Reviewer Comments for HESSD 10, 9309-9361, 2013: ***Proposal of a lumped hydrologic model based on general equations of growth – application to five watersheds in the UK***

This paper presents the application of a fairly general model of complex systems to the problem of simulating streamflow. The paper itself is poorly written and does not motivate the study in a meaningful way, however the ideas are new as far as I know. The authors draw some interesting parallels between parameters of the growth model and certain conceptual hydrologic theories, but in the end the model is only applicable to a limited set of conditions (humid environments), and then only some of the time (regimes of rainfall excess). The authors promise to address limited applicability in the future, but at this point I am left wondering why I should look forward to these developments.

The demonstration of the proposed model shows that it may be more resistant to overfitting than IHACRES, however the demonstration is not large enough to draw general conclusions and there is no analysis of significance. The analysis of parameter interaction is exceptionally poor and the analysis of parameter identifiability is only marginally better. The authors do not motivate their model as filling any particular lack in ability of current methods, and note that in its current form it is not suitable for general application. I do not see any science here, where is the question that needs answered by the proposed solution?

I recommend major revisions because I don't see any errors in the math or the logic, and the idea is novel and may have potential (as discussed by the authors in their conclusion) in the future. But I hesitate to offer to review this paper again because the authors have failed to spark in me any interest in their subject or results. However, if four major improvements were made, it could be worth taking another look:

1. Motivate the model as (i) filling a gap in our understanding or ability and (ii) by explicitly illustrating a conceptual link between watershed behavior and all parts of the model.
2. Show convincingly that the model has some advantage – perhaps due to resistance to overfitting during calibration. This will require a larger set of test experiments, and a more robust investigation of parameter interaction and identifiability.
3. Address the issue that it is only applicable to watersheds in humid climates.
4. Reorganize the manuscript so that it presents a logical flow of ideas.

Elaboration on Major Points:

Basically, the argument made in the paper for this model is unconvincing. The authors should use the introduction to motivate us to conceptualize watersheds as being isomorphic with growth models. They do a good job of drawing connections between certain model *parameters* and hydrologic phenomena but the motivation for the model itself is somewhere between completely missing and severely under-developed.

The introduction (both section 1 and section 2.1 – I cannot tell any distinction between material presented in these two sections) meanders through a history of hydrologic modeling, touching on many diverse topics that are unrelated to the contribution of this study. For example: the model does not help with the problem of ungauged basins, it does not provide insight on the philosophy of an "optimal" parameter set (nor does it allow us to bypass this issue), and no muti-criteria objective functions are used in the application study. All of these topics apparently lead into the motivating sentence at the end of paragraph 15 on page 9312: "The problem of the … determination of the correct structure of a model is one of the major challenges in hydrology." This apparently means, to the authors, that we should address the problem by proposing a plethora of new conceptualizations. My question is; <u>"what gap in our understanding does this model fill?"</u> The authors note this on page 9315 paragraph 5, but fail to outline their "modeller's objectives".

In the introduction, just tell us what you are going to do and why, and dispense with the history of hydrologic modeling and description of issues which are not related to your methods or results. The real meat of what we want from an introduction has been replaced by trivia.

Section 2.2 is a history of the mathematical method with no relevance to a hydrologic application. Your contribution – which is the application of an existing method to a different problem, requires a motivation for this application, not a detailed history of the method in diverse fields. As far as I can tell, the only sentence with any relevance to the present study in section 2.2 is the last one about Chow's use of the s-hydrograph – the resemblance that is referred to, however, is not explained in any way. If you were to motivate the use of this model by some science question then add some of this history in the paper - that would be ok. But history does not substitute for a motivation or science question.

Section 3 is the first time we actually get to material that is at all relevant to the paper. In this section, which claims to describe the homology I am looking for, the only justification for the model is that growth models describe time-varying behavior, and streamflow is a time-varying quantity. There is no further justification for the "unbounded function and bounded one". Nor for the "limiting factor" except to note, without reference, that "it is consistent with the observation that a watershed will tend to generate a constant discharge." First of all, what observation? Second, all you are saying is that watersheds are dynamic systems with a steady state. So again, I argue that almost any dynamic systems model will be able to represent time-varying quantities and steady-state behavior. What do growth models give us that we don't already have?

