Response to reviewers

Review "Frequency domain water table fluctuations reveal recharge in fractured aquifers depends on both intense and seasonal rainfall and unsaturated zone thickness"
by Guillaumot et al.

Reviewer 1

I enjoyed reading this paper. It was a pleasure to see that the famous Ploemeur site continues to contribute fundamental insights! The approach is both novel and thoughtful. The authors do an excellent job of presenting their findings and also addressing potential limiting assumptions that underly the analyses. I am convinced that they have demonstrated that their approach offers a useful way to understand how precipitation and evapotranspiration are modified within the vadose zone to produce recharge. My only suggestion is that the authors may help readers to see the underlying approach and to understand the significance of the work with a couple of schematic diagrams and some added background explanations. For example, given that the method is heavily reliant on frequency domain analyses, it may be useful to show the power spectrum of recharge at the two sites and to describe it in terms tha a general reader can appreciate. Later, this could help to explain, again at a more intuitive level, how the method filters the effects of pumping and leads to the conclusions regarding the contribution of P-ET to recharge at different frequencies. I would also have appreciated a paragraph to explain the coherence and transfer function results. Can the authors help the average reader to make sense of these so that they can appreciate the results that follow? Finally, I would recommend that the authors provide some thoughts on the applicability of the method to a broader range of sites. Is it, for any reason, limited to fractured rock settings? Does it require a highly instrumented site with a long record? What practical benefits might other researchers and practitioners realize if they apply this approach at their site? All of this is simply aimed at broadening the impact and readership of the work ... the underlying science was a pleasure to read!
Dear reviewer,
Thank you for your positive feedback. We addressed your different comments below. Our answers are in blue, sentences from the manuscript are in italic.
Sincerely,
Luca Guillaumot, on behalf of all co-authors.

My only suggestion is that the authors may help readers to see the underlying approach and to understand the significance of the work with a couple of schematic diagrams and some added background explanations. For example, given that the method is heavily reliant on frequency domain analyses, it may be useful to show the power spectrum of recharge at the two sites and to describe it in terms that a general reader can appreciate.
Thank you for this suggestion. We plotted the power spectrum of recharge (see figure below) obtained from soil models and from groundwater levels in Ploemeur and Guidel. First, we observe a similar pattern between the three signals: (1) there is a peak at T = 365 days, (2) recharge amplitude increases with the time period (T). In addition, we observe that the

amplitude is higher at high frequency in Guidel compared to Ploemeur. We added this figure along with the previous description to the Supplementary file. However, we think this figure is a bit redundant with Figure 9 showing the transfer function response (recharge/(P-PET)) in the frequency domain.

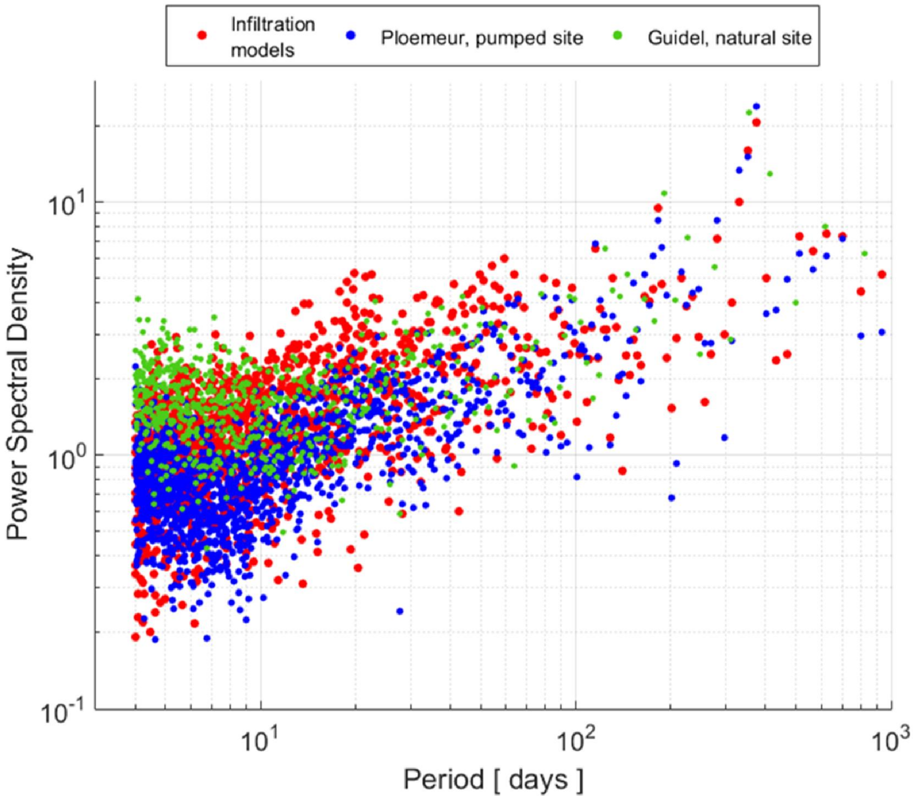


Figure 1: Power spectrum of recharge estimated from infiltration models and from groundwater levels fluctuations in Ploemeur and Guidel sites.

As suggested, we added more explanations concerning the frequency domain analyses to help readers (see next comment).

I would also have appreciated a paragraph to explain the coherence and transfer function results. Can the authors help the average reader to make sense of these so that they can appreciate the results that follow?

Coherence and transfer function are introduced in section 2.6 “How unsaturated zone transforms precipitation into recharge”. Then, associated results are presented in section 5.3 “The unsaturated zone and recharge fluxes”. Finally, these results are discussed in sections 6.2.1 and 6.3.1.

