Legend

Reviewers' comments

Authors' responses

Direct quotes from the revised manuscript

Reply to Reviewers' comments (Reviewer#3)

Reviewer #3: This is a review of "An in-situ daily dataset for benchmarking temporal variability of groundwater recharge" by P. Malakar et al. This paper describes the development of a benchmark dataset of groundwater recharge per unit specific yield (RpSy). The authors apply an established Water Table Fluctuation / Master Recession Curve method and QA/QC measures to produce daily groundwater level variation time series data for 485 sites.

As the authors note, there are currently no daily timescale data sets in the literature with which one can compare estimated/modeled results for groundwater recharge. Due to usefulness of the results, I recommend acceptance of the paper after minor revisions.

Response: We thank the reviewer for their thoughtful evaluation. We greatly appreciate the recognition of our efforts in creating a benchmark dataset. We have addressed the suggested minor revisions to further strengthen the manuscript and enhance the readability and utility of our figures and results. Thank you for your recommendation for acceptance following these adjustments.

Reviewer #3: My main request is that more analysis and discussion be given to the RpSyu data, which is the version of the data that includes time varying specific yield, and which is provided in this data set alongside the RpSy data. The authors state that because the correlation between these two data sets is greater than 0.8, they will not include RpSyu data in the comparisons with USGS data, Ppt, ET, etc. It is not surprising that RpSy and RpSyu would generally correlate with each other, but without a comparison between RpSyu data and the USGS data, a user cannot determine whether the additional complexity of the RpSyu calculation adds any value. Some analysis here would help the user decide whether to use the provided RpSy or RpSyu. At a minimum, the map of R2 between the USGS data and the RpSy data should have an equivalent map for RpSyu, and there should be some summary statistics that help readers understand whether RpSy or RpSyu more closely matches the temporal variation in recharge.

Response: We would like to thank the reviewer for highlighting the importance of comparing RpSyu and RpSy data with USGS recharge data to assess the additional complexity's value in RpSyu. Following the reviewer's suggestion, we have generated a spatial map of the correlation between RpSyu and USGS recharge and included summary statistics and distribution plots of correlations for RpSy and RpSyu. These analyses allow users to assess the utility of using RpSyu vs. RpSy. The average correlation between RpSyu and USGS recharge is comparable to that of RpSy. Overall, the results suggest that while the two datasets are closely related, differences exist. This additional information will aids users in selecting the dataset best suited for their needs.

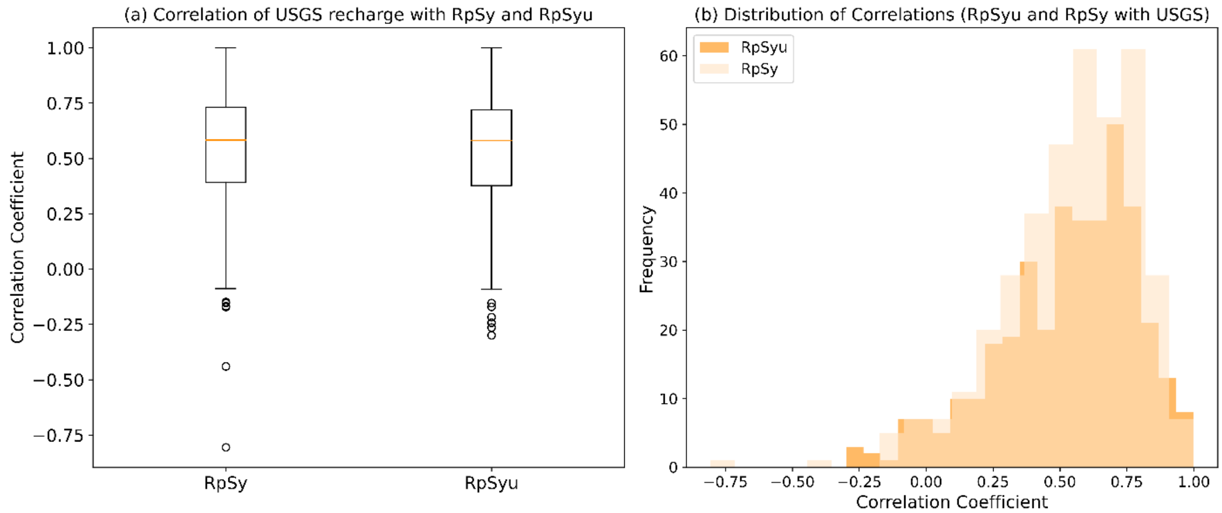

Figure S5: Comparison of RpSy and RpSyu correlations with USGS recharge data. (a) the correlations spread for both RpSy and RpSyu, with the orange line indicating the median correlation; (b) histograms show the correlation distributions for RpSy and RpSyu with USGS recharge.

