

This manuscript addresses the important issue of quantifying inter-basin groundwater flow (IGF). Recognizing that water balance residuals may contain both data errors and IGF, it accounts for data uncertainty and uses Gaussian process regression (GPR) with distinct inputs and covariance structures to probabilistically decompose the residuals. This design helps avoid attributing the residuals to a single source a priori. However, the IGF estimates depend on the priors and error model, and the robustness of the key findings and their physical interpretations requires further validation. I acknowledge that directly monitoring IGF in large catchments to validate the estimates is challenging. Therefore, the manuscript should clearly distinguish model inference from independent physical validation and use more cautious and consistent wording when presenting the relevant conclusions and their implications. Please see the comments below.

Comment 1: Within a mass conservation framework, water balance residuals calculated from observed or estimated data may reflect both data errors and inadequately observed fluxes or storage changes, including inter-basin groundwater flow. Some studies redistribute the residuals according to the error covariance of the water balance components and impose a strict water balance closure constraint (Pan and Wood, 2006; Zhang et al., 2018). Other studies, assuming that data errors are negligible or insufficient to explain the main imbalance, combine the residuals with other evidence and interpret them primarily as net inter-basin groundwater flow (Fan, 2019; Liu et al., 2020). I appreciate that the additive-GPR framework used in this manuscript allows IGF and data errors to coexist, uses distinct inputs and covariance structures for probabilistic decomposition. Nevertheless, the inferred IGF should still be interpreted as a conditional estimate under the specified priors and error model, because systematic data biases associated with the hydrological state may also be assigned to the IGF component. Although the authors acknowledge in the limitations section that the decomposition depends on assumptions about the inputs and smoothness, I recommend that this interpretive limitation be stated consistently in the Results, Discussion, and Conclusions. In particular, when discussing the magnitude and direction of IGF and its interactions with other hydrological processes, the authors

should clearly distinguish model inference from independent physical validation and avoid wording that implies unique identification of the actual groundwater fluxes.

Comment 2: The finding that 47% of the catchments satisfy $|IGF/P| > 0.2$ is important, but the associated uncertainty should also be reported. Please clarify how many catchments still support exceeding the 20% threshold after accounting for the uncertainty intervals.

Comment 3: The relationship between catchment area and the seasonal variation in IGF requires quantitative support, including effect sizes and statistical tests. Figure 4 shows high STL-based seasonality strength in some individual large catchments, but this does not establish a general tendency for IGF seasonality to increase with catchment area. High values of STL-based seasonality strength are concentrated mainly in some catchments in Africa and Asia, which may be related to the strength of climatic seasonality.

Comment 4: The Abstract and Discussion suggest that neglecting IGF reduces the reliability of climate projections from Earth system models. However, this study does not include model evaluation, comparisons with and without IGF, or climate projection experiments. I recommend reframing this conclusion as an implication for future research on land surface and Earth system models.

Comment 5: The proportion of catchments exhibiting P–IGF and IGF–ET links identified by PCMCI+ varies with the significance level. I recommend assessing the sensitivity of the analysis results to the significance level and clarifying how the conclusions change when stricter significance thresholds are applied.

Comment 6: I recommend revising the wording used to interpret the hydrological mechanisms. The P–IGF and IGF–ET links identified in this study provide valuable clues for understanding interactions between groundwater and other hydrological

processes. However, specific mechanisms should be described more cautiously so that the interpretations remain consistent with the evidence available from monthly data. For example, a zero-lag P–IGF link at the monthly scale may be related to a rapid response of the groundwater system following precipitation, and hydraulic pressure propagation is one physically plausible explanation. However, the current analysis cannot clearly distinguish among rapid recharge, hydraulic pressure propagation, and the effects of hydrological variables omitted from the analysis. Similarly, the significant negative IGF–ET association found in 47% of the catchments provides a basis for discussing the potential contribution of net groundwater inflow to evapotranspiration, but the specific pathways supplying this water have not been directly validated.

Fan, Y. Are catchments leaky?, *WIREs Water*, 6, e1386, <https://doi.org/10.1002/wat2.1386>, 2019.

Liu, Y., Wagener, T., Beck, H. E., et al. What is the hydrologically effective area of a catchment?, *Environ. Res. Lett.*, 15, 104024, <https://doi.org/10.1088/1748-9326/aba7e5>, 2020.

Pan, M. and Wood, E. F. Data Assimilation for Estimating the Terrestrial Water Budget Using a Constrained Ensemble Kalman Filter, *J. Hydrometeorol.*, 7, 534–547, <https://doi.org/10.1175/JHM495.1>, 2006.

Zhang, Y., Pan, M., Sheffield, J., et al. A Climate Data Record (CDR) for the global terrestrial water budget: 1984–2010, *Hydrol. Earth Syst. Sci.*, 22, 241–263, <https://doi.org/10.5194/hess-22-241-2018>, 2018.