

HESS 2015-538

Dear editors, we are grateful for the opportunity to resubmit the manuscript in a version that address all the valuable suggestions made by reviewers. We have followed the advice to split the manuscript into two parts – which will, however, have to be read in conjunction; both are included here in a single file, as advised.

Our specific responses to all suggestions are listed in a table below

	Reviewer's comments	Authors' responses
Reviewer 1		
1	The paper presents an interesting attempt to develop a simple measure of flow persistence and address the issue of the ability to detect and attribute variations in daily riverflows to the effects of land-cover change in large catchments (riverbasins), a well-known issue in hydrology.	Thanks! We aim to derive an indicator of watershed health that mainly requires a series of data on rainfall and the temporal pattern of river flow at a point of interest.
2	The authors then use a model for estimating flow persistence to try to demonstrate how difficult it can be to identify the effects of land-use change in four tropical catchments. Their illustration of the sample sizes needed to identify effects is also interesting but difficult to interpret as too little information is given on the catchments used as examples	We have provided further details on the subject.
3	I am in full agreement that the issues of detectability and attribution are important for hydrologists to investigate because decision makers need evidence that catchment restoration can reduce flood risks and increase flow persistence by redirecting water from flow paths with rapid responses to rainfall to those with slow paths. This evidence can then be used to demonstrate that those benefits are being realised.	This is indeed the overall direction of the paper.
4	The authors also note that these effects are well documented in hillslope and small catchment studies but there is little evidence of such effects at in large catchments (river basins).	Agreed
5	I have no problem with their argument that some form of measure of hydrological change is important, but (a) their method of deriving the persistence indicator and (b) of applying their model to the four catchments is not adequately explained.	We have provided further details on the subject. See further explanation below.
6	I have divided my review into two sections based on the two main components of the paper: (a) flow persistence and (b) flow change detection and attribution. There are some typographic errors but I have not gone into these as the paper needs rework.	Thanks – we hope the current version addresses all concerns and suggestions and is clean

7	<p>A) The one part of the paper addresses both flood risk and flow buffering by measuring aspects of the flow responsiveness of a catchment using a simple index (Fp) of the flow persistence. Flow persistence is defined as identical to the ‘recession constant’ (pg 7 line 7).</p> <p>However, I would argue that flow persistence is only half the picture, what is needed is actually a measure of flow responsiveness to rainfall because flooding can depend on how rapidly the flow increases versus factors that constrain that flow and cause water levels to rise. Flow responsiveness is also directly influenced by antecedent wetness, rainfall event depth and duration, and other factors which they do discuss but do not seem to incorporate in their approach.</p>	<p>We agree with flow persistence is focused on half the story (peak flows and subsequent decay), and that the speed at which peak flows are attained matters beyond the peak itself. However, since our aim is to derive watershed indicator using parsimonious information, we could not study the flow responsiveness to rainfall at timescales less than a day. If shorter-term dynamics matter, their study will require detail information on the specific space-time pattern of rainfall and flow measurements at least on hourly time-scale. Our study focuses on situation where daily records are the only information exists empirically. We explained and discussed further in our paper (see e.g. line 423-425).</p>
8	<p>Maybe I missed it, but I did not find a clear statement that Fp is only being calculated for the descending limb of a hydrograph. Yet this must be so because the range for Fp is constrained to the range 0-1. I would argue that to understand flood risk you also need to measure the ascending limb of the flows (the rapidity of the rise in response to rainfall patterns). I think would be probably be necessary to have indices for both the ascending and the descending limb as they are rarely symmetrical.</p>	<p>We agree that the Fp parameter (at daily scale) alone is not sufficient to predict flood dynamics at shorter timespans (e.g. hours), as details of spatial and temporal storm patterns interact with characteristics of the streambed, beyond what Fp captures. However, the finding that the proportion of fresh rainfall (minus soil water storage capacity linked to preceding Et) that comes down as riverflow is (1-Fp), allows us to infer an important component of flood predictions: the peak daily flow volume (given rainfall). We agree that a further empirical parameter at hourly (or similar timescale) can add further value – but for the empirical data sets we used only daily records are available. Indeed we don’t assume that the ascending and descending limbs are symmetrical.</p>
9	<p>Attempting to parameterise Fp for the various flow pathways seems to me to add complexity but not much insight given their focus on large catchments where the relative importance of different flow paths can vary considerably across space and temporally.</p>	<p>This part is presented as an aid in the interpretation of the aggregate Fp, not as a way of empirically deriving it. We have tried to state this more clearly in the revised manuscript, section 2.3 (lines 288-312).</p>