We also modified and added in the text,

Additionally, we quantify RpSyu or recharge per specific yield that considers a time-varying specific yield. In instances where specific yield fluctuates temporally due to precipitation induced variations in groundwater table depth, RpSyu is expected to be more effective in capturing daily recharge fluctuations. This is particularly relevant in regions with a shallow groundwater table or in soils with fine textures, such as clayey soils, which have a large capillary fringe. In these conditions, the specific yield is significantly reduced. Part of the reason is that the capillary fringe retains water tightly thereby reducing the freely drainable portion of water. Also, when the groundwater is shallow, the unsaturated zone above the capillary fringe is either minimal or absent. As a result, the soil's ability to release water is constrained. These conditions could be common in regions experiencing large fluctuations in water table depth, such as areas with large season precipitation, intensive irrigation, or heavy groundwater pumping. However, since RpSy and RpSyu are the first of their kind to provide observational data-based recharge equivalents at a daily resolution, direct validation is not feasible. The scale mismatch between RpSyu and USGS recharge data, and inherent assumptions in the USGS product also preclude a direct one-to-one comparison. **[Page: 11, Line: 277-288]**

Table S2: Summary Statistics for Correlation between RpSyu, RpSy, and USGS recharge.

	Correlation_RpSyu	Correlation_RpSy
mean	0.531049	0.541389
standard deviation	0.256786	0.251975
minimum	-0.29906	-0.8052
5% (first quartile)	0.376744	0.391686
50% (median or second quartile)	0.58046	0.582712
75% (third quartile)	0.719417	0.730783

maximum	0.999441	0.999503

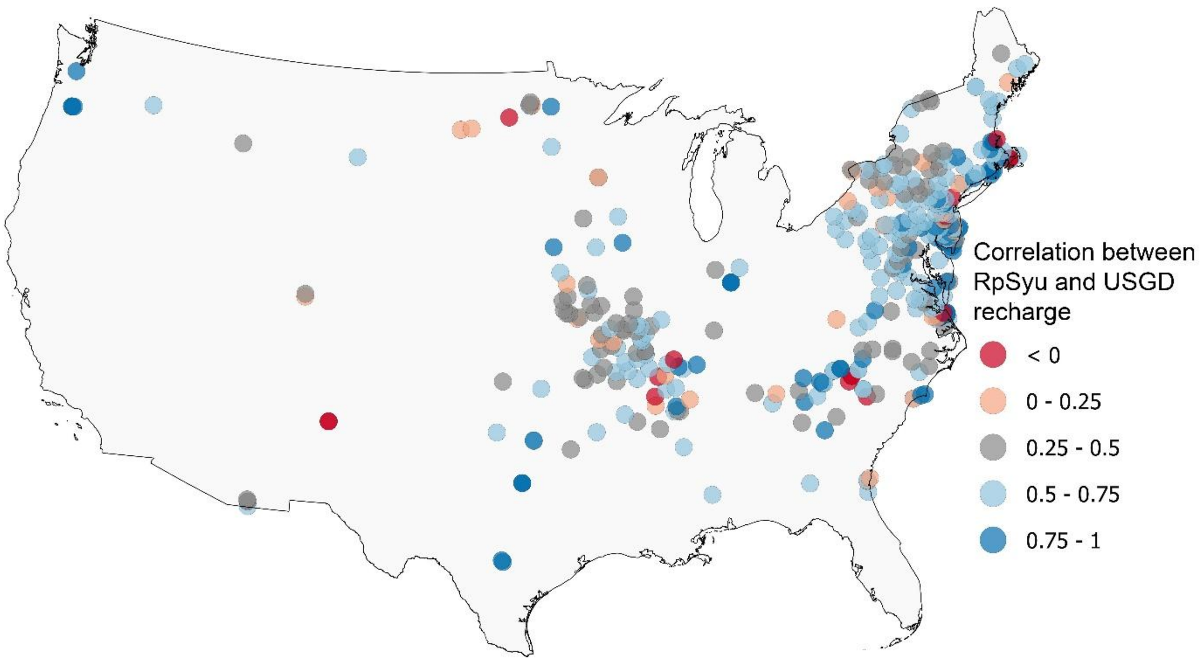

Figure S6: Spatial variation of temporal correlation between RpSyu and USGS recharge.

