Interactive Discussion: Author Response to Referee #1

Groundwater recharge in Brandenburg is declining – but why?

T. Francke, and M. Heistermann
NHESS Discussions, doi:10.5194/egusphere-2025-222

RC: ***Reviewer Comment***, AR: *Author Response*, ☐ Manuscript text

Dear madam or sir,

thank you very much for your referee report, and for the time and effort you spent to examine the manuscript. In particular, we appreciate the detailed suggestions along the entire manuscript text.

Similar to referee 2, you specifically mention the overall impression that parts of the manuscript were difficult to follow. We will hence attempt, in a revised version, to particularly improve the comprehensibility.

Please find both your comments and our responses below in a point-by-point reply.

Thanks again for your feedback, and your support in improving the manuscript.

Kind regards,
Till Francke and Maik Heistermann

Point-by-point replies

RC: ***This is a very interesting paper that analyses the groundwater recharge behavior in 5 catchments in Brandenburg. The paper fits very well within the scope of the special issue.***

AR: Thank you for the positive reception of the manuscript.

RC: ***I think the paper has great potential, but it is somewhat hindered by missing information and lacking methodology descriptions. This makes the paper somewhat difficult to follow, I often had to backtrack quite a lot in the text to see if I missed some detail or information. Therefore I am sending below my recommendations to improve the manuscript.***

AR: We will revise the manuscript in order to improve readability and comprehensibility. To that end, the comments of both referees are very helpful.

RC: ***My main question is about the model setup. Are the models set up on a per catchment basis, or over a grid or other spatial sub-units? This was very confusing to me – sometimes I was sure everything is aggregated over catchments, but in some parts further inhomogeneities were resolved. It would be important to include more details about this in the text. Please find my specific comments below.***

AR: A similar comment was made by referee #2. We suggest to add a sub-section at the beginning of the methods section in order to give an overview of the modelling approach. Specifically, this sub-section will clarify how the spatial average of the simulated GWR is obtained per catchment, based on the hydrotope concept.In the following, we provide the draft for this new subsection 3.1:

> The main methodological approach of our study is to simulate series of GWR in each study catchment for the period from 1980 to 2023. From these series, we computed the trends (section 3.6) in GWR and compared them to the corresponding trends of observed discharge. A good agreement between both would indicate that the simulation model is able to explain the discharge trends as outlined in Fig. 1 and Tab. 1 (see also section 1 for additional background).
>
> To obtain GWR, the one-dimensional Soil-Water-Atmosphere-Plant (SWAP, van Dam et al., 2008, section 3.4) was used to simulate the surface water balance and the resulting percolation of water through the unsaturated zone down to the groundwater table. This daily "bottom flux" was aggregated to obtain the annual sum of groundwater recharge.
>
> In order to obtain the spatial average of GWR per catchment, we followed the concept of "hydrotopes", i.e. spatial sub-units that are considered as homogeneous with regard to (i) climate forcing, (ii) soil texture, (iii) land use, and (iv) groundwater depth. For each catchment, climate and soil were assumed to be uniform across the entire catchment (i.e. *one* class each per catchment): the climate forcing was based on the nearest of the four selected climate stations (section 2.3), and soil texture was represented by the dominant soil texture class in the catchment (section 2.6). Land use and groundwater depth, however, were assumed to be heterogeneous across the catchment: land use was represented by *two* classes (forest and grass-/cropland) and groundwater depth by *13* classes of groundwater depth (see section 2.6). This resulted into a total of 26 hydrotope classes (1 climate x 1 soil x 2 land uses x 13 groundwater depths). By spatially intersecting all four layers, we quantified the areal fraction of each hydrotope class per catchment. Running the SWAP model for each of the 26 hydrotope classes, the daily GWR per catchment was then obtained as the area-weighted average of the simulated daily bottom flux per hydrotope.
>
> In the following subsections, we further explain the treatment of missing hydro-climatological data (section 3.2), the precipitation correction (section 3.3), the SWAP model and its parameterisation (section 3.4), the specific design of the simulation experiments (section 3.5) and the calculation of trends (section 3.6).

RC: ***L33: These few paragraphs could use some references to support the text.***

AR: We would like to point out that the corresponding paragraph ends with the statement "And, in fact, there is strong evidence for such a decline." This statement essentially sets the scene for the two subsequent paragraphs which introduce two types of evidence including references (first, the decline in groundwater levels, second the decline of discharge). This should be kept in mind because otherwise the above statement would in fact appear to lack support by reference. We emphasize the link to the subsequent paragraphs by adding a colon at the end of the statement instead of a full stop.

Having said that, we agree that the paragraph ll. 33-35 could, in itself, be better supported references. For this purpose, we selected two additional references, so that the entire paragraph will become:

> Any long-term shift of Brandenburg's already unfavourable vertical water balance towards lower rates of net groundwater recharge (GWR) would put water resource management in Brandenburg at risk – including the German capital Berlin in its centre (Somogyvári et al. 2024; Pohle et al. 2025; Somogyvári et al. 2025). And, in fact, there is strong evidence for such a decline: first, [...]

RC: ***L36: groundwater***

AR: Will be fixed throughout the text.

RC: ***L42: Does the term "combination" here means a gauging station after multiple rivers join together?***

AR: As already pointed out in l. 42 of the preprint, we would like to refer to section 2.2 for further explanation (in order to keep the introduction concise). In section 2.2, we will enhance the original explanation (ll. 136 ff.) in order to improve the comprehensibility (also based on a comment by referee #2). The corresponding sentences in section 2.2 will hence become:

> In three cases, adjacent catchments were aggregated in order to further reduce the effect of uncertain belowground watersheds, and to increase the area over which the water balance is computed. "Aggregating" two (or three) catchments means that their areas were merged to one coherent area, and that the observed runoff at the gauges of each catchment was summed up in order to obtain the total runoff from the merged area. Tab. 1 gives a comprehensive overview of the used discharge gauges and the aggregation of catchment areas and runoff gauge observations for our analysis.

RC: ***L50: can you provide a reference?***

AR: The preprint already contains, in line 66, the corresponding reference which establishes that surface runoff is negligible in Brandenburg (LUA Brandenburg, 2001). In order to improve the readability of the manuscript, and avoid the need to jump back and forth while reading, we will also add the corresponding reference after the statement in ll. 50-52 of the preprint.

RC: ***L52: what about cross-flow with lower aquifer layers? The recent study of Tsypin et al., 2024 showed that deeper layers interact with the uppermost at certain geological settings, mainly where the Rupelian clay layer is eroded. These flows could act as a water source or sink as they flow upwards or downwards at different locations.***

AR: Thanks for this interesting reference. However, it is difficult for us to assess the potential role of such flows between aquifer layers with regard to long term trends. The authors of the study themselves do not explicitly associate the fluxes between pre-Rupelian and Quaternary aquifers with any long-term trends in groundwater levels. Instead, they hypothesize that these trends "can be explained by (a) changes in water balance; (b) anthropogenic factors; (c) memory effects in the unsaturated zone." In our view, the present manuscript is not a suitable venue to speculate about the relevance of such local cross-fluxes for the water balance of the Quaternary aquifers.

RC: ***Fig. 1: Which wells were used to create this map? Is the data coverage homogeneous, or are there any areas with fewer information?***

AR: As pointed out in the figure caption, the map was adopted from Landesamt für Umwelt Brandenburg (2022). That reference does not provide further information to answer the question. In general, however, there are less gauges for very deep groundwater tables.

RC: ***This figure would be a better place to show the used gauge locations than fig. 5.***

AR: We are not entirely sure what the referee means by ”gauge locations”, but assume that the locations of the climate stations are meant. Still, we are confused because Fig. 5 only contains a small inset map for support the interpretation of the other figure panels while the locations of the climate stations have already been introduced in Fig. 2. In any case, we would prefer not to introduce additional elements in Fig. 1a because it is only intended to give a spatial overview of groundwater level trends as adopted from Landesamt für Umwelt Brandenburg (2022).