10	<p>I did not find Figure 1 helpful as an illustration of the causal pathway or for placing the wide range of factors into their context. It has too much detail presented at the same level, is not adequately explained, and has too many terms that are not explained in the caption. A figure should be able to stand on its own with its caption, but in this case it does not, even with an extensive caption. A part of the problem is the use of two sets of arrows with different meanings in ways that are confusing. A more common way of representing this kind of diagram is a flow chart where the factors that influence (solid arrows) the (relative) magnitude of the water flows (hollow blue arrows) are represented in ways which make their role much clearer. I do not understand why rainfall is not made the start of the diagram and why the caption ends with number 0 rather than beginning there as one would intuitively expect. It is also not clear that land cover has various influences at both the “plot” level (whatever that may be) AND at the hillslope/landscape level. It is also not clear to me why there is a blue water flow arrow directly from rainfall to #2 and to #3 without passing through the landscape (and why #2 is in brackets). Why are some of the arrows broad and others not? Is the “triangle” to the left of human population density an arrow? Why are human population density and topography (subsidence) linked? What is the relevance of subsidence? Why is topography placed here and not within the sets #0 and #1 given its importance as a factor in the generation and flow of surface and subsurface runoff at both plot and hillslope levels?</p>	<p>We have rearranged the elements in Fig 1, and hope it now supports the understanding of the relations we try to highlight in the manuscript.</p> <p>We have replaced “plot” by “patch”.</p> <p>Subsidence due to groundwater extraction in urban areas of high population density is a specific problem for a number of cities built on floodplains (such as Jakarta and Bangkok). It is a rather specific situation, but economically important flooding risks that were at first attributed to changes in upland land use are now understood to be generated in the urban areas. We have add a brief explanation and reference to this point of detail (line 154-157).</p>
11	<p>For a paper that attempts to explain how land cover changes affects catchment flow responses I find it inexplicable that there is almost no reference to: (a) the very extensive body of hillslope hydrology research into flow pathways and the temporal effects of different water partitioning and surface/subsurface on flow response to rainfall inputs; and (b) how hillslope responses might scale up to larger parts of catchments and large catchments. Even the brief mention of the different ways in which overland flow can be</p>	<p>Our primary purpose with the paper is not to explain how land cover changes affect hillslope hydrology, but rather to start at the other end and evaluate the information content of a metric that is derived form observations of river flow alone – we are indeed aware of the large body of work (and various models) at hillslope scale, and have added some further references in section 2.3, but the paper is not meant to be a review of all we know about ‘floods’, but to evaluate</p>