Further, I do not see how the argument of steady state leads to equation (3) – there is no difference between P and PET in that equation, as discussed in the leading text. The multiplication between Q and Qeq is not motivated other than to state blankly that this is how growth models are.

Now, the motivation for the Qeq parameter expression in (4) is good! There is hydrologically-relevant explanation as to where this comes from. I would like a reference for this c_eq idea though, listed in the sentence above equation (4).

In section 3.2.1 the authors are still trying to draw a connection between the growth model (and its classical quantities like carrying capacity and population) with hydrology. But again, no explicit isomorphism is explained. What is system saturation here, and is it different than hydrological saturation discussed in the previous section? (I know the answer, but it is up to you to lay this out for the reader.) The point is, by section 3.2, we should already be speaking in terms of hydrology considering section 3.1 was supposed to be all about isomorphism (or homology). Paragraph 15 is all redundant material.

Related to organization of the experimental sections, First of all, section 4 is titled incorrectly; it not results. It is a set of experiments including methods and results. It would be a lot easier to understand and to absorb the implications of your experiments if you were to lay out the methods first to give us some idea of why we are going to run this particular collection of experiments, and what we expect to learn. The type of disorganization found in this section is why we typically use a methods section and a results section – sometimes it is ok to not take this conventional route, but this was not one of those cases. From the introduction paragraph in this section I have no idea where we are going to go, or why.

For example, why are the experiments not homogeneous in their presentation? The introduction paragraph of section 4 only tells us that we will investigate 4 basins, so if you are going to treat the four basins differently, the reader needs to know how and why before we get started.

From the very beginning of the paper tell us explicitly that the model you develop is only applicable to humid climates, and in the abstract tell us that the model is only capable of reproducing behavior related to excess saturation. These are very severe limitations. I understand that you want to present an "approach" but from a practical perspective this approach is not yet very promising or very general, so it is necessary to stress that from the beginning. If it were me, I would wait to publish until I had a viable model. This is most definitely still just an idea with little to no practical value and I am personally of the opinion that it does not warrant a publication at this point. I know that you disagree, but my question is why don't you finish building your model and resubmit once it's viable – or at least a complete idea – or at least explains something that we didn't already know?

Page 9328, related to "however …" and "strictly speaking …": I agree with the authors that Figure 3 tells us very little about parameter interaction. A better illustration of parameter interdependence would be an OF response surface – actually three of them, one for each parameter pair. In addition to response surfaces, we need to see joint parameter distributions, not just "optimal" parameters. A simple study showing different "optimal" parameters over multiple calibration periods does not really address the issue of parameter identifiability.

Details at Specific Locations in the Paper:

Page 9313 paragraph 15, first line: Darcy was before the mid-twentieth century.

Page 9315 paragraph 25: Bayesian methods do not, in general, make use of the concept of a Pareto optimum. There may be certain techniques founded on Bayesian principles that do, but these would be the exception, not the rule. What exactly is the class of techniques you are talking about here – Bayesian or multiobjective? And why is it relevant to your paper? In addition, please note the relevant page(s) in your book reference (Beven 2012).

Page 9316, first line and on: Related to the words "this type" - what non-statistical approach have you been talking about? Further, all Bayesian methods – by definition – allow for the consideration of epistemic uncertainty. Further, there is no such thing as an "epistemic nature of a residual", a residual is a measure of distance between a two estimates, and may be a result of the fact that there is uncertainty in either estimate. The uncertainty may be epistemic in nature, but the residual is simply a measure of distance.

Page 9317 paragraph 15: "Savageau (1980) …" This sentence seems promising but really gives the reader no information. After reading it I know nothing about Savageau's argument other than that the behavior of complex systems arises due to the interaction of system processes and components – which is an obvious truism. What does Savageau do with the growth model and why should I care? Further, why do I care about single variable vs. multi-variable growth models – what does it have to do with modeling hydrology?

Further, about "what all applications … have in common is … gradual variation … over time governed by a limiting factor." So you are saying that you choose a growth function simply because it simulates smoothed time-varying behavior with a limiting factor. So do an uncountable number of other models – these are basic quality of most dynamic systems models. Regarding the limiting factor: what does that have to do with hydrology and why do I care?

Further, what is an equation of evolution and what does it have to do with the subject of the paper?