Reviewer #3 (Other comment 1): The interannual variation data in Figure 5 should be presented in table form, showing the R2 between the data sets (and including columns for the RpSyu data).

Response: We thank the reviewer for the comment. Following the reviewer's comment, we have added the table.

Table S3. Inter-annual variation of normalized annual recharge (normRpSy), precipitation (normPpt), and Ppt-ET (normPpt-ET).

Year	RpSy	RpSyu	Ppt	Ppt-ET	R^2 for Rpsy	R^2 for Rpsyu
1983	103.0	103.1	109.2	124.3	RpSy Vs. Ppt	RpSyu Vs. Ppt
1984	125.2	92.9	112.8	133.5	0.839	0.029
1985	79.9	108.4	86.4	62.8	RpSy Vs. Ppt-ET	RpSyu Vs. Ppt-ET
1986	84.8	103.4	85.5	69.9	0.837	0.026
1987	101.1	106.8	103.6	108.1		
1988	79.2	107.8	82.2	60.7		
1989	104.2	98.6	106.9	115.1		
1990	102.9	106.5	94.6	86.3		
1991	105.9	117.0	106.0	111.4		
1992	90.6	102.7	94.8	92.2		
1993	109.2	80.8	99.8	101.9		
1994	113.1	94.9	106.3	114.4		

1995	87.1	102.9	85.1	69.4	
1996	120.3	89.4	115.8	133.7	
1997	110.6	120.6	100.4	101.9	
1998	113.8	106.8	105.6	111.2	
1999	82.0	99.9	88.7	73.9	
2000	85.3	104.0	89.4	78.4	
2001	82.0	114.9	83.7	67.4	
2002	74.7	98.2	86.5	70.5	
2003	120.0	118.5	116.7	140.7	
2004	114.4	109.0	111.6	121.0	
2005	101.9	110.7	93.6	83.9	
2006	104.1	86.3	106.7	105.9	
2007	98.7	89.4	92.7	79.5	
2008	100.2	85.3	106.8	111.4	
2009	100.4	111.6	100.6	100.5	
2010	111.5	113.4	109.0	113.9	
2011	112.8	87.9	114.0	121.5	
2012	86.2	110.2	89.2	69.2	
2013	101.4	90.3	102.6	104.7	
2014	99.0	99.0	97.9	95.2	
2015	96.7	99.5	99.7	97.6	
2016	103.3	84.4	104.9	104.1	
2017	100.4	104.7	101.6	95.8	
2018	108.8	105.3	112.3	94.1	
2019	125.3	102.3	119.0	140.6	
2020	110.0	102.0	109.0	114.8	
2021	108.7	104.6	111.7	114.0	

We further added the corresponding figure 5 for RpSyu estimates.