RC: ***L67: ”any long term trend in discharge can thus be interpreted as a long term trend in GWR.” Can you elaborate on this a bit further: what is the long-term criteria, why short term changes cannot be used in such way? Is there a model that can show this behavior/or a past study where this was investigated?***

AR: We are a bit confused by this comment since the criteria (multiple decades) and also the reasoning between the use of long-term trends are elaborated in the subsequent sentence:

> Any *long-term* trend in discharge can thus be interpreted as a long-term trend in GWR. The emphasis is on long-term (in the sense of multiple decades) because any short-term dynamics, e.g., across events or seasons, might differ substantially.

In our view, this statement does not require support by modelling results or references, but maybe we could explain the statement a little bit better by modifying and extending the two sentence as following:

> Any *long-term* trend in discharge can thus be interpreted as a long-term trend in GWR. The emphasis is on long-term, in the sense of multiple decades, since we can assume that over long time periods, any GWR will also end up in the surface water bodies that drain the system. Short-term dynamics of river discharge (e.g., across rainfall events or seasons) are, however, governed by the travel times of water along the different transit paths in the vadose zone and the aquifer.

RC: ***L108: reference to the dataset***

AR: The reference to the dataset, including some background, was provided in section 2.5 in the overall "Data" section. To improve readability, we will already introduce the reference after ”we use a recently published LAI dataset”, corresponding to l. 107 of the preprint.

RC: ***L125: reference needed***

AR: We will add the corresponding reference (LBGR, 2024).

RC: ***L129: Can you support this statement with a reference, or by referring back to the introduction***

AR: There is already a reference to the introduction at the end of the sentence.

RC: ***L136: What does partly mean here?***

AR: Based on this comment and a comment by referee #2, we understand that the explanation should be improved, so we will change the corresponding sentence as follows:

In three cases, adjacent catchments were aggregated in order to further reduce the effect of uncertain belowground watersheds, and to increase the area over which the water balance is computed. "Aggregating" two (or three) catchments means that their areas were merged to one coherent area, and that the observed runoff at the gauges of each catchment was summed up in order to obtain the total runoff from the merged area. Tab. 1 gives a comprehensive overview of the used discharge gauges and the aggregation of catchment areas and runoff gauge observations for our analysis.

RC: ***L143: I am just curious here: what about not interpolated, but dynamically downscaled datasets, such as CERv2 (https://www.tu.berlin/klima/forschung/regionalklimatologie/mitteleuropa/cer) . Would they not be a better choice for gridded data?***

AR: The mentioned dataset is certainly interesting, and the big advantage of such a dataset would be that it is physically consistent in space and time. Furthermore, Bart et al. (2024) in fact demonstrated considerable improvements in comparison to CERv1, particularly with regard to precipitation (with a mean deviation of 0.1 $\mathrm{mm\,d^{-1}}$ from DWD stations). Yet, to our knowledge, it has not yet been systematically assessed how well CERv2 captures the long-term trends in the hydro-climatological variables that are key to the assessment of GWR trends (solar irradiation, temperature, humidity, wind speed, precipitation). Apart from these uncertainties, CERv2 is not suitable for our study because it only starts at 1980 while we require a long and consistent warm-up starting in 1951.

While we consider such a discussion interesting, we think that it will lead too far away from the scope of our study, so we would prefer not to go into such depths in the context of the revised manuscript.

RC: ***L161: parenthesis missing***

AR: We will remove the superfluous closing parenthesis.

RC: ***Were there any weighting used when multiple gauges were included? I think using a buffer area as a criteria here could easily introduce further uncertainties to the data (as precipitation in Brandenburg could be very heterogeneous, especially during extreme rainfall events) – does including them make a big difference? (does it worth it to use them instead of just the climate station)***

AR: The rain gauge observations are not weighted. The observed series of each selected rain gauge was used individually to force the model. Any weighting would correspond to an interpolation, which we intend to avoid. We are fully aware that, for individual events, single gauges are not able to reproduce the mean areal precipitation of a catchment, specifically in case of convective heavy rainfall during the summer months. Yet, for our analysis, individual precipitation events are not the focus, as we focus on long-term behaviour. Instead, it is more important to use homogeneous forcing series. And while precipitation can substantially vary in space for individual events, the spatial variability of average (seasonal) precipitation is much lower.