Section 2.6 (now line 206-228) was corrected in the revised manuscript to both describe and « popularize » equation 7 and 8. That is why we added some explanations about frequency domain analysis in this part. We also moved several sentences from Discussion to this section following remarks from reviewer 2:

“*The classical approach consists in defining statistics on the distribution and intensity (e.g. number of days without rain, cumulative sorted rainfall), but does not often yield satisfactory results. Considering the relationship between P−PET and recharge as a frequency dependent function is a simple but effective way to consider the impact of rainfall distribution on recharge. This function defines the relative efficiency to generate recharge between a single rainfall*

event and a long-lasting wet season, and has been recently tested by Schuite et al. (2019)." (l 209-214)
"*These transfer functions comparing flux coming in vs out of the unsaturated zone allow to infer its role in the recharge dynamics. Switching to the frequency domain offers the additional advantage to visualize how precipitation is converted into recharge at each frequency.*" (l 224-226)

Following your remark, and because Figure 9 constitutes a main result of our study, we better explained coherence and transfer function results (section 5.3) in the revised manuscript:
"*On figure 9, the coherence and transfer functions (Eq. 7 and 8) between P−PET fluctuations and RF inform on the efficiency of the transformation of rainfall events into recharge. These functions therefore illustrate the unsaturated zone response to rainfall in frequency domain. From Figure 9, results can be summarized as follows: recharge estimated from soil models and recharge estimated from WTF have similar long-term behavior, recharge estimated from soil models is too sensitive to rainfall at short-term, recharge estimated from WTF is more sensitive to short-term events on the natural site compared to the pumped site*." (l 402-407)

In addition, Figure 9 was slightly modified to add "*Recharge*" and "*P-PET*" on *y* axis. Then, more explanations and discussions are given in section 5.3, 6.2.1 and 6.3.1.

Later, this could help to explain, again at a more intuitive level, how the method filters the effects of pumping and leads to the conclusions regarding the contribution of P-ET to recharge at different frequencies.
Pumping fluctuations are already included as a boundary condition of the model (in *x=0*) so that the effect of pumping is taken into account (see line 104-107 and 253-254 in the previous manuscript). Pumping is equal to "Q" in equations 2 and 5. We improved the manuscript to highlight this point.
First, we provided more steps when developing the analytical groundwater model in Appendix A2, so that the pumping boundary condition appears clearly now.
We also mentioned explicitly this point in section '2.2 Defining the 1D flow model for each field site' (one part of these sentences were mentioned later in the previous version):
"*Here, we consider a field site (Ploemeur) where pumping wells are concentrated within a "pumping zone" constituting the outlet of the system. Therefore, the propagation of the pumping through the aquifer is simulated in a physical way.*" (l 115-117)

Finally, we added in section '3.2 Models setting up':
"*At x = 0, we impose transient pumping rates based on recorded pumping data.*" (l 267-268)

In addition, we added *Q(t)* on figure 1.

Finally, I would recommend that the authors provide some thoughts on the applicability of the method to a broader range of sites. Is it, for any reason, limited to fractured rock settings? Does it require a highly instrumented site with a long record? What practical benefits might other researchers and practitioners realize if they apply this approach at their site? All of this is simply aimed at broadening the impact and readership of the work ... the underlying science was a pleasure to read!
Thank you for this comment. The applicability of the method is an important point. We are convinced that the method can be employed to other regions with different aquifer complexity

provided that aquifer response can be approximated by a Dupuit equation. It would be very interesting to test it in karstic system where heterogeneity is more pronounced and Dupuit equation more critical. Due to the fixed 1D model structure, a critical aspect is that the method will be not relevant for boreholes located in areas where the water table pattern changes across seasons (see line 460-470 in the previous manuscript).

The method requires long-term water table records and a first guess of weekly or monthly recharge rates which can be roughly estimated from precipitation and temperature data (line 493-497 in the previous manuscript). In order to inverse recharge, it is more suitable to use high frequency water table records (daily). Finally, benefits are twice: estimating aquifer-scale characteristic time from a single point and estimating recharge (then its relationship with other variables as we did it by comparing recharge to P-PET on two sites).
Following your comment, the last part of the abstract was revised as follows:
"*In spite of the heterogeneous nature of aquifers, parameters controlling WTF can be inferred from WTF time series making confident that the method can be deployed in different geological contexts where long-term water table records are available.*" (l 15-18)

Moreover, the last sentences of the conclusion are now:
"*In this study, the method is applied in crystalline contexts that display fractured aquifers, highly heterogeneous, which is challenging. Thus, similar approaches could be deployed in different geological contexts where GW levels time series are available over long time scales. In particular, it could be very interesting to test it in karstic aquifers. This method constitutes a useful alternative to study GW flows and recharge processes and their sensitivity to imposed boundary conditions, namely, precipitations and water use*." (l 577-581)

Reviewer 2

Major comment

The paper analyzes the effect of recharge on groundwater level fluctuations in aquifer wells in the frequency domain through Fourier transforms of an analytical groundwater model. The authors invert the model to derive recharge fluctuations from water level fluctuations in wells, which is quite interesting. In the model calibration phase, the mathematical efficiency of the analytical model is taken full advantage of by carrying out a massive number of model runs that explore the entire parameter space on a regular grid.

The approach is original, interesting, and well-suited for HESS. At times, the explanation of the model and the mathematical techniques is a bit brief (see detailed comments). The method is applied to a pumped and an unpumped aquifer near the Atlantic coast of France, in a relatively humid climate.
Dear reviewer,
Thank you for your interest in our manuscript. We answered below to the main comments. The revised manuscript address all these points as well as minor comments. Our answers are in blue, sentences from the manuscript are in italic.
Sincerely.
Luca Guillaumot, on behalf of all co-authors.

From the fitting results in Fig. 6 it appears that the parameter identifiability would benefit from replacing the aquifer length as a fitting parameter by the characteristic time, but this is not discussed or explored in the paper.