	<p>generated (e.g. Hortonian versus variable (saturated) source areas) fails to cite the original research papers and the insights they provide in the catchment responsiveness.</p>	<p>a very parsimonious model, that can in a single index capture important first-order predictive value of the influence the watershed has, as modifier of stochastic rainfall.</p>
<p>12</p>	<p>Despite reading section 2 a few times I am still not entirely clear on the logic of the various deductions that are made about low flow, seasonality and the influence of varying F_p on the form of the hydrograph. Perhaps this is because the text is not always very clear. For example, the authors present the following (page 7 line 19 onwards): “If we consider the sum of river flow over a sufficiently long period, we can expect ΣQ_t to closely approximate ΣQ_{t-1}, and thus $\Sigma Q_t = F_p \Sigma Q_{t-1} + \Sigma \epsilon$ (equation 2) From this relationship we obtain a first way of estimating the F_p value if a complete hydrograph is available: $F_p = 1 - \Sigma \epsilon / \Sigma Q_t$ (equation 3)” The only way I can derive equation 3 from 2 is to assume that $\Sigma Q_{t-1} = \Sigma Q_t$ and so ΣQ_t can be substituted for ΣQ_{t-1}. However, if this is so, then the only way equation 2 can hold is if $F_p = 1$ and $\Sigma \epsilon = 0$. If this is so, how can this relationship then be used to estimate F_p? Or am I missing something here? In section 2.4 I assume that a model with a set input of daily rainfall and flow responses to that rainfall was used to create the ascending flow limbs so that F_p values could be used to generate the descending flow limb? And so that F_p could be varied? I also had similar difficulty in following parts of the methods section. For example, on pg 12 line 9 the term Q_{add} is abruptly introduced without an adequate explanation of its meaning. This is followed by the ‘apparent Q_{add}’ and $F_{p,try}$, again with no proper explanation. I should not have to go and find the paper cited (in fact a user manual) or to download the spreadsheet for an adequate explanation of the terms or to find a proper explanation of the FlowPer algorithm.</p>	<p>This is a crucial point in the derivation, we hope that a number of changes we made in the presentation make the text easier to follow.</p> <p>We hope that replacing the notation ϵ that may have, inadvertently, suggested that this is a negligible term by Q_a the misunderstanding is avoided.</p> <p>If we indeed assume $\Sigma_1^t Q_t = \Sigma_0^{t-1} Q_t$, we obtain $\Sigma Q_t = F_p \Sigma Q_t + \Sigma Q_a$ and hence $\Sigma Q_a = (1 - F_p)(\Sigma P - \Sigma E)$. The easiest way to obtain this relationship at the level of annual sums, is to have it hold true at event level: $Q_a = (1 - F_p)(P - Ex)$ with the caveat that the way ΣE is partitioned over terms calculated at each day with rain (ΣEx) may require some nuance (where not all antecedent ET is compensated in a single storm on soils that don’t easily rewet, for example).</p> <p>The key point, however, is that the stochastic $(P - Ex)$ term is to be multiplied with $(1 - F_p)$, which leads to direct influences on peak flows (if F_p does not vary with Q).</p> <p>We have made the text more explicit in the way Q_a is calculated for each $F_{p,try}$ value: Every pair (Q_t, Q_{t+1}) yields an estimate of Q_{add} derived from $Q_a = Q_{t+1} - F_{p,try} Q_t$ Each $F_{p,try}$ value thus generates a frequency distribution of inferred Q_a values (some of which will be negative for</p>

		relatively high F_p try values), and the algorithm retains the value that minimizes $\text{Var}(Q_a)$.
13	The second part of the paper deals with the application of the GenRiver model for assessing the impacts of land cover on river flow and its attribution and detectability. The entire model and its application is only introduced in the methods but its structure and use should really be described already in the introduction. The model is said to be spatially explicit but it is not clear how that is realised in practice (i.e. it a distributed model?).	The use of the model to explore F_p derived from hydrographs for scenarios other than current land use indeed had to be more explicitly announced in the introduction – which we have done. However, it remains a ‘tool’ for exploring F_p properties of hydrographs, rather than being a focus on its own. By separating the manuscript into two parts, further explanation has been added on the the GenRiver model in part II. With a full specification of the model available for who wants to get into the detail, we have increased the description of key features that likely influence the results obtained.
14	Tables 1-4 do not provide an adequate description of the study catchments – what does dominant land cover mean? Although the authors note the importance of knowing what changes where in a catchment in relation to flow paths and times and attributing responses to changes and other factors, Table 1 does not give any indication even of how much of each land cover there is the baseline situation. Why not provide summaries or maps? Nor are we given information on where, in relation to this baseline state, the changes in land cover are made for the different scenarios. Why not provide a summary or a map? In Table 4 there is a repeated use of “some” in describing the changes made. This to me is not acceptable.	Thanks for the suggestion – we have add a map as Fig. 1 in Part II that visualizes the baseline situation, and a table of (plus figure as appendix 2 of part II) on what the scenario’s mean in local context. A point of warning may be needed here, as the case study catchments have not been subject to a multi-year intensive measurement campaign of the type that reviewer probably finds necessary to fully trust results. We do provide the level. Of correspondence between F_p ’s derived from measured and modelled hydrographs, and on that basis present the further model results as illustrations of what can be expected, rather than as statements of fact.
15	We are not given an adequate explanation of how the single values of each of the 13 parameters of the GenRiver (Table 2) were obtained. Those parameter values are all ones that would vary a great deal spatially and with different land cover types (e.g. interception), but only a single value is given with no indication of their variability in the study catchments or how representative each value is. Providing definitions of the terms in a user manual the	We have provided full specification of the parameter values used for the calculations, while the model is freely downloadable, so we’ll meet reasonable standards of repeatability – but of course further tests of the F_p summary statistic on hydrographs obtained with other models for other situations are likely needed before this method can be more widely accepted by the community.

