined at the first time appearance.

6) Line 494: There is a typo: 3.32.9
7) Figure 14: There is no need to give two color legends.

8) Check the correctness of the references. For example, In Line 801: Zongxing L. should be Li, Z., Li is the family name and Zongxing is the given name. Similar mistakes lake Hongyi L. and Yongge L.