Thank you, we agree with this suggestion, it would be interesting to replace calibration parameters *T*, *S* and *L* by *T*, *S* and t_c. Note we re-organized this section (4.1.2) to be very clear on the fact that characteristic time is not a calibration parameter and we added the following sentences:

"*These results suggest the parameter identifiability would benefit from replacing the aquifer length as a fitting parameter by the characteristic time. It could be achieved by re-organizing equation 2. However, this approach would not reduce the number of calibration parameters and requires more sampling as the characteristic time range is larger than length range.*" (l 346-349)

The model contains three parameters (*T*, *S* and *L*). These parameters have a physical meaning and can be compared to values from the field. Once we inverted parameters, we chose to highlight the characteristic time (a combination of the three parameters) as we found it is well constrained by observed groundwater level fluctuations and it allows to summarize the behavior of the aquifer (line 111-112 in the previous manuscript). One of the reasons explaining that we did not replace aquifer length by characteristic time during parameter inversion is that characteristic time ranges over more than five orders of magnitude given the plausible ranges of the three parameters.

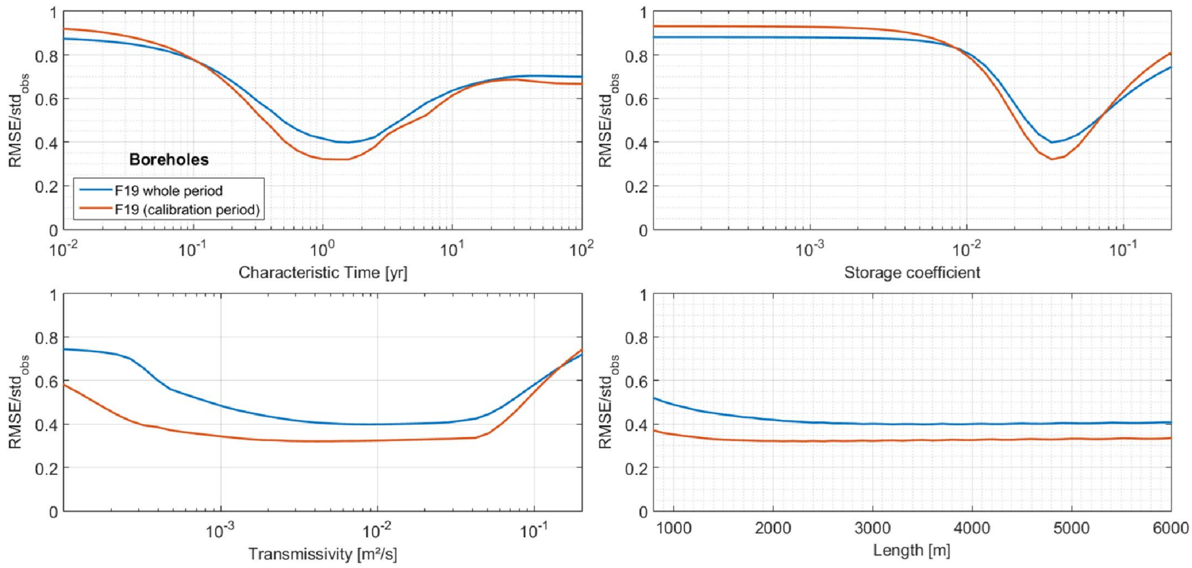
A considerable weakness of the paper is that the model is calibrated for both aquifers, but not validated, making it difficult to assess the model performance.

Indeed, we did not split observations for calibration and validation periods. Note we highlighted in Supplementary the impact of the size of the study period on storage coefficient estimates. This shows that the studied period should be as long as possible when estimating parameters. However, following your remark, we run the calibration over half the period from 1996 to 2004, instead of from 1996 to 2012, for borehole F19 located on the Ploemeur site. Results are illustrated by figures 2 and 3 below comparing parameter estimates and simulated water table fluctuations. They are also included in the revised Supplementary ("*5 Model calibration on the first half of the period and validation on the second half*"). This is mentioned in the main manuscript at the end of section "2.4 GW model parameters calibration":

"*In addition, we assessed the model performance by comparing parameter estimates and simulated water table when calibration is based on the first half of the observation period (see calibration-validation in SI)*." (l 184-186)

Estimated parameters appear very similar when obtained over the first half of the period or over the whole period. Consequently, similar water table fluctuations are obtained. To complete, results in term of parameters could be slightly altered in function of the studied time windows.

Figure 2: nRMSE in function of aquifer model parameters for borehole F19 in Ploemeur. Blue curve: model calibrated over the period 1996-2012. Orange curve: model calibrated over the period 1996-2004.



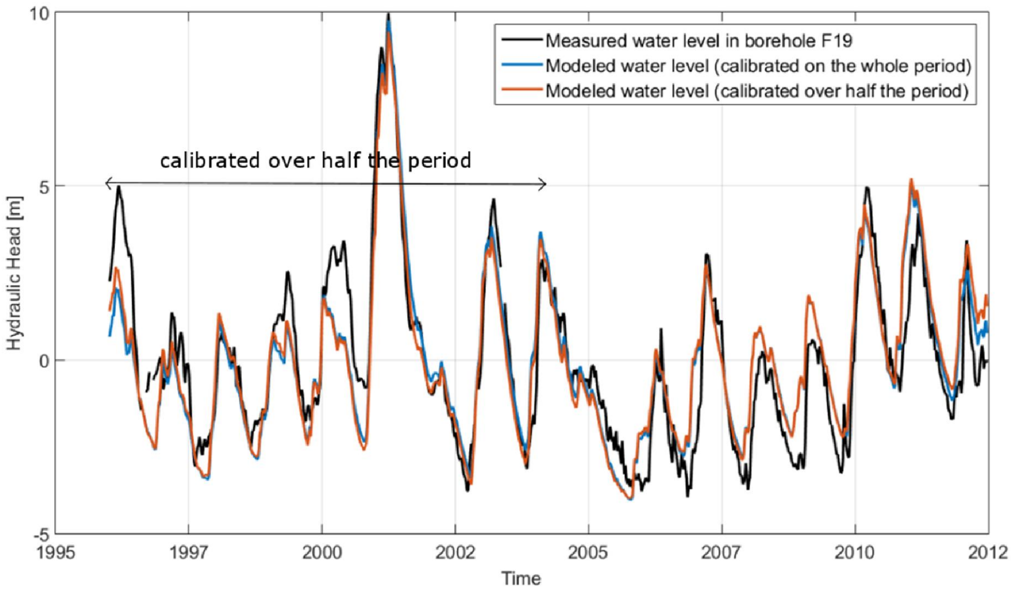


Figure 3: Comparison between observed (black line) and simulated water table fluctuations. Blue curve: model calibrated over the period 1996-2012. Orange curve: model calibrated over the period 1996-2004.