9319 paragraph 20: "especific"

Page 9321: It would be good to give a short explanation why (8) is the desired behavior of a watershed with Qeq=0. Since this is a cornerstone in the development of the model, it would be good to let the reader intuit the need for this form rather than just refer them to a reference.

Page 9321 paragraph 5: of course particular cases occur if you choose particular values of certain parameters. I think you mean that certain special cases of (8) have been named and studied. Furthermore, I think some equation numbers are wrong here and in the next paragraph. Also, for (7) to take the form of (8) when Qeq=0 and r defined by (9), it is necessary to use the convention that 0/0=1; this should be stated.

Page 9321, paragraph 10: Seems that you mean equilibrium discharge instead of maximum obtainable discharge. What if we are in the falling limb of the hydrograph? This is the first time "maximal obtainable discharge" has been discussed and I do not see this concept in any of the equations.

Page 9322: Move the "Eq 4" reference to follow the "As has already been discussed" clause. More importantly. Now we are back to the definition of the equilibrium rate. There is no logical flow in this article – we jump back and forth between pieces of the model. It should be structured to:

(1) Lay out an argument for using this model including explaining isomorphism (or homology).
(2) Define the model.
(3) Define the parameters and their conceptualizations.

Please avoid jumping back and forth between concepts unless absolutely necessary. I see no reason why this Qeq related to c_eq concept needs to be discussed in technical detail in two separate sections of the paper.

Furthermore, at this point (Page 9322 paragraph 10) we have no explanation as to why we would want to extend the concept of the Budyko functions. This goes back to pg 9313 – you said there also that we would extend these concepts but not why (I see that your goal is to make this parameter local in time, but this is not explained). Specifically, the distinction between "mean conditions" and instantaneous conditions needs to be made explicit. Again, I recommend putting all this c_eq material in one single discussion and laying it out succinctly from beginning to end.

Page 9324, last sentence of paragraph 8: "have little sensitivity to" or "be relatively insensitive to"

Page 9324, paragraph 15: Is this the same \tau as in equation (17)? This delay parameter should also be dependent on topography, size of the watershed, vegetation, soil types/depths, and other characteristics of the vadose zone.

Page 9325, paragraph 5: How is the logistic equation used to integrate anything? Why are times notated as subscripts here but as functional inputs elsewhere?

Page 9325 paragraph 20: …the mathematical "nature" … Furthermore, I am not sure that I agree with the argument in this paragraph – or at least I don't see it. My interpretation is that this is not an SDDL because (1) it is a deterministic equation (not stochastic), and (2) Qeq is dependent on P, not Q, so the increment of the process does not depend on past values of the process. Related to (1) you have not considered uncertainty at all in formulating the model – only given lip service in the introduction. Please do not introduce and suggest treatment of a model that you do not develop or apply. Related to (2) At most you might consider it an SDE (if it really was stochastic), but not an SDDL.

Page 9326: So even though there is all that talk of multi-objective and application-specific optimization in the introduction all we do is fit a weighted residuals-based objective function. Notice that the "constraints" you mention are just the parametric form of the model, so this optimization is no different than any single-objective residuals-based optimization of any conceptual hydrologic model. The constraints are simply due to the rather involved numerical method for solving the master differential equation, they are not actually related to other modeling objectives. You have simply parameterized a model with three parameters and optimized those three parameters. Also, that is a very long description of Matlab's *fmincon* function ☺ does that lengthy description really tell us anything we need to know to understand your contribution or understand or reproduce the results of your experiment?

Page 9327 – Reference to Table 1: Caption in the table: this table appear to use three word for the same idea: intervals, sub-periods, and sub-series. Is there a difference between these three things? Please choose a single vocabulary and stick with it. It is ok, in the text, to draw connections with Littlewood's vocabulary, but don't do this in the table caption. Or if you feel you absolutely must draw this connection in the table caption, then the Littlewood reference (year) is necessary in the caption. You state in the text that the table uses the "same nomenclature as Littlewood." I am confused. Also conflicting vocabulary in the text near this area.

In the table: What is the "#X"? What is the "#1-8"? What is "Mean #1-#8" and why is it below "#X"? and what is "#1-6" and why are the dates in parentheses? You explain the last one later in the text, but this information should be in the caption.

Also, what is a "calibrated sub-period"? Tell us exactly what you did – don't make us guess.