	<p>reader would have to look up is simply not acceptable. Table 3 also gives values for three important parameters for each of the land cover types with no explanation of what their sources and ranges are (BTW surely interception [Table 2] differs between forest and annual crops and so is land cover specific?). Table 3 also introduces the term relative drought threshold with no explanation of what it means and how the model uses it. The legitimacy, accuracy and representativeness of these values, together with the land cover changes, are critical to our confidence in the model outputs and thus in the analysis of the detectability and attribution of the changes in flows to changes in land cover. A study should be repeatable and this hypothetical modelling exercise certainly is not given the information included in the paper. In summary I am not entirely sure what to recommend overall. The idea of deriving a simple but robust measure of flow change (i.e. flow responsiveness) which can be causally related to land cover changes is sound, and necessary.</p>	<p>We hope that the discussion gives a fair assessment of the level of evidence, avoiding an oversell.</p>
16	<p>Flow persistence (Fp, recession) is an interesting measure and can be related to changes in the relative importance of different water flow paths, but it is also evident that it is not straightforward to derive and could be masked by the effects of location and catchment heterogeneity. I do think that a measure of the flow recession is not sufficient, the nature of the whole response to rainfall needs to be assessed for flood risk. The flow persistence component of the paper needs careful thought to make sure that the measure(s) are clearly and thoroughly explained. Even so, I still am left with the question of whether a simpler approach would not be to examine the slopes of the flow recession curves (in relation to rainfall event sizes and sequences [antecedent conditions]) for possible shifts due to land cover changes. Alternatively, using shifts in the flow duration curves as measures of changes in the relative importance of flow pathways, as has been done elsewhere, would be more effective and understandable.</p>	<p>It seems that reviewer presents a multi-parameter description as 'simpler' than our single metric? Again, our focus is to assess the information content of a metric of flow predictability that appears to align well with the way downstream observers describe and experience changes in the conditions in upper watersheds. If our primary focus would be to assess the effects of land use change on flood risks as such, for any of the catchments studied or for the wider geographic domain in its totality, we would probably embark on further data collection... A key point here is that in many situations (historical) rainfall data are inadequate – at best a few point records are available, but no spatially weighted average, and there are many degrees of freedom in getting reasonable answers for wrong reasons in multiparameter models.</p>

17	<p>Another alternative would be to use the relationship between rainfall event sizes and sequences (e.g. antecedent wetness) and flow response to those events and sequences to infer changes. The modelling component needs a lot more information to back up the chosen parameter values for both the hydrological (Table 2) and land cover-specific values, as well as specific information on the extent of the land cover changes and their spatial configurations. It also needs to provide information on how well the outputs it generates for the different land cover types compare with the findings of other studies (i.e. how well does the model perform). Overall I need more information on the model structure and setup to interpret how well it performs in this application. This would require expanding the paper substantially.</p>	<p>We have added information on the model in the methods and appendix and provided the main parameters with its values in Table 2 and 3.</p>
18	<p>Overall, my conclusion is that perhaps this paper attempts to cover too much ground and should be two papers: - One on the issue of catchment responsiveness to rainfall as a measure of land cover change, including flow persistence - One on modelling of the effects of land cover change on river flow responses and the difficulties of detecting and attributing changes in flow responsiveness to changes in land cover (and relating this back to the changes in the relative importance of flow paths linked to the changes in land cover).</p>	<p>As suggested we have split the paper into two: 1) on flow persistence concept as indicator of watershed health and 2) on the application of flow persistence (Fp) concept on 4 watersheds. We believe summarizing hydrographs resulting from different land use scenarios into Fp can play a role as 'metric of watershed health'.</p>
Reviewer 2		
1	<p>Summary Overall, I think that this study addresses a critical knowledge gap with important implications. However, the conceptual foundation regarding watershed health and flow predictability requires a closer examination. The derivation of Fp and the process used to create the GenRiver model parameters needs more discussion, clarification, and justification if the reader is to have any confidence in the results. I think that if the authors were to accomplish these revisions then the paper would simply be</p>	<p>Thanks for the interest, we have followed the suggestion to split the manuscript into two parts "I. Theory, II. Applications", and appreciate the detailed suggestions and comments.</p>