Our approach aims to estimate the informative content of the water table time-fluctuations in terms of parameters and recharge rather than to provide predictions. We shows that water table fluctuations recorded in different boreholes contain the same aquifer-scale properties in spite of the aquifer heterogeneity. Moreover, we tested all possible combinations of parameters within extended ranges of potential parameter values with the finest systematic sampling of the parameter space. Thus, for these models we will not find any other better simulations than those we obtained. We argue that the more data we have the less equifinality issues occur.

Overall, the line of thought is a bit hard to follow some times because individual sections are not as focused as they can be. Throughout the paper, the clarity can be improved by thinking about what exactly the authors wish to convey to the reader and how to do so clearly. In the detailed comments I indicate where I got completely lost. I hope this will lead to a more structured, coherent paper.
Thank you for your remark. Following your detailed comments, the revised manuscript clarifies several points of the methodology and more details are provided regarding the development of the analytical model, including more sub-steps and references in Appendices A1, A2 and A3 describing model geometry, boundary conditions and analytical resolution.

Sections 5, 6, and 7 are disappointing. They lack focus and structure and do not convey the main strengths of the study. That does not mean these strengths are not there. I very much like the modelling approach and the intended use. The Results and Discussion sections need to bring that out more strongly though.

Thank you for your comment. We understood that this main comment and the previous one relate to several detailed comments (mainly the four comments below):

- Section 5.3 uses terminology with which I am too unfamiliar to understand what points are being made. Throughout the paper, the English is a bit off, but here it somehow becomes so much so that I can no longer decipher the meaning of several sentences. This section needs to be thoroughly rewritten to be accessible to readers outside the immediate field of this paper, and the English needs substantial improvement.

- Sections 6 and 7 require over five pages of text. That is rather long. They also introduce a large number of new references, which indicates that the paper is not well organized. The Introduction and the Methods sections should cover most of the literature needed in the paper.

- Section 6.2.1. This is some kind of recap, mixed with a discussion. I like the evaluation of the soil models but do not really know what to do with the rest of this text. It reads like a brainstorm session.

- Section 7: Only the last two paragraphs are conclusions, the rest is more of a summary.

The whole section 5 describes results about recharge estimates. Section 6 hosts discussion, while section 7 constitutes the conclusion. Section 5.3 is very important, as underlined by reviewer 1. We analyze recharge fluctuations (RF) in frequency domain to highlight the different behaviors of the recharge obtained from soil models and obtained from groundwater levels in function of the time scale. The main advantage is to summarize the recharge signals for each frequency rather than comparing each recharge event individually and at different time steps.

While section 6.1 discusses methodology, sections 6.2 and 6.3 deal with processes understanding in Ploemeur and Guidel. In particular, the transformation of infiltration into recharge through the unsaturated zone and the impact of pumping. The first part of Section 6 (section 6.1.1) contains additional references mainly in order to compare estimated parameters with the literature. They are very specific to the studied site and they do not need to be presented before.

We took into account your remarks in order to clarify section 5.3 and to reduce sections 6 and 7 and make them more fluid for the readers. Sections 6 and 7 are now composed of 147 lines instead of 170 lines. See below the main modifications appearing now in the revised manuscript:

- Section 6.1.3 'Limitations' was too long (25% of section 6) and was reduced in order to avoid losing readers. We deleted the following part in section 6.1.3, because it repeats the first paragraph of section 5.2:
  "*This flexibility is a real strength, but some numerical issues can arise. High frequencies in potential recharge are damped when modelling WTF (equations 2 and 3), so that high frequency uncertainties have a limited impact. Conversely, in the backward model, observed WTF (h(xr , ω)) are multiplied by frequency (equations 5 and 6) so that any high frequency error (or noise) in observed WTF will be amplified in RF estimation. To avoid such issues, we used temporal averaging in temporal domain to mitigate such noise amplification, but alternative methods can be used.*"

- At the end of section 6.1.2, we removed "*Finally, knowing the spatial variability of aquifer properties could be unnecessary to model groundwater systems and fluxes between hydrological compartments.*" because a similar idea is mentioned in section 6.1.3.

- We removed "i*ncluding a spatially-distributed approach (Leray et al., 2012)*" (section 6.1.1) because it is mentioned later.

- We removed "*Correlation among well observations due to lateral GW flow brings insights into the aquifer averaged characteristic time.*" (section 6.1.1) because similar ideas are mentioned in Discussion (section 6.1.2).

- Indeed, several new references are not necessary. In section 6.1.1, *Bresciani et al. (2016)* and *Gauvain et al. (2021)* were removed as readers can see from our own definition of characteristic time that this property is function of aquifer length. In section 6.1.2, we removed "(*Cuthbert, 2010; Healy, 2010; Cuthbert et al., 2016, 2019b*)" as these references to the WTF method appear in Introduction.