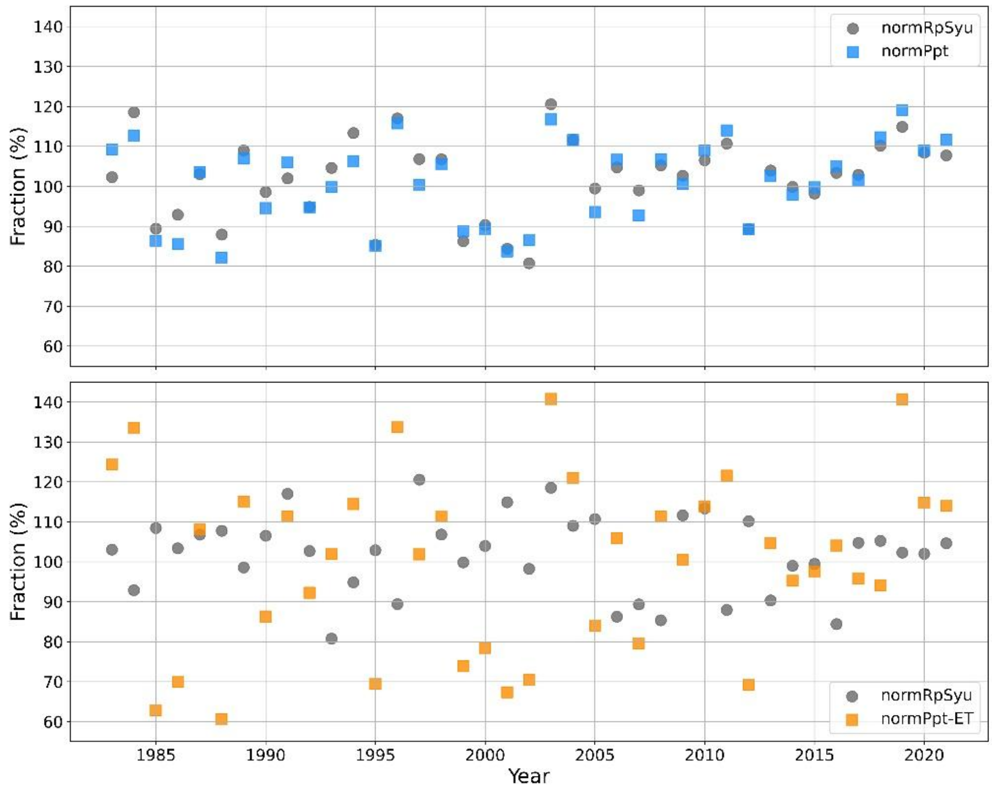

Figure S8: Inter-annual variation of normalized annual recharge (normRpSyu, shown using grey solid dots), precipitation (normP, blue squares), and Ppt-ET (normP-ET, orange squares).

Reviewer #3 (Other comment 2): Can you comment more in the discussion on the appropriate way to use these data to evaluate models? You make some mention already, but more clear statements on this point would be useful. Such as, temporal variation but not magnitude between these data and recharge estimates, what to do if you have some specific yield numbers to apply for a given area, etc.

Response: We appreciate the reviewer's suggestion to provide clearer guidance on the probable use of the RpSy dataset for model evaluation. In response, we have expanded and modified the discussion to offer more details on how researchers and practitioners can effectively utilize this dataset. We added and modified,

While the RpSy data does not offer direct recharge estimates, it still captures the variations and changes in groundwater recharge over time at daily to coarser temporal resolution. Hence, despite the limitations, uncertainty, and associated caveats discussed in section 2 and 6, the RpSy dataset can be used to validate temporal consistency of recharge estimates derived from empirical methods (Reitz et al., 2017a; Reitz and Sanford, 2019a), physically-based land surface models (Anurag and Ng, 2022; Li et al., 2021; Niraula et al., 2017) or integrated hydrologic models (Kumar and Duffy, 2015; Kollet and Maxwell, 2006; Kumar et al., 2009; Therrien et al., 2010). The RpSy dataset can be utilized for analysing the timing, frequency, and duration of recharge events. Since RpSy provides fluctuations at a daily scale, researchers can use the temporal patterns to assess whether the abovementioned models have the ability to accurately simulate groundwater recharge variability. The match or mismatch between the temporal alignment of RpSy based recharge and model based recharge outputs can provide an assessment of the model's capability to replicate the event based response to hydroclimatic forcings. Furthermore, the data may also be used to validate the functional relationship between recharge and associated factors as represented in land surface and global hydrologic models. Gnann et al. (2023) demonstrated that theoretical and empirically based functional relationships for recharge differ significantly from global water models. Even when a model produces highly accurate predictions, it may still poorly simulate the strength of functional process couplings. In other words, it may produce right results for the wrong reasons. Such models are likely to underperform during periods when the forcing characteristics are different than those in the training data. The derived benchmark RpSy data, along with forcing variables, can be used to validate the functional relationships represented in models of recharge, using one of the several diagnostic methods such as information theory (Ruddell et al., 2019), causality mapping (Barnett and Seth, 2014; Runge et al., 2019; Runge, 2018), and convergence cross mapping (Ye et al., 2015). The RpSy data may also be used to temporally downscale long-term recharge estimates from observations, thus facilitating generation of recharge inputs for groundwater models (Kim et al., 2008). In circumstances where high confidence specific yield values are available and/or obtainable from field measurements, hydrogeological surveys, or from literature, the RpSy data can be converted into recharge estimates (Recharge = RpSy × Sy). In these cases, a direct comparison can be made between the magnitude of modelled recharge and RpSy based recharge. Additionally, the data may be used for an improved understanding of the role of different forcing and antecedent hydrologic conditions on groundwater recharge and, thus helping manage groundwater aquifers under water-stress conditions. **[Page: 12, Line: 308-333]**

Anurag, H. and Ng, G. H. C.: Assessing future climate change impacts on groundwater recharge in Minnesota, J. Hydrol., 612, 128112, https://doi.org/10.1016/J.JHYDROL.2022.128112, 2022.