Page 9328, "Fig 1 shows …": median and quartiles of what? You still have not explained exactly what you did and at this point I have to learn from the figure caption that you mean and quartiles of "optimal" parameters in each of the 8 sub-periods. Also, on here, it would be helpful to mark the parameter values that result from using all of the data for calibration (last line of Table 1).

Page 9328, related to Figure 2: You have chosen to illustrate a *portion* of the best-performing sub-period. Might is also be nice to see scatterplots – or some other explicitly comparative plot to get an idea of where, in the hydrograph, the errors in model estimates are large and small? The CDF plots are nice, but they say nothing about timing.

Also, what is an "adjusted interval"?

Page 9328 "that showed a linear correlation": Why would you only consider linear parameter interdependence – this is basically the type of parameter interdependence that it <u>not</u> interesting from an optimization perspective.

Figure 4 and Table 2 are essentially redundant. My suggestion is that only the figure is necessary because the table is too dense to be meaningful. Just a suggestion.

Page 9329 un-numbered section: Previously the authors have made references to Littlewood, but only discussed their methods in that context – not their results. While reading the previous section I often wondered why they don't compare their results with Littlewood – here I see that they do, but I have had to read through a mess of material to get to the heart of the matter. Here is the first explicit description of Littlewoods methods – after a whole section that drew implicit analogies.

The authors seem to be using "transfer function" (or TF) to mean a model, which is fine but why the redundant vocabulary?

Page 9329 relate to Fig 6: what is a VERHULST 3param? Why is this name first appearing in this figure caption and legend? How many parameters does IHECRAS have and how were they calibrated?

In the text there are three different descriptions of the NS – as "Nash-Sutcliffe", as "the values of variance explained by the model" (Page 9327), and "coefficient of determination" (Page 9329). Perhaps it would be good to have one sentence that explains how you want us to interpret your evaluation statistic and then use consisted vocabulary to refer to that statistic. This is just an aesthetic issue unless I am wrong in interpreting the :coefficient of determination" as equivalent to NS (it is interpreted this way when the regression for the COD is slope 1, bias 0). Or are you now using a new evaluation stat – so that the 10th and 11th bars of Figure 6 are not comparable to all the others and Fig 6 has an incorrect caption?

I have no idea what *experiments* produced bars 10 and 11 of Fig 6. This is a completely separate issue from the fact that I don't know what evaluation statistic these bars plot (NS or a more general definition of coefficient of determination).

It would be nice to split Fig 6 up into separate charts for calibration-period results and evaluation-period results. Really, the evaluation results are of primary interest and they are kind of lost amongst all the other data in this figure.

Page 9329, you refer to the "delay parameter" of IHACRES, but have previously told us nothing about any parameter of this model. This entire description of Littlewood's experiment is lost on me because you have provided no background on the IHACRES necessary to put this discussion in context.

Page 9332: Using "data" from 1986 to 1989. What does "always starting and ending in a hydrologic year" mean?

With respect to all experiments: what initial state value was used?

Page 9334: I am very confused by section 4.4.1. First of all, we are still in section 4, so please don say "as discussed in section 4…" Why is there a bold section header interrupting a single, continuous train of thought (i.e. about optimizing \tao)? I imagine that the rest of my confusion comes from the redundant use of notation \tao.

Page 9335, Section 4.4.3: now we age getting back to model development equations in the "Results" section.

Page 9336, paragraph 10, last line is missing a word. Similarly, the first sentence of paragraph 15 has several grammatical errors.

Page 9336, paragraph 20: where do you get the conclusion that the model is "well-conditioned for optimization?" You have not tested at all for equifinality – in no way do your experiments address this issue.

How is delaying a rainfall equivalent to interpolating to a lower timestep? It seems like the method you use for delay may require a smoothed (interpolated) rainfall time-series or may not. Furthermore, I still fail to see why this constitutes a delay equation.

Also, why should we want a model that delays rainfall rather than explicitly recognizes transit time? The latter is a more intuitive conceptualization.

Page 9338, paragraph 20: This topic is "the" object …

Page 9338 Last paragraph: proved is a really strong word. What do you mean by an "infinitesimal scale"? At an infinitesimal scale we have quantum mechanics, not growth laws. This is the reason that conceptual models like growth models are – with 100% certainty – **not** "correct" models of the system and can never be "proved". These types of models are, and always will be, convenient conceptualizations – useful but never exact, never unarguable, and absolutely never to be taken as fact.