- We moved the following part of section 6.2.1 to section 2.6 (Methods), in the same time it helps to clarify methods:
  "*The classical approach consists in defining statistics on the distribution and intensity (e.g. number of days without rain, cumulative sorted rainfall), but does not often yield satisfactory results. Considering the relationship between P−PET and recharge as a frequency dependent function is a simple but effective way to consider the impact of rainfall distribution on recharge. Indeed, this function defines the relative efficiency to generate recharge between a single rainfall event and a long-lasting wet season, and has been recently tested by Schuite et al. (2019).*" (l 210-214)

- Then, we gathered sections 6.2.1 and sections 6.2.2 and applied several modifications. The new section 6.2 is now called: "*6.2 Relation between precipitation and actual GW recharge*". (l 514-537)

- Indeed, the beginning of section 6.2 (line 499-502) was not appropriate here, it was a kinf of summary/introduction. We moved one sentence in section 6.1.1 where it is more appropriate (l 445-447). We removed the second part as it was repeating a lot section 6.1.2.

- Finally, we reduced Conclusion (section 7), as suggested, in order to focus on the main key points and perspectives. As underlined by reviewer 1, we discussed more about method applicability. (l 558-581)

We hope that this will help to clarify the manuscript.

Detailed comments

The title is informative but a bit long.
We replaced the title by:
"Frequency domain water table fluctuations reveal impacts of intense rainfall and vadose zone thickness on groundwater recharge"

L. 18: Something is missing.
We removed the copyright statement, as it will be done during editing.

You use 'indeed' a few times, but it is unclear to me why.
Thank you. We corrected in several parts.

L. 41: Potential recharge suggests this to be the maximum possible recharge in analogy to 'potential evapotranspiration', e.g., in the absence of evapotranspiration and overland flow. It would be equal to the rainfall rate, but with damping and delay due to the flow through the unsaturated. zone. I am quite sure this is not what you mean. Perhaps use another term, like 'plausible' or 'probable'. From line 88 it appears 'instantaneous recharge' might also be adequate.
Thank you for these suggestions. "Potential" could refer to "potential evapotranspiration" and then reader could think that "potential recharge" equals the rainfall rate. We understand this point but "potential" describes perfectly what we want and is explained anyway line 41-42:
"*We argue these estimates provide only "potential GW recharge" (Figure 1) because of potential storage changes and lateral flow in the deep unsaturated zone (Besbes and Marsily, 1984; Cao et al., 2016).*" (note we modified references here)

Cao et al. (2016), Han et al. (2017) or Martinsen et al. (2022) also used this term in the same sense. Note we want to keep the word "recharge" in order to be consistent when comparing results.

Dongmei Han, Matthew J. Currell, Guoliang Cao Benjamin Hall, Alterations to groundwater recharge due to anthropogenic landscape change, Journal of Hydrology (2017)

Martinsen, G., Bessiere, H., Caballero, Y., Koch, J., Collados-Lara, A. J., Mansour, M., ... & Stisen, S. (2022). Developing a pan-European high-resolution groundwater recharge map–Combining satellite data and national survey data using machine learning. Science of the Total Environment, 822, 153464.

L. 75: 'Although simple...' Is not the simplicity of the model the reason it has the adaptability you highlight here?
The aim was simply to highlight that the model can be used with different boundary conditions (no flow, constant/variable flow, constant/variable imposed head) and/or geometry (1D, radial, semi-radial). The model is simple in the sense that it needs few parameters, it runs fast and the geometry is simple as well. We modified to clarify:
"*The GW model (section 2) offers the advantage of requiring few parameters, running fast and adapting to different boundary conditions as illustrated by the two application sites.*" (l 75-77)

Fig. 1: I do not understand 'backward analytical inversion'. Is there a forward analytical inversion, as this term implies? If yes, what are the differences between the two, and are there any references you can provide?
Thank you for pointing this out. Indeed, there is no 'forward analytical inversion' but only a 'forward analytical model'. In the text we referred to "analytical inversion" or "backward model". We replaced here by 'analytical inversion' in Figure 1 and in the text.

The boundary conditions in the aquifer sketch do not seem to match those described in the caption.
We modify a bit the sentence in the caption to clarify:
"*model boundary conditions are constant head at x=L and either constant head or imposed flow rate at x=0 in order to represent natural (green case) and pumped systems (blue case).*"

L. 105: Both imposed heads are not allowed to vary in time, correct?
We added "*constant*" to the sentence to highlight that imposed heads are not allowed to vary in time:
"*In what follows, we consider a constant imposed head (hL) at x=L and either an imposed flux (Qpump(t)) or a constant imposed head (h0) at x=0 depending on the site considered…*" (l 105-107)

We also modified a bit Figure 1 to clarify this point.

L. 120: This sentence is difficult to follow. Also, the wells are located in a 2D system (the map, if you will), while your solutions are one-dimensional. Is the well distance measured along the coordinate that is represented in the 1D system? If so, than you essentially spread out a point sink (the well) over the full width W. It would be a line sink in a 2D system then. This is not to criticize this approach, just to make sure I understand it correctly. If I am correct, than 'distance to the pumping wells' should be changed to reflect it is the distance along a single coordinate only, not a true Euclidian distance.
Sorry, we updated the sentence line 120:
"*Equations 2 and 3 describe WTF as a function combining the Fourier transform of the recharge and the pumping (R and Q) modulated by the transient events frequency (ω). Therefore, high frequencies contained in R an Q are dampened in WTF in function of t_c and x_r (the monitoring well relative position inside the domain, see on the right in figure 1).*" (l 123-125)

Yes, the coordinate *x* is equal to the distance between the monitoring well and the pumping wells barycenter (the three pumping wells are located in the same zone and are quite close compared to the distance *x*). Yes, we spread out a point sink (the three pumping wells) over the full width *W*. It is a 1D system as hydraulic head would not depend of the fictive *y* axis. *W* is necessary in order to impose the pumping rate in cubic meter. Supporting Information deals with the impact of *W* and compares results with a radial system. *W* has no impact if *W* is not too small (500 m). The pumping flow is too concentrated when *W* is small, so that pumping impact is overestimated.