	<p>too long and cover too much ground. Breaking the research into two separate papers is probably a better course with one focusing more on the conceptualization and creation of the Fp term and one on the application of it within the GenRiver framework.</p>	
2	<p>As I understand the article, the authors are attempting to develop a single measure of watershed health called 'flow persistence' (Fp). This Fp parameter measures the volatility of daily river flow in response to land cover change within large catchments.</p>	<p>We have tried to make the manuscript more clear in the steps involved: Fp is a measure of 'volatility' (or its complement), that can be used to quantify one aspect of the impacts, at multiple scales, of land cover change.</p>
3	<p>One of the key objectives of the study is to determine the value of specific land cover types in terms of flood mitigation. The study itself is broken down into two phases: (1) the derivation of a river flow algorithm, and (2) the application of the algorithm within four watersheds with different rainfall and land cover characteristics.</p>	<p>In the current manuscript the Fp definition in terms of a recursive river flow model is presented as step 0, before the empirical steps on how actual (or simulated) flow data can be used to derive an estimate of Fp (step 1), illustrated with applications in four case studies (step 2).</p>
4	<p>The key points that need to be addressed include: 1. A better justification that flow predictability does in fact correspond with watershed health; 2. A better explanation of the flow persistence derivation; and 3. A much more thorough explanation of the Fp algorithm application within the four catchments.</p>	<p>We thank reviewer for the positive suggestions and have I tried to clarify these points in the resubmission. "does in fact correspond with watershed health", is however a complex question, as "watershed health" hasn't been satisfactorily quantified in absolute terms anywhere. What we claim is that changes in Fp, towards lower values match "degradation" and towards higher values match "restoration", from a downstream perspective. The metric matches a common way of describing degradation as loss of predictability. In unpacking the concept, however, we find that the specific pattern of rainfall in a given year interacts with the condition of the watershed in generating a river flow pattern, as captured in Fp. Our conclusion that multiple years of "paired catchment" type data (different</p>

		watershed conditions, same rainfall pattern) are needed to reject null-hypotheses that land cover use effects are neutral. So – we qualify our claim that Fp is an “indicator”, not a “measure” of watershed health. But, so far we don’t know of a better simple indicator.
5	The study addresses a significant point of contention in the literature: the influence of land cover (particularly forests) and flooding at the watershed scale. If the Fp model is properly justified and performs adequately, then it would undoubtedly increase our understanding of the linkages between land cover and flood risk. The benefits of such an approach are clear as it would make for a much more parsimonious model of river flow that would greatly enhance the monitoring and prioritization of specific landscape management decisions.	We thank reviewer for the interest – indeed our target is a key parameter for a parsimonious model that can communicate key functional property of the way a watershed functions, given variable rainfall.
6	That said, the paper requires substantial work to adequately address the points above and may need to be split into two separate papers. I will address the three points I mentioned above in greater detail below. There are quite a few of typographical and grammatical errors in the paper, but I will leave these alone for now as the paper requires substantial work.	As reviewer 1 also suggested splitting the paper along similar lines, we have followed this suggestion.
7	Point 1 In the paper the authors use persistence, predictability, and watershed health interchangeably.	We hope the current text is clear in presenting “persistence” as a first descriptor of what Fp measures (the part of today’s flow that can be counted on for tomorrow, regardless of additional precipitation); 1-Fp controls the part of new rainfall that adds to the stream, and as such controls the predictability of overall flow; the way that Fp is an indicator of watershed health is at a further level of interpretation.
8	One of the key assumptions of the paper and previous watershed rehabilitation efforts is that increasing the presence of natural land covers (particularly forests and wetlands) will restore the natural flood regime with lower peak flows and less damaging flood events.	Actually we took this to be a testable hypothesis rather than a priori assumption.