Barnett, L. and Seth, A. K.: The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference, J. Neurosci. Methods, 223, 50–68, https://doi.org/10.1016/j.jneumeth.2013.10.018, 2014.

Gnann, S., Reinecke, R., Stein, L., Wada, Y., Thiery, W., Müller Schmied, H., Satoh, Y., Pokhrel, Y., Ostberg, S., Koutroulis, A., Hanasaki, N., Grillakis, M., Gosling, S. N., Burek, P., Bierkens, M. F. P., and Wagener, T.: Functional relationships reveal differences in the water cycle representation of global water models, Nat. Water 2023 112, 1, 1079–1090, https://doi.org/10.1038/s44221-023-00160-y, 2023.

Kim, N. W., Chung, I. M., Won, Y. S., and Arnold, J. G.: Development and application of the integrated SWAT–MODFLOW model, J. Hydrol., 356, 1–16, https://doi.org/10.1016/J.JHYDROL.2008.02.024, 2008.

Kollet, S. J. and Maxwell, R. M.: Integrated surface-groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model, Adv. Water Resour., 29, 945–958, https://doi.org/10.1016/j.advwatres.2005.08.006, 2006.

Kumar, M. and Duffy, C. J.: Detecting hydroclimatic change using spatio-temporal analysis of time series in Colorado River Basin, J. Hydrol., 374, 1–15, https://doi.org/10.1016/j.jhydrol.2009.03.039, 2009.

Kumar, M. and Duffy, C. J.: Exploring the Role of Domain Partitioning on Efficiency of Parallel Distributed Hydrologic Model Simulations, J. Hydrogeol. Hydrol. Eng., 04, https://doi.org/10.4172/2325-9647.1000119, 2015.

Li, B., Rodell, M., Peters-Lidard, C., Erlingis, J., Kumar, S., and Mocko, D.: Groundwater recharge estimated by land surface models: An evaluation in the conterminous United States, J. Hydrometeorol., 22, 499–522, https://doi.org/10.1175/JHM-D-20-0130.1, 2021.

Niraula, R., Meixner, T., Ajami, H., Rodell, M., Gochis, D., and Castro, C. L.: Comparing potential recharge estimates from three Land Surface Models across the western US, J. Hydrol., 545, 410–423, https://doi.org/10.1016/j.jhydrol.2016.12.028, 2017.

Modern monthly effective recharge maps for the conterminous U.S., 2003-2015 | USGS Science Data Catalog: https://data.usgs.gov/datacatalog/data/USGS:5cd0a1b1e4b09b8c0b79a51c, last access: 13 May 2022.

Reitz, M. and Sanford, W. E.: Estimating quick-flow runoff at the monthly timescale for the conterminous United States, J. Hydrol., 573, 841–854, https://doi.org/10.1016/j.jhydrol.2019.04.010, 2019.

Reitz, M., Sanford, W. E., Senay, G. B., and Cazenas, J.: Annual Estimates of Recharge, Quick-Flow Runoff, and Evapotranspiration for the Contiguous U.S. Using Empirical Regression Equations, J. Am. Water Resour. Assoc., 53, 961–983, https://doi.org/10.1111/1752-1688.12546, 2017.

Ruddell, B. L., Drewry, D. T., and Nearing, G. S.: Information Theory for Model Diagnostics: Structural Error is Indicated by Trade-Off Between Functional and Predictive Performance, Water Resour. Res., 55, 6534–6554, https://doi.org/10.1029/2018WR023692, 2019.

Runge, J.: Causal network reconstruction from time series: From theoretical assumptions to practical estimation, Chaos, 28, https://doi.org/10.1063/1.5025050, 2018.

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., Zhang, K., and Zscheischler, J.: Inferring causation from time series in Earth system sciences, Nat. Commun. 2019 101, 10, 1–13, https://doi.org/10.1038/s41467-019-10105-3, 2019.