L. 121-125: Is that the reason you kept the boundary conditions time-invariant? Is this a reflection of the applicability of the superposition principle?
The imposed head boundary condition is time-invariant because "*Here, river level is shallow (typically 10 − 50 cm) and represents conceptually a constant head, as suggested by limited WTF observed at borehole PZ19 close to the river (Fig. 3). In such context, the assumption of imposed constant head at x = 0 is reasonable (boundary conditions colored in green on Fig. 1).*" (section '3.2 Models setting up').

In another context, it is possible to solve the equation with a time-varying boundary conditions. But in this case, it would require to use recorded river height variations.

L. 157. Apparently, I misunderstood you at L. 120. Is the aquifer length then related to the half-distances between pumping wells in the aquifer? This needs to be explained better.
It seems that the lines mentioned here do not correspond to the comment. L is not related to the half-distances between pumping wells in the aquifer. There is no reason to link the aquifer length and pumping wells location. We hope our different updates clarified this point. See for example our answer to the next comment. The re-organization of several sentences improved our explanations.

In addition, we slightly improve the right panel on figure 1, thus, it highlights the *x* position of one monitoring well inside the model domain (*L*).

L. 158: In a similar fashion, is the aquifer length related to half-distance between adjacent rivers?
L/2 is related to half-distance between adjacent rivers. We modified a bit the sentence line 158:
"*In the pumping case, x is defined as the distance* ***between observation well and*** *the pumping wells*". (l 162-163)

N.B. Reading on, I see that you discuss this in section 3.2. Perhaps give some of that info here and refer for details to section 3.2. I do not find your choice for a constant head BC terribly convincing. Is it mathematically or physically necessary? If so, please point that out more clearly.
Please, see our answer three comments above.

Section 3,2 was modified a bit:
"*At x=0, we impose transient pumping rates based on recorded pumping data. At x=L, we assume a constant hydraulic head (blue case on Fig. 1) representing typically the nearest river in that direction.*" (l 268-269)

Section 2.4 was also modified a bit:
"*Regarding length L, we refer to the density of the river network in the region. In Brittany, the hillslope length is typically 1 km (Lague et al., 2000). For Guidel site, length L varies between 20 and 2000 m with 100 values. This range is smaller than for Ploemeur site because L is expected to be smaller in 'natural conditions' compared to pumping conditions where water table drawdown tends to extend the system (Fig. 1).*" (l 174-176)

L. 157 and 158: It appears to me that if you base aquifer lengths on distances between rivers and wells, the boundary that is not the river or the well should be a no-flow boundary, not a prescribed head boundary.
Both options are possible. For the pumping site, an imposed head was more appropriate because pumping variations are changing the upstream area of the pumping. In addition, the no-flow boundary would fix the (unknown) upstream area. For the natural site, we keep the same boundary condition in order to compare results with Ploemeur. Thus, in Guidel, aquifer lengths is the distance between two adjacent rivers. Explanations are given in section 3.2.

L. 164-165 (and possible elsewhere): Units are usually not in italics, and the multiplication dot is unusual.
Thank you for pointing out this issue. We removed italic for units. We will correct the multiplication symbol following editorial team recommendations.

L. 177: '...along time.' I do not understand.
Ok, we deleted.

L. 182: The double division signs of Q/W/L are confusing. I presume you want to divide Q by the aquifer area.
Yes, we replaced by "*Q / (WL)*". Thank you.

L. 183-185: Unclear sentence, please rephrase.
We modified:
"*Finally, the monitoring well position (xr = x/L) also plays an important role when estimating recharge from observed GW levels as GW levels variations integrate both vertical and lateral recharge.*" (l 193-194)

L. 203: From this sentence it is not clear how anomalies are defined.
Ok. "Anomalies" are defined in section 2.1. We replaced by:
"*...P-PET (also expressed as variations with respect to the long term mean).*" (l 216-217)

Eq. 7: The dependency of Cxy on omega is missing. One of the Pxx should be Pyy.
Thank you. We strongly modified section 2.6 in order to be more consistent. (l 208-229)

Section 2.6 is so brief it is difficult to follow for me, as a reader who is not familiar with this technique. Please expand and add some references that give the basics of the methodology.
We strongly modified section 2.6 (l 208-229) in order to clarify and according to these comment. We added one reference, other references related to frequency analysis but between recharge and groundwater level are mentioned in Introduction (*Gelhar, Jimenez-Martinez, Townley, Dickinson…*). The two functions introduced in section 2.6 aim to characterize the relation between precipitation and recharge estimated in this study.

Fig. 2. Panel c has wells that are not mapped on the cross-section depicted in panel b. The dashed blue line in pane c is the phreatic level, right? Please mention this in the caption or the legend.
Yes, all wells are not mapped on the cross section, it could be confusing to project all wells on the cross-section. Thank you, we added the phreatic level in the legend.

L. 234-235: How does this observation (high sensitivity of deep fractures to recharge) affect the nature and the role of the unsaturated zone model for modifying the recharge signal?
This observation suggest that the fractured system is well connected with the unsaturated zone. We can argue that our modelling approach would be less relevant if the fractured system was isolated from the unsaturated zone. (l 247-249)

Fig. 4. Why does the SURFEX recharge appear to increase and then fall before heavy rainfall?
This is because SURFEX recharge is coming from a large-scale French model. Precipitation shown in Fig. 4, correspond to local records used only for GR4J and Thornthwaite models. We completed in section '3.4 Climate data':
"*...precipitation and … potential evapotranspiration (PET) estimates. Both are used to generate potential recharge from Thornthwaite and GR4J soil models while climate data used by SURFEX are derived from large-scale climate simulations.*" (l 293-295)

We also completed the caption in Fig. 4:
"*Daily precipitation (grey shaded bars) is recorded by a METEO-FRANCE weather station and is used to derive recharge from Thornthwaite and GR4J models.*"

Section 4: You only show the results of the calibration. A validation step is missing. This is a major weakness of the paper.
Please, see our response and modifications in your main comments.