In the revised manuscript, we will modify the respective line (l. 82 of the preprint) in section "Design of simulation experiment" to

- ...
- precipitation observed at other precipitation gauges (individually, i.e. without any further weighting) in or close to the study catchments (ref. section 4)
- ...

RC: ***L196: The used dataset is a categorized dataset for groundwater distance. Does its non-continuous nature create any issues for your analysis?***

AR: It is true that the dataset is categorical data. Yet, for the definition of hydrotopes, we would have to create categorical data anyway since each hydrotope is defined by its membership in distinct classes of the relevant variables (climate station, soil texture, land use, depth to groundwater). We think that the definition of categories is reasonable (meter resolution for shallow groundwater, wider class boundaries for deep groundwater). The resulting uncertainties should, on average, not cause any systematic errors: for shallow groundwater tables, the groundwater depth has a strong effect on evapotranspiration (due to the interaction of surface and groundwater); however, for shallow groundwater, the groundwater depth data is sufficiently resolved (1 m increments down to a depth of 5 m). For larger depths, the groundwater depth only affects the transit time of percolating water through the unsaturated zone. Using classes here will not cause a systematic error. In the revised manuscript, we will explain this issue in a bit more detail after l. 199 of the preprint.

RC: ***General question for methods: Was the analysis done over a grid, over catchment or any other spatial unit? It is not clear for me from the text. Also, for a better comprehensibility, this section could really use a flowchart on how the different parameters interact, or even a conceptual figure on how the modeling concept looks like.***

AR: As pointed out above, we will introduce a new sub-section at the beginning of the methods section in which we will outline the overall methodological approach, with specific focus on the spatial model set-up.

RC: ***L214: I really like this concept for filling in missing data. Can you give more details on how this package was used (as it is a specific module for photovoltaic modelling, it is not straightforward how to use it in this setting). Which function was used for modelling with what parameters?***

AR: Beginning in l. 212, we essentially report all relevant information with regard to model training. The fact that the package is geared towards applications in the energy sector is irrelevant as to its application for retrieving clear-sky radiation for any specific date. Clear-sky radiation was modelled for the training location (Potsdam) and for the application locations (other gauges) based on geographic coordinates, altitude, and datetime at 20 minutes intervals. The results were then aggregated to daily values. The random forest model was using 100 trees with a maximum node depth of 10. We will add the additional information in the revised version of the manuscript as follows:

> [...] we trained a random forest (100 trees and a maximum node depth of 10) to model solar irradiance from the following predictive features: clear-sky radiation as simulated from geographic coordinates, altitude, and datetime by the Python package pvlib (Holmgren et al., 2018; Anderson et al., 2024; using the function ”get_clearsky” and aggregating from 20 min resolution to a daily sums of radiation) [...]

RC: ***L228: same question: what random forest model was used and how?***

AR: We assume the question refers to L258 of the original manuscript. The random forest model was using 100 trees with a maximum node depth of 10. As predictive features, we used the local daily ERA5-Land wind speed estimates and the wind speed observed at the climate station in Potsdam. The missing information will be added to the section.

RC: ***L266: How did you check that this spin up period was enough for the model?***

AR: The required spin-up time corresponds to the time required for any signal (in terms of a flux) to reach the groundwater table for the initial condition being set to a hydraulic equilibrium along the entire profile. The required transit time is longest for forest locations (lowest GWR) and, evidently, for the hydrotopes with the deepest groundwater (up to 70 m).

RC: ***L270: was the modelling run on a grid? – I am really confused about this at this point***

AR: Please see our response above: We will add a sub-section at the beginning of the methods section to outline the overall approach.

