Review of Beven et al. 2020 for HESS by David Milledge

I enjoyed reading this paper and found it both interesting and informative. It is different from the papers that I am used to reading because it reads more as a careful and balanced reflection on a model rather than a report of new findings. But I think it is valuable and will be a useful resource to those who use or are considering using Topmodel in the future as well as those who must make sense of its predictions. I have four major comments, none of which are critical to publication but all of which I feel would improve the paper. There are then many minor comments and suggestions most of which are either typos or suggested rewording.

Major comments

Assumption A1

The assumption that "that the storage for any given value of Sbar is configured *as if* it was at a steady state with a steady homogeneous recharge rate (L88)", and its implications comes up in three different places within the article. It is an important point because it relates to a central assumption and one of the primary perceived weaknesses of Topmodel. I found this discussion particularly helpful in my thinking on Topmodel but I also found it confusing in places.

On my first read through I felt the first discussion of A1 on L88 didn't give enough detail. In particular I was confused by the language around "configurations" and "configured as if". I didn't understand how Sbar could be configured as if it was at steady state (L88) nor how configurations are dependent on storage (L89) nor how the two ideas related to one another. Did this mean that Sbar is varying only slowly in time? How slowly does it need to be? What controls the sensitivity of Sbar to rainfall and what is the sensitivity of the saturated zones to Sbar?

The later treatment of the assumption (L320) is more detailed and I understood this section better. It might be enough just to point to the later section at L89 for more detail. In this L320 paragraph I still struggled to understand what you meant by configuration. I understood it to mean that: 'the two-dimensional phreatic surface over the flow strip is that which would result from steady recharge over that flow strip'. However I wasn't confident in my understanding so clarifying this would be helpful. The main outstanding question for me at the end of the paragraph was: how close to 'as if' is near enough? You mention this with reference to Kirkby (1997) but a more complete restatement of his examination and findings would be useful here.

Assumption A1 is revisited on L755, and I found this the clearest expression of the steady state assumption within the paper. It may be that the other sections had laid the groundwork but I think you should consider re-stating this expression earlier in the paper.

Assumption A2

Topmodel uses tan(beta) to calculate lateral subsurface flux (L49 and equation 1). Others, usually modelling steep landscapes, have used sin(beta) to make the same calculation (e.g. Montgomery and Dietrich, 1994; 2002; Borga et al., 2002; Chirico et al., 2003). In some cases they explicitly claim that there is a choice between "the original ln(A/tan(beta)) or the more physically correct ln(A/sin(beta))" (Montgomery and Dietrich, 2002, p2). It might be helpful to respond to this claim, perhaps explaining why the difference, whether you consider one more physically correct than the other and if so what the implications are for situations in which they can or should be applied.

Assumption A3

It would be useful to have a longer discussion of whether the exponential transmissivity function is an assumption introduced by the authors (as is suggested L91-2) or one that is required within the derivation (as Kirkby (1997) seems to suggest). There are clearly advantages to being able to use alternative transmissivity functions, