*Interactive comment* on "Statistical Characterization of Environmental Hot Spots and Hot Moments and Applications in Groundwater Hydrology" by Jiancong Chen et al.

Jiancong Chen et al.
nigel_chen1993@berkeley.edu
Corresponding to: rubin@ce.berkeley.edu

We appreciate the reviewer's efforts in reading our manuscript and providing useful comments and recommendations. We will provide summary notes to the reviewer's general comments and responses to the reviewer's detailed comments and questions. Please see below.

Summary notes:
1. **Why stochastic?** The processes governing the HSHMs are very likely subject to some uncertainty. There may be uncertainty in the parameters and also in the governing equations. This is the reason to adopt a stochastic formulation. In this way, the uncertainty can be modeled and even reduced by taking advantage of the information provided by the data. The mechanism for modeling and reducing uncertainty are built into our approach. For example, we can use prior information from similar sites, and we can use local measurements for conditioning.

2. **Defining places in space and time as either being or not being HSHMs**. Our characterization of HSHMs is not binary, because we use probabilities. In our approach, all space and time intervals in the investigated domain are associated with probabilities to be or not to be a HSHM. Notice that the location of the HSs is uncertain due to the combined effect of physical system heterogeneity and limitations in its characterization. In addition, even if the positions of the HSs are known without uncertainty, hot moments may depend on other factors (e.g., solute pathway, retention time), which can also be uncertain.

3. **Arbitrary cutoffs**. In our approach, the cutoffs are defined by the user and can be modified at will. One can change the cutoff values based on prior information and based on risk tolerance. For HSHMs that have negative influences, thresholds are often introduced in environmental regulations in order to identify levels of contamination above which to consider a site as contaminated. In addition, activation thresholds may be used to identify the thresholds for reactions that are necessary for biogeochemical driven HSHMs (see also point 5 below).

4. **Binary view of HSHMs.** As stated in point 2, we model the HSHMs stochastically. For example, we can have a zone with high probability next to other zones with lower probabilities in terms of HSHMs occurrence. Thus, we do adopt a continuum approach by creating HSHM probability maps. In another note, we suggest that there might be situations that require focusing on a particular area because of a need to focus on the site investigation efforts. Thus, in our approach, we can identify areas that are more critical/sensitive compared to others, and this could assist the project managers in defining priorities. For example, at the Rifle site (Wainwright et al. 2010), geophysical

datasets indicated the presence of naturally reducing zones (NRZs), which may have higher level of uranium and nitrate. Based on this information, site investigation and parameter estimation were both goal oriented, which reduces efforts and uncertainties in quantifying the corresponding HSHMs.

5. **Improving understanding**. We will expand our discussion in order to improve understanding. In particular, we make the following points: (a) Our framework can be used to investigate HSHM sites and identify the process and parameters controlling the HSHMs. (b) As the reviewer noted, there's uncertainty associated with the HSHMs. Using our approach, we can identify which models and parameters work, using for example Bayesian model comparison and identifying the best performing models or whether the current understanding of a certain HSHM is lacking. (c) The probabilistic approach offers great advantages of addressing the uncertainty on HSHMs and reducing it. (d) We will also add information on where to get the threshold parameters.

6. **Why a statistical framework**? Based on the experiences in the hydrology community, the coupling of probabilistic concepts with the physics led to a tremendous progress in our ability to model the complex phenomena taking place in the subsurface. Similar observations have been made on multiple disciplines within earth sciences. There is a vast body of knowledge accumulated in hydrology and what we want to show in this paper is that this knowledge could also bring enormous potential to HSHMs investigations.

7. **Simple language**. We will add plain language discussion in the revised manuscript.

In the following section, we will provide a detailed response to the reviewer's specific comments and questions.