9	<p>The authors do a good job of documenting previous studies that have illustrated the complexity of the linkage between reforestation and river flows. Moreover, the ability of wetlands and riparian forests to absorb rainfall, slow streamflow, and attenuate peakflows is supported by many studies and is fairly well understood. However, these types of stream corridor ecosystems also require a particular type of disturbance regime that creates opportunities for species specific recruitment processes and establishes landscape and topographic heterogeneity that are critical components of watershed health. These disturbance regimes are often characterized by variable flow patterns with various flood magnitudes required for specific types of ecosystem level processes. Most efforts to create a stable and predictable flood regime have been anthropogenic in origin through engineered based interventions like dams, and retention and detention ponds, which are also some of the primary drivers behind the degradation of watershed health. A perfectly stable flow regime could, theoretically, be established by a highly integrated system of engineered solutions (albeit until they are either overwhelmed by a storm or undermined by system failure) within a very ecologically degraded watershed. Likewise, there could, theoretically, be examples of ecologically well connected and healthy watersheds with fairly volatile and unpredictable daily flow regimes. To overcome this, the authors need to discuss what exactly watershed health means and whether or not a predictable flow regime is the product of an ecosystem service. I could see an argument in which the shape of a storm specific hydrograph within a healthy watershed should be fairly predictable, however, to study this would preclude the advantages proposed by this paper (i.e. the application of the Fp algorithm in data sparse regions). I also agree that more human development and less natural systems generally leads to more flashy river flows as a result of decreased buffering capacity, however this study examines flow rates at daily intervals which washes out the ability to assess this linkage. Or maybe I'm</p>	<p>Thanks for these thoughts. We agree that stabilizing riverflow beyond what the “natural condition Fp” indicates is not without risk for the ecological functions of the river and its biota, and have added some comments to this effect (line 466-471).</p> <p>Quantitative estimates of Fp derived from a limited sampling period do depend on specifics of the actual rainfall. As rainfall is not generally known at the required spatial and temporal resolution to disentangle these relationships, we have to accept that Fp does not only respond to the land use mosaic, but also to rainfall. We have stated this more clearly upfront.</p>
---	---	--

	<p>missing something. The authors do point out later in the paper that Fp of zero (i.e. low predictability) would be the result of erratic rainfall (page 7 first paragraph). This is somewhat confusing because most of the introduction and discussion is focused on using Fp as a way to summarize “complex land use mosaics”. Two paragraphs later the authors state “a decrease of Fp indicates watershed degradation.” So how much of the decrease in Fp is explained by watershed degradation as opposed to just more erratic rainfall? I know the authors say that the GenRiver model is spatially explicit, but this is a little vague. Does this mean that spatial autocorrelation in precipitation is controlled for or that the model is spatially distributed? I understand that to have an Fp equal to 0 would require erratic rainfall, but the authors need to be consistent when describing what proxy measurements that Fp is suitable for.</p>	
10	<p>Figure 1 gets at the interconnection between many different elements that influence the hydrological cycle, and the authors break up the components into ecosystem structure, function, and human land use/perceived ecosystem service. However, I find the figure difficult to navigate and poorly described in the study. The different color arrows with different shades and outlines is one of main culprits of the confusion. The graphic needs to be simplified, it should probably start with rainfall, and terms like “plot-level” should either be defined or removed.</p>	<p>Thanks, we have tried to simplify the graphic</p>
11	<p>Point2. An Fp value ranging from 0-1 essentially represents the buffering capacity of the watershed, but there are also characteristics that influence how rapidly water reaches the stream. In this sense, Fp is only represents half of the picture.</p>	<p>We don't quite understand the reviewer here: Fp quantifies a property of flow dynamics, and integrates over multiple aspects of the subsystems along the rainfall-vegetation-soil-streams pathway. The Fp concept is close to 'buffering' – but this is itself a scale-dependent concept (variability of outflow relative to variability of inflow, depending on time scale and system boundaries).</p>