Therrien, R., McLaren, R. G. G., Sudicky, E. A. A., and Panday, S. M. M.: HydroGeoSphere: a three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute transport, Groundw. Simulations Group, Univ. Waterloo, Waterloo, 322p., 2010.

Ye, H., Deyle, E. R., Gilarranz, L. J., and Sugihara, G.: Distinguishing time-delayed causal interactions using convergent cross mapping, Sci. Rep., 5, https://doi.org/10.1038/srep14750, 2015.

Reviewer #3 (Other comment 3): Figure 4a and 4c are identical – must be some error in the figure production.

Response: We are grateful to the reviewer for pointing this out. Upon review, it was found that we inadvertently made a mistake during the production of the combined figure, by pasting identical figures in panels 4a and 4c. We have corrected the figure.

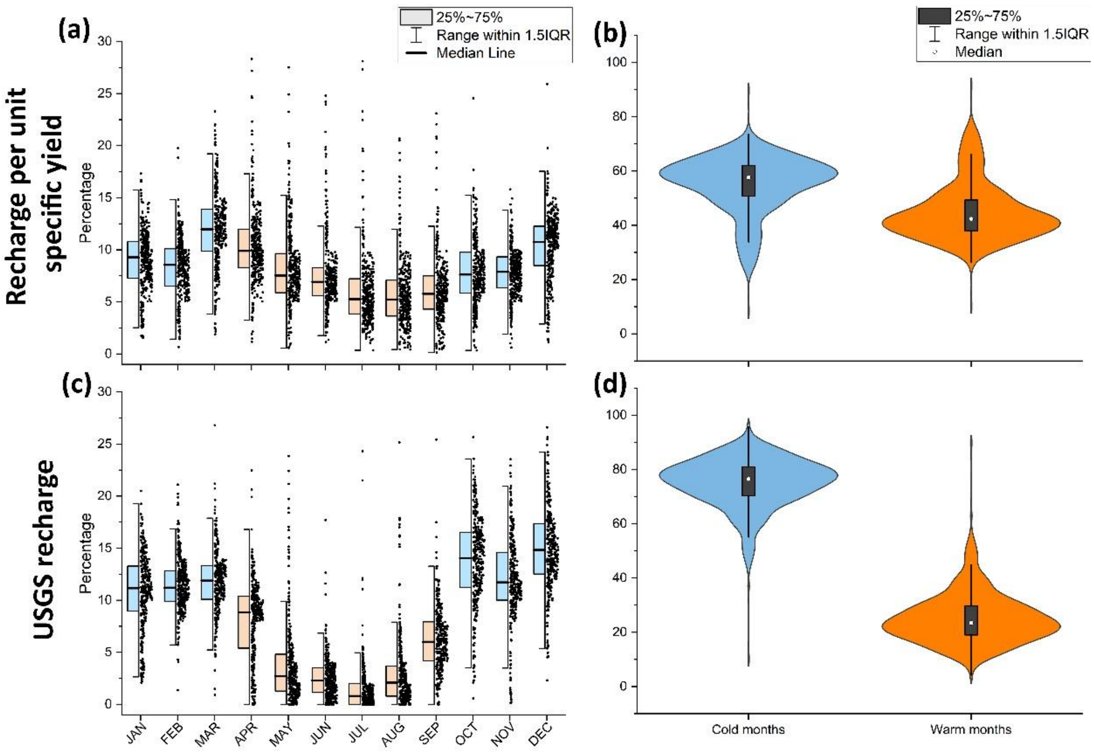

Figure 4: Fraction of recharge in different months and seasons (i.e., Cold seasons (Oct to Mar), Warm-season (Apr to Sept)) relative to the total recharge(/equivalents) for RpSy (top, a and b) and USGS (bottom, c, and d) recharge products. In this plot, USGS recharge data for the grids with RpSy estimates are used. IQR indicates the interquartile range.

Reviewer #3 The paper has a few minor/grammatical errors and could use a proofreading. A few examples:

L25: strike "the rate of"
L36: "in East Africa is" to "in East Africa are"
L36: "ratio" to "fraction"

Response: We appreciate the reviewer's careful review and attention to detail. We have thoroughly proofread the manuscript and corrected minor and grammatical issues, including the specific changes identified (e.g., removing "the rate of" in L25, changing "in East Africa is" to "in East Africa are," and using "fraction" instead of "ratio" in L36). Thank you for taking the time to point out these errors; your feedback has helped us improve the clarity and overall quality of the manuscript.