Detailed responses:
*L40: I am having a hard time with the definition used for HSHM as it depends on the event having a negative effect on something (health or environment). But what if a HSHM does something beneficial like remove a pollutant. Is that not considered here? Maybe a slight modification of the definition is all that's needed.*

Response: We will modify the definition to include the beneficial perspectives (L41). Table 1 in the revised paper will exhibit an extra column indicating whether a specific HSHM has a positive, neutral or negative impact on the ecosystem.

*L90: Here again the focus is on the negative side of HSHMs, I suggest taking a more balanced view that includes their benefits as well.*

Response: We will expand the HSHM definition to cover both positive and negative perspectives of HSHMs. Please also see our response to L40 comment.

*L100: I would suggest adding 1 sentence providing a non-jargon definition of indicator statistics. It will make the work more accessible.*

Response: We agree with the reviewer and will make necessary modifications to make the reading more accessible.

*L135: While I appreciate the development of a rigorous statistical framework, I question the utility of the binary definition of whether a place/time is or is not a HSHM. Do we care more about the definition or its influence? The influence is some continuous function of the magnitude to which it deviates from background conditions. It would be more powerful to define a statistical framework that captured this more continuous perspective. At a minimum, I think the authors should discuss the limitation of their framework and provide ideas for developing a more continuous approach. For example, maybe one could continuously vary the Cth and Rth from equation 1 to examine outcomes across a continuum of thresholds?*

Response: As discussed in items 2, 3 and 4 above, our framework is flexible as it can incorporate different conditions that trigger HSHMs. The cutoff values are chosen by users and can be modified at will. With this flexibility, one could definitely vary $C_{th}$ and $R_{th}$ values, and examine how the probability of HSHM occurrences changes correspondingly, as suggested by the Reviewer. Thus our approach indeed captures the continuous perspectives as specified in the previous answers.

*L145-L150: All examples are for concentrations (Cth). It would be good to provide some examples for rates (Rth).*

Response: In the revised manuscript, we will include multiple examples for rates. For example, $R_{th} = 0$ can be used for chemical reactions that have significant negative impact on ecosystem (e.g., nuclear reactions). $R_{th}$ values can also be obtained based on similar studies, such as denitrification and carbon cycling rates summarized in Harms and Grimm (2008).

*L160: Please briefly explain why type B includes the spatial component instead of only including the temporal.*

Response: The idea here is to identify where the dynamic conditions exist in conjunction with spatial zones to trigger a hot spot. For example, following the reviewer's comments, a zone of high concentration may be a location for HSHMs, only that we do not know where it gets triggered. An example here could be nuclear waste remediation sites where natural attenuation strategies are in place. While contaminants can be held in place; within the zones where contamination occurred, yet some temporal conditions may trigger the formation of these HSHMs. Here it is worthwhile to note both the temporal conditions, but also the spatial domain of HS. Furthermore, the hot spot may also depend on variables different from the concentration of the species of original interest. For example, a nuclear contamination site that has historically looked at uranium can now be potential hot spots for strontium.

*Equation 3: This seems a bit circular to me. It seems like this says that a location is a hot spot because it has the conditions (e.g., concentration) needed to be a hot spot given our* *defined threshold of what counts as a hot spot. So it's a hot spot because it's a hot spot. Maybe this can be clarified in terms of how this isn't circular? In other words, explain further why it is useful to call some place a hot spot based on defined criteria. Why don't we just define the location based on its levels of continuous variables relevant to a given situation? This goes back to my comment above about the very binary nature of this approach. I am not yet convinced that this is really moving us forward a great deal. Though I am keeping an open mind as I read.*

Response: As mentioned in L160 response, the location in time and space may be unknown. Equation (3) is related to spatial variability and uncertainty in site characterization, which leads to uncertainty in identifying the locations critical for HSHMs. The definition is needed to define the corresponding statistical random variables.

*L225: I like the statistical framework here, though it presumes that we have complete (or very good) knowledge of the spatial and temporal factors governing HSHM 'activation' and I wonder if that makes this framework difficult to use? That is, if we already know the conditions that lead to HSHM behavior, then we already know that, and I am not clear on what we are learning from this framework.*

Response: We agree with the reviewer that work in the HSHM community thus far has been remarkably site-specific. To enable the transferability of HSHM features from one site to the other, we have proposed this statistical framework. For example, static indicators at riparian sites could be quite similar – riparian buffer strips or microtopographic depressions. By using a statistical formulation to capture these spatial zones and applying them to a new site under corresponding dynamic conditions can help us pre-identify potential zones of HSHMs. It is also important to note that the impact of HSHMs does not depend only on the fact that they may exist in a certain compartment, i.e. riparian and hyporheic zones, but also on their location and duration in the active state, which may be intermittent. All these factors are uncertain because we don't know the exact location of HS and for how long they are active under new conditions and new sites. Thus, our stochastic approach is beneficial to enhance applicability to other sites.

*L265: Again I am not understanding what we are learning here. The hot spots have been defined as NRZ with specific quantitative conditions. So what more is equation 10 telling us? I was expecting to see a figure or analysis here that went to the next level of understanding through the use of eq. 10.*

Response: Equation 10 is an example how we can construct a static indicator quantitatively. Please also see our responses to L225 comment.

*L340: Unless I am missing something, the examples are based around meeting specific conditions in space and/or time, and saying that if all conditions are met, then a HSHM should occur. That's all fine, but again, what are learning from that? It seems like this boils down to an if-else statement that is built around previous analyses of a given* *system. That seems really straightforward to the point that I fell like I am missing something. Maybe more of the implications can be drawn out through these sections?*

Response: There is a challenge in knowing what the conditions required to trigger HSHMs are, but knowing the conditions may not suffice to predict when and where. This is where our proposed approach comes in. As mentioned above, our proposed framework is unified and allows us to investigate a variety of HSHMs, with complex dynamics multi-dimensional dynamics and under diverse conditions of uncertainty. It also can be easily integrated with the concept of Bayesian conditioning in order to reduce uncertainty and to develop site investigation schemes with information from similar sites. We do not promulgate an if-else approach, because we assign probabilities over the entire domain. The proposed equations and formula in this section are mainly presented to show how the framework can be utilized and how the corresponding indicators can be constructed.

*Section 4.2.3: To be honest, I am not savvy enough to follow the math in this section. I can only assume that it is correct, and maybe other reviewers can go through it. Regardless of whether it is correct or not, however, I do not understand the purpose of the formulations. Maybe they come clear further into the paper. At this point this section and the previous two seem esoteric, and I am not sure what the work is really driving towards.*

Response: In previous sections we focused on evaluating the probability of HSHMs occurring at a given time $t$. This allows us to evaluate when and where we could observe the highest probability of HSHM occur. As hot moments can persist over time periods, estimating the corresponding probabilities for given time intervals becomes also quite important. And this is the main reason for introduced section 4.2.3.

Specifically, Equation (22) – (26) describes the dynamic indicator and an analytical stochastic solution for the HSHM. These equations can be simplified into Equation (27) – (30) if the hot spot can be defined by a simple geometry as described in Line 458. The deviations of these equations are based on stochastic theories, which are well documented and extensively verified cf., Dagan. (1989) and Rubin. (2003).

*Section 4.3: Not clear on what 'w' is in this case. More generally, I continue to struggle to understand what we are learning. I really want to get on board and I feel strongly about HSHMs as important features, I just am struggling to connect the conceptual dots.*

Response: In most world conditions, HSHMs occur within a volume rather than a single spot. And this is why we introduced 'w' in the mathematical formulations, which represents the corresponding dimensions of this control volume.

*L515: Can you show a figure of this? pretty hard to understand as is.*

Response: Hot spot $\Omega$ was placed $21I_{YH}$ away from the source, and the dimension of $\Omega$ is $(2I_{YH}, 2I_{YH}, 2I_{YV})$. Figure 3 presents the configuration of this example, where the red box is the candidate hot spot $\Omega$.

*L520-540: Okay, so now we start to see some results from the framework, in which the time course of HSHM development is linked to variation in conductivity. This is nice, though I must wonder whether the formal statistical framework is necessary. Could this be done just as well with a Monte Carlo approach? What I am missing is a convincing argument that the formal framework is needed. Could one not just run a simulation and sample it to characterize the spatial distribution of biogeochemical rates, use that to determine the frequency and magnitude of hot spots and then do that through time to show the time course?*

Response: Our approach can definitely be applied using Monte-Carlo (MC) approaches. We present a framework, and it can be applied using analytical models (when available) or using MC simulation. These approaches are complementary rather than exclusive. One can use our framework to define the flowchart for the Monte-Carlo analysis. Although MC approach can be used for implementation, our approach goes further than MC because it can easily incorporate Bayesian concepts such as conditioning as well as utilization of prior information from other sites. For example, knowledge from previous nitrogen HSHM studies can be implemented with the proposed framework and guide new HSHM investigation at other new sites. More details will be provided in the revised manuscript.

*L595: I don't recall seeing any results showing how the framework can be used to study uncertainty. This seems important, but not presented.*

Response: As we stated in point 1 and 6 above, our proposed framework incorporate uncertainty through modeling the dynamics as stochastic processes and through modeling the parameters as random variables. For example, in Section 4.4, we show how the uncertainty surrounding the hydraulic conductivity influences the probability of HSHM occurrence in the subsurface.

*L605: I think it would be useful to expand on the discussion through the manuscript in terms of how the framework provides understanding of mechanisms. Through much the paper it seemed that the mechanisms were known a priori and were actually used to define conditions that result in HSHMs. I don't fully understand how we are gaining more mechanistic understanding, but I am open to hearing more.*

Response: We appreciate this comment. In the revised manuscript, we have incorporated certain Bayesian statistics theories into the indicator formulation and expanded the discussion correspondingly. The wide range of approaches used for modeling HSHMs reported in the literature are helpful in gaining a better understanding of HSHMs, however it is challenging to evaluate and rank the suitability of the models for realistic scenarios. This is where our study becomes useful. The flexibility of our proposed framework enables us to compare the performance of competing models and select appropriate models for new sites. And we would cite here a couple of examples. First, Bayesian model averaging approaches (Volinsky et al., 1999) could be implemented to obtain a combined and less risky estimation of HSHMs at new sites. Second, model comparison criteria, such as the Akaike information criteria (AIC, Akaike, 1974) and Bayesian information criteria (Schwarz, 1978) can also be applied to compare and rank the performance of different HSHM indicator models and their ability to explain observations. For example, smaller AIC and BIC values indicate a better match between a HSHM model and data. Large AIC and BIC values would suggest an incomplete and possibly even faulty model. Through this process of model inter-comparison, we could gain better understanding of the underlying mechanism, which, in essence, is the learning process that the reviewer rightfully wishes us to show.

Reference:

Akaike, H.: A New Look at the Statistical Model Identification, IEEE Trans. Automat. Contr., doi:10.1109/TAC.1974.1100705, 1974.

Dagan, G.: Flow and Transport in Porous Formations., Springer Verlag, Berlin., 1989.

Harms, T. K. and Grimm, N. B.: Hot spots and hot moments of carbon and nitrogen dynamics in a semiarid riparian zone, J. Geophys. Res. Biogeosciences, 113(1), 1–14, doi:10.1029/2007JG000588, 2008.

Rubin, Y.: Applied Stochastic Hydrogeology, Oxford University Press, Oxford, UK., 2003.

Schwarz, G.: Estimating the Dimension of a Model, Ann. Stat., doi:10.1214/aos/1176344136, 1978.

Volinsky, C. T., Raftery, A. E., Madigan, D. and Hoeting, J. A.: David Draper and E. I. George, and a rejoinder by the authors, Stat. Sci., 14(4), 382–417, doi:10.1214/ss/1009212519, 1999.