12	<p>The authors go on to create separate Fp's for each flow pathway, which is probably necessary for large catchments as each flow pathway likely do have large influences over space and time. However, this seems to be overcomplicating a model that was originally being created out of need for greater parsimony. If these pathway specific indices are necessary, then more discussion and justification is required in the text.</p>	<p>We offer the weighted average of pathway-specific Fp values as an additional way of interpreting results, hoping that for some readers this will help understand what Fp is. It is not an essential component of the main argument, and we will state this more clearly.</p>
13	<p>R2.13. The authors use vague language that needs some more clarification. Line 19 on page 7 contains "flow over a sufficiently long period". What is a sufficiently long period? Wouldn't a sufficiently long period wash out the "flashy" fluctuations that the authors are trying to explain with changing land cover/watershed degradation? If the 'sufficiently long period' is preventing what the study is attempting to explain, then I do not see how equation three could be derived. Maybe I'm missing something, but wouldn't the stochastic term represent all unexplained variations in the predicted river flow? Line 28 on page 7 explains that the stochastic term is equal to the sum of peak flows. Couldn't other unpredicted river flows have other anthropogenic origins that contribute to the river flow stochasticity (e.g. dam operations/failure, irrigation, urban water use, etc.)? The authors also mention new variables like Qadd and Fp, try without adequately discussing what they actually represent.</p>	<p>We obviously created some confusion here and have tried to more clearly separate the recursive model of river flow (and its associated Fp) as a "principle", before getting into issues of data and empirical estimates of Fp in defined data sets.</p>
14	<p>Point 3 Figures 2-9 were very readable and for the most part stand on their own, however the table were poorly formatted and vague. Table 1 does not provide land cover proportions by land cover type (other than forest). Percent developed land, existing flood control infrastructure, and population would all be helpful information. Not sure what 'dominant land cover' means. Do these watersheds have a history of damaging floods?</p>	<p>Thanks, we have improved the presentation of Table 1 in part II and defined the terms used, with references to further studies.</p>
15	<p>Why were the parameters in table 2 chosen? The authors do not provide an adequate discussion of how these parameter values</p>	<p>As also commented on by reviewer 1, the reference to the existing manual of the GenRiver model was clearly not</p>

	were estimated. Why were the defaults in GenRiver used for each of the land cover types in table 3? What process or methodology did GenRiver use to estimate these values? What does 'relative drought threshold' mean?	sufficient here, and we have now summarized key equations and assumptions of the model, before using it to illustrate how Fp can be interpreted for alternative land cover scenarios.
16	The use of the word 'some' in table 4 is simply too vague when describing the scenarios. The reader is left wondering what the magnitude of change would be within each of these scenarios. All of the information in tables 1-4 are critical components of the GenRiver model. The legitimacy and accuracy of the model is weak without proper documentation and justification of the underlying model parameters. The authors must correct this if we are to have any confidence in the results.	In Appendix 2 we have added information on the associated land use distributions of the different land use scenario described in Table 4.
17	Table 5 and its corresponding discussion regarding the sample sizes required to reject a null hypothesis is interesting, but not enough information was given to make this section clear. The methodology is clear enough, but the implications were not really discussed. The statement beginning on line 21 on page 14: "In practice, that means that empirical evidence that survives statistical tests will not emerge, even though effects on watershed health are real" is vague and needs some more clarification.	Thanks for the interest in these results – we have brought in some quantitative terms in these overall, qualitative conclusions
18	Lastly, in table 6, the authors provide broken links to the detailed reports of rainfall and river flow data. Moreover, there is very little discussion on the accuracy and metadata of each of these data sources, all of which have different origins.	Unfortunately the website to which links are provided has been off-line for some time, we have double checked all links function again in a resubmitted manuscript.