L. 332: 'L/T appears ... in equation 2'. So does S/T, through the characteristic time. I am not sure this argument is valid.
Yes, S/T also appears in equation 2 (through the characteristic time), but we argue that S, L and T can compensate each other inside characteristic time, not only S/T. In addition, S appears alone on the left side of equation 2 (in "R/S"), and recharge has a bigger weight compared to the pumping component. In order to avoid confusion and because this point is more a discussion point we reduce the sentence to:
"*Conversely, transmissivity and aquifer length are poorly estimated.*" (l 342-343)

Note that we reorganized this section (4.1.2) to be very clear. (l 332-350)

Fig. 6: The characteristic time is fully defined once the storage coefficient, transmissivity, and aquifer length are determined. How can it then be a fitting parameter along with the other three? (Earlier in the paper, it was not.) From Fig. 6 it would appear a good idea to use the characteristic time instead of the aquifer length as a fitting parameter and determine L from the other three.
Thank you for this comment. The figure introduces confusion because one can think that characteristic time is a calibration parameter of the model. Only T, S and L are calibration parameters. We modified the caption of Fig. 6:

"*Evolution of the minimal normalized RMSE for Ploemeur wells as a function of model parameters: storage coefficient (top-right), transmissivity (bottom-left) and aquifer length (bottom-right). Evolution of characteristic time, a combination of the three parameters, is plotted on the top-left graph. We selected the best prescribed potential recharge model for each point. The impact of the prescribed recharge is drawn by the shaded area on borehole F7*"

Note that we reorganized this section 4.1.2 (l 332-350) to be very clear on the fact that *tc* is not a calibration parameter. Now, we first present results on calibration parameters, then on *tc* (and saying that it is a combination of the three calibration parameters).

For the second part of the comment, please refer to the main comments.

L. 358-362: This paragraph starts clear, but then you lost me. What does 'inter-compared' mean, for instance? And how can you reduce noise amplification during backward modelling by doing something that apparently makes graphs easier to read?
We simply aggregate results at a monthly scale, so that high frequency variations disappear. We updated this part to clarify (see next comment).

L. 363: How does Fig. 8 show similarity between wells? Also, the 'noise' can be easily explained physically. Why is it called noise instead of fluctuations or temporal variations caused by pumping?
Fig. 8 shows that recharge estimated in Ploemeur (in red) agrees well with recharge estimated in Guidel (in blue). Fig. 8 also shows that the recharge signal is coherent between different wells as illustrated by shaded area in red and blue. We completed the sentence:
"*Figure 8 shows analytically estimated RF at a monthly scale for both Ploemeur and Guidel sites, including uncertainties linked to parameters uncertainty and observation well variability. Phase and amplitude of RF are quite similar between the different wells as illustrated by shaded areas. They are also similar on both sites, as illustrated by the overlapping of red and blue lines on Fig. 8, although WTF are very different on each site (Fig. 3).*" (l 385-388)

Indeed, we used the word "noise" because the reasons of the high frequency variations are uncertain (high frequency variations in pumping, not detected precipitation events). Anyway, because the relative contribution of each pumping well at a daily scale is unknown, it generates noise in the estimated recharge. We proposed this modification:
"*The analytical GW model appears as a low-pass filter in equations 2 and 3, smoothing out high frequency pumping and recharge variability (these variables are divided by frequency in equations 2 and 3). When computing recharge from the backward model (Eq. 5 and 6), high frequency WTF variability is amplified (h is multiplied by frequency in equations 5 and 6). Thus, noise linked to observation uncertainties in WTF is amplified. The high frequency (daily to monthly) variations of recharge increases when the borehole is well connected to the pumped fractured zone, as suggested by difference in RF between F 9 and F 19 (SI), respectively 519 m and 268 m far from the pumping station. This can be expected as observed WTF contain high frequency variations linked to short-term pumping rate variations (and the associated relative contribution of each pumping well), which are difficult to model (Fig. 5)…*" (l 376-384)

Note we re-organized a bit this section ('5.2 Recharge fluctuations estimate for Ploemeur and Guidel sites').

L. 372-375: I can find neither a graph nor a table backing up these statements.
Indeed, we compared RF estimated from GW levels and recharge obtained from soil models from lines 368 to 375. From line 372-375, we compared weekly time series with statistical analysis. We did not show additional figures because these results is summarized by Figure 9. We modified the beginning of the paragraph to mention it:
"*Main differences appear at the monthly scale or at shorter time scales. Therefore, we compared here RF at a weekly time scale (not shown). …*" (l 392-393)

Note we re-organized a bit this section ('5.2 Recharge fluctuations estimate for Ploemeur and Guidel sites').

L. 380: Thornthwaite does not show anything in winter 2002.
This is the advantage of the WTF method. WTF can see recharge events not detected from soil models.

L. 383 and 387: What are rainfall and P-PET efficiency?
Thank you. We better explained:
"*On figure 9, the coherence and transfer functions (Eq. 7 and 8) between P−PET fluctuations and RF inform on the efficiency of the transformation of rainfall events into recharge. These functions therefore illustrate the unsaturated zone response to rainfall in frequency domain.*" (l 402-404)

Moreover, this sentence is now appearing in section 2.6 (Methods):
"*This function defines the relative efficiency to generate recharge between a single rainfall event and a long-lasting wet season, and has been recently tested by Schuite et al. (2019)*". (l 213-214)

And these sentences are still in section 2.6:
"*The transfer function H(ω) describes the amplitude ratio between output and input in the frequency domain as… equation 8…Thus, the transfer function quantifies P−PET efficiency to recharge GW, ie. a proxi of rainfall efficiency.*" (l 224-225)

Section 5.3 uses terminology with which I am too unfamiliar to understand what points are being made. Throughout the paper, the English is a bit off, but here it somehow becomes so much so that I can no longer decipher the meaning of several sentences. This section needs to be thoroughly rewritten to be accessible to readers outside the immediate field of this paper, and the English needs substantial improvement.
Please, refer to our answer to the main comments.