RC: ***L282: why was this setup needed?***

AR: The background is explained in section 2.4 (ll. 153-156). To make this clear, we had already added a reference to section 2.4 at the end of l. 282 in the preprint.

RC: ***Table 3: this table is a bit complicated and difficult to understand. Please consider revising it.***

AR: We agree that the table is a bit bulky. In our view, though, it is the most straightforward way to document the different model realisations in a transparent way. In order to make the table more comprehensible, we will revise the table caption, adding more explanation on how to read the table.

RC: ***L320: maybe: "(small triangles)"***

AR: We will change "triangles" to "small triangles".

RC: ***Figure 4: can you add a legend instead of explaining everything in the captions?***

AR: We deliberately chose to explain the meaning of the symbols in the caption because a corresponding legend would have to been too text-heavy.

RC: ***L318: are the blue/red lines are simply drawn by connecting the points of the corresponding LAI setup?***

AR: Yes, that is correct.

RC: ***L327: This could also be referred to as a linear relation between precipitation and GWR***

AR: We would prefer to not explicitly frame the relationship as "linear" (see ll. 370-71). We understand that l. 327 is a little bit unwieldy, yet we chose the wording deliberately to make it more unambiguous.

RC: ***L338: Can you explain further how the significance was calculated?***

AR: We will add the following explanation to section "Trend calculation":

> [...] The significance of the trends is likewise computed according to Sen (1968), and expressed as the p-value.

RC: ***Figure 5: This figure is very complex, but I think the chosen visualizations are really the best way to show the findings. You could make the figure a bit less busy by removing the gauge locations from the map. Also positioning the legend next to the map is not ideal.***

AR: In the revised manuscript, we will remove the gauge locations from the map, as suggested. We will also switch the positions of map and legend which, in our opinion, further improves the structure of the figure.

RC: ***Could you consider adding another figure where (some of – maybe just the main configuration) the modelled GWR timeseries are shown against the discharge timeseries for the catchments? It would help***

the text in my opinion…

AR: We will add the requested figure to the results section "Trends in GWR", and a brief description of the figure.

RC: ***L351: rephrase "While this is plausible…"***

AR: We will rephrase as follows:

> This is plausible. Still, trends from HYRAS-DE-PR need to be considered with care (as repeatedly mentioned before) because they are based on temporally varying sets of gauges [...]

RC: ***L361: what are the blue and red lines, are they trendlines or just connectors between the points?***

AR: The lines simply are connectors between the grey circles, showing which circles are the result of simulations with or without an LAI trend. They are intended to help the visual differentiation of these two groups, which is otherwise difficult, as the colour of the dots is already used for encoding their significance. We had also tested adding a coloured outline to the dots, but this proved to be much less legible.

RC: ***L365: which offset? Could you be more exact?***

AR: As "offset", we refer to the ordinate value at an abscissa value of 0. We will add this information in the revised manuscript in parenthesis add change the wording to "intercept" to make this clear.

RC: ***Can one state here that there is an unknown water loss from the system as the discharge trend is steeper?***

AR: This would be one explanation. Uncertainty in the precipitation would be another. Overall, we would like to reserve section 5 for the discussion of such hypotheses.

RC: ***L370: Does it mean more humid conditions or a faster change towards humid conditions?***

AR: Both. The Stepenitz-Löcknitz catchment is an example for a more humid catchment where the slope of the lines is the steepest among all five study catchments. However, for the remaining catchments (most pronounced for the Ucker-Welse), we observe that the lines become a little steeper for higher positive precipitation trends which effectively implies an increase of humidity over time.

RC: ***L373: point out whiskers on the legend***

AR: We would prefer to explain the whiskers in the figure caption as it is difficult to unambiguously distinguish them in the legend from the lines denoting the two simulation groups (with/without LAI trend, see also reply to comment L361 above).

RC: ***L380: where was this pointed out?***

AR: This was pointed out in section 1. In the revised version, we will add a corresponding cross-reference.

