

This manuscript presents an interesting and timely evaluation of how multi-source observational datasets can be used to constrain the pywatershed in snow-dominated headwater catchments. The integration of diverse data products (such as ASO SWE and OpenET) to address the pervasive issue of model equifinality is a significant undertaking that holds substantial value for the hydrological community. The study's focus on identifying which observations provide truly informative vs. potentially misinformative constraints is particularly important for advancing operational modeling practices. However, while the research is promising, there are several points that warrant further discussion and clarification before the paper can be considered for publication. I recommend Major Revision.

Major Comments:

Q1: The sensitivity analysis identifies forcing correction parameters, such as *tmax_cbh_adj* and *tmax_allsnow*, as highly sensitive across all target observations. Please explain whether this heavy reliance on "tuning parameters" to correct meteorological input errors might be masking structural deficiencies in the model's snowmelt or runoff generation mechanisms.

Q2: The Morris analysis indicates that almost no runoff-related parameters were identified as sensitive in this study. Please explain if the pywatershed model overemphasizes meteorological forcing corrections while potentially neglecting the regulatory role of catchment topography and geological attributes on runoff flow paths.

Q3: The results show that the adjacent East and Taylor River catchments respond quite differently to the same observational datasets. The author suggests that the performance discrepancy between the adjacent East and Taylor River catchments may result from differences in hydroclimatic variables such as runoff ratio or aridity. However, these variables are not explicitly quantified or compared for the two study basins in the manuscript. To support this claim, the author should provide a table or figure comparing the key physiographic and hydroclimatic attributes (e.g., mean slope, forest cover, runoff ratio, and aridity index) of these two basins to explain why the model's physical consistency varies so significantly between them?

Q4: Does the structure of the pywatershed model may struggle to remain compatible with the simultaneous integration of high-spatial-resolution (ASO) and high-frequency (SMAP) remote sensing data? Table 2 reports several instances where the number of behavioral parameter sets is zero when certain observations are included. My primary concern lies in the potential for model structural failure under these advanced constraints.

Q5: While Actual Evapotranspiration (AET) is the only variable to consistently improve simulation performance, it is evaluated using monthly averages. Could this improvement be an artifact of the temporal averaging smoothing out daily error variances?

Minor Comments:

Q1: The layout of Figure 1 appears somewhat cluttered and would benefit from a clearer visual hierarchy to help readers quickly identify the study's geographic context.

Q2: In Figure 4, the Spearman correlation R values and p-values overlap with data points in several

subplots, hindering legibility. Please adjust the placement of these statistical annotations to improve clarity.