Sections 6 and 7 require over five pages of text. That is rather long. They also introduce a large number of new references, which indicates that the paper is not well organized. The Introduction and the Methods sections should cover most of the literature needed in the paper.
Please, refer to our answer to the main comments.

L. 499-502. What is the point of this paragraph?
Please, refer to our answer to the main comments.

Section 6.2.1. This is some kind of recap, mixed with a discussion. I like the evaluation of the soil models but do not really know what to do with the rest of this text. It reads like a brainstorm session.
Please, refer to our answer to the main comments.

Section 6.2.2. What is 'linear behavior' in this context and how does it link different phenomena? Is a 'linear coefficient' (which normally is a single-valued number, not a curve), in reality the slope of a linear relationship that you have at this point not properly defined? This is an example where lax formulations obscure what I believe to be a useful message.
Thank you for this remark. This paragraph was not clear as we did not define the 'linear behavior'. It was referring to the end of section 5.3 where we also used the term 'linear behavior'. The linear behavior is a linear regression with an affine function between potential RF and RF estimated from GW levels. We modified this paragraph in order to explain better: "*The proposed approach allowed computation of both RF and associated uncertainties at seasonal time scales to re-investigate the relationship between wet season P−PET and recharge. In addition, annual RF estimated from GW levels can be described as a fraction of potential RF using a linear regression. Thereby, the linear coefficient can be seen as a potential recharge partitioning coefficient. This partitioning coefficient should differ between Ploemeur and Guidel assuming their potential recharge is similar. On Ploemeur, 74 to 80 % of the potential recharge could be converted into groundwater recharge. For Guidel, this value ranges from 84 to 100 %. The remaining part would be attributed to lateral flow within the unsaturated zone between the soil and the water table.*" (l 531-537)

We also modified the end of section 5.3:
"*RF estimated from GW levels can be described as a fraction of potential RF using a linear regression*" (l 429-430)

L. 530: I know of very few examples, other than interflow, of lateral flow in the unsaturated zone.
We think that the sentence "*The remaining part would be attributed to lateral flow within the unsaturated zone between the soil and the water table*" (l 536-537) is appropriate because it better describes what we want (the definition of interflow is too large) and there is no direct observation of this flow.

Section 6.3.1. This does not really arise from the Results, does it? Also, it is a bit obvious.
This section proposes to explain the different recharge behaviors (mainly figure 9) obtained from groundwater levels in Ploemeur and Guidel. We argue that the results make sense because the unsaturated zone thickness is more important in a natural context. The approach allows to highlight and quantify the role of the unsaturated zone.

L. 534: This info can be moved to the site description earlier in the paper.
Thank you. We added the average unsaturated zone thickness to the site description. (l 256-257 and l 263)

Section 7: Only the last two paragraphs are conclusions, the rest is more of a summary.
Please, refer to our answer to the main comments.

Appendix A. You do not present a formal Fourier transform of the governing equation and the boundary conditions, which would perhaps make it easier to understand the line of thought and your choice to make the prescribed heads, but not the prescribed flux, time-invariant.
The Fourier transform of the governing equation is given by equation A3. We added the Fourier transform of the boundary conditions as well as more explanations in Appendices A1 and A2.

I can follow most of the development, but nevertheless would like to have more references and a few more steps in the derivations. I am not intimately familiar with the techniques you use.
Ok, we added more steps and references.

You never clarify how many terms of the infinite sums are required to achieve convergence, or what criterion you used to define convergence.
This is a technical aspect, the number of frequencies (ω) used is limited by the number of time step in the data. This point appears in section '6.1.3 Limitations': "*The number of frequencies is limited by the WTF sampling (Nyquist frequency)*" (l 510). Following your comment, we added this sentence in Appendix A1 as well.

L. 585: '...of the aquifer' There is a symbol I do not know after 'aquifer'. Typo?
Thank you, we corrected: "*S the storage coefficient of the aquifer [−]*" (*S* is without units)

Eq. (A2): The only time-dependent variables appear to be R(t) and h(x,t). Is that correct?
Yes. Pumping Q(t) is also time-dependent but will appear in Appendix 'A2 Defining boundary conditions'

L. 593: The steady-state part of Eq. (A1) is that the second derivative of h with respect to x equals a constant recharge rate, correct? In that case, should C3 equal the mean recharge rate?
Not exactly, the second derivative of *h* should be *-R / T*. From equations A8 and A9, C_3 equals h_L and h_0.

L. 600: Mixed-type (Cauchy BC) are not permitted?
We provided several additional solutions in Matlab codes, corresponding to known hydrogeological contexts. But not this one.

Eq. (A9): Are the first three terms the parabolic groundwater level found for steady flow towards drains modeled as fully penetrating ditches?
Yes, it can be seen as the groundwater level in steady state between two parallel canals.
We added:
"*The steady state part (left part of equation A9) was obtained by inserting $h_{mean}(x)$ (equation A5) in the two following boundary conditions:*
- *at x = 0, $h_{mean}(x = 0) = h_0$*
- *at x = L, $h_{mean}(x = L) = h_L$*" (l 640-643)

It is not clear to me how you arrive at Eqs. (A10) and (A11).
We modified one sentence and added an intermediate step in order to clarify. (l 645-663)

L. 630-633: This is not a real test, is it? You modeled numerically the simplified set-up that permitted the analytical solution. This used to be done when in the 1970s and 1980s to test the accuracy of numerical models but was never intended to evaluate the analytical models.
Yes, we tested the analytical solution regarding water table (Appendix B1), but the main interest was to test the ability to recover recharge from water table (Appendix B2).

Supporting information: If you have a radial aquifer which is pumped in the middle, does that not lead to a singularity at the center of the radius of the well is zero?
Yes, it means that we can not obtain results at x=0 in a radial system. We used observation wells not located in x=0.