RC: ***How could the model have GWR estimations for different depths – was it over a grid?***

AR: Please see our response above: we will add a brief overview of the modelling concept at the beginning of the methods section, which will also serve to clarify this issue.

RC: ***5 limitation and uncertainties: I really like this section, but it could have a more exact title like "Explaining the gap between trends"***

AR: While the section has a focus on explaining the gap between observed discharge trends and simulated GWR trends, it is not entirely limited to this issue. Instead, we would like to cover the full range of potential error sources included in the data and modelling chain and their interpretation. We would hence prefer to keep the section title more generic.

RC: ***L437: a similar behavior was pointed out locally by Somogyvari et al., 2024 for a lake system in the region. They also used a combination of different factors as an explanation for system water loss, together with an environmental tipping point. Did you consider such explanations?***

AR: In our study, we only accounted for the effect of climatic drivers and trends in LAI, while other factors are only touched in the discussion. Still, it is encouraging to see that the results from the two studies are consistent as to the effect of vegetation (in terms of NDVI or LAI). We suggest to add a reference to Somogyvári et al. (2024) in the corresponding discussion.

RC: ***L511: Could you rank the different explanations based on plausibility?***

AR: We understand this comment in reference to all sub-sections of section 5. As of now, we would prefer not to provide an overall assessment of relative importance, since governing processes might also vary regionally while other mentioned processes are not yet fully understood. In the different sub-sections, however, we already stated in case we considered a certain explanation to be less relevant (e.g., section 5.4 with regard to water consumption by irrigation). However, we have no rigorous way of ranking them.

References

Bart, F., Schmidt, B., Wang, X., Holtmann, A., Meier, F., Otto, M., Scherer, D.: The Central Europe Refined analysis version 2 (CER v2): evaluating three decades of high-resolution precipitation data for the Berlin-Brandenburg metropolitan region, Met. Z., 33(5), 339-363, 2025.

Landesamt für Umwelt Brandenburg: Wasserversorgungsplanung Brandenburg - Sachlicher Teilabschnitt mengenmäßige Grundwasserbewirtschaftung, `https://lfu.brandenburg.de/sixcms/media.php/9/Wasserversorgungsplan_barrierefrei.pdf`, 2022.

LBGR: Landesamt für Bergbau, Geologie und Rohstoffe Brandenburg, Bodenübersichtskarte des Landes Brandenburg im Maßstab 1:300,000, `https://geo.brandenburg.de/?page=LBGR-Webservices`, last accessed: 30 April 2024, 2024.

LUA Brandenburg: Flächendeckende Modellierung von Wasserhaushaltsgrößen für das Land Brandenburg - Studien und Tagungsberichte, Band 27 | Startseite | LfU, Tech. Rep. Band 27, Landesumweltamt Brandenburg, `https://lfu.brandenburg.de/lfu/de/ueber-uns/veroeffentlichungen/detail/~15-10-2001-flaechendeckende-modellierung-von-wasserhaushaltsgroessen-fuer-das-land-brandenburg-stu#`, 2001.

Pohle, I., Zeilfelder, S., Birner, J., and Creutzfeldt, B.: The 2018–2023 drought in Berlin: impacts and analysis of the perspective of water resources management, Nat. Hazards Earth Syst. Sci., 25, 1293–1313, https://doi.org/10.5194/nhess-25-1293-2025, 2025.

Somogyvári, M., Scherer, D., Bart, F., Fehrenbach, U., Okujeni, A., and Krueger, T.: A hybrid data-driven approach to analyze the drivers of lake level dynamics, Hydrol. Earth Syst. Sci., 28, 4331–4348, https://doi.org/10.5194/hess-28-4331-2024, 2024.

Somogyvári, M., Brill, F., Tsypin, M., Rihm, L., and Krueger, T.: Regional-scale groundwater analysis with dimensionality reduction, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-4031, 2025.

Tsypin, M., et al.: Modeling the influence of climate on groundwater flow and heat regime in Brandenburg (Germany)." Frontiers in Water 6 (2024): 1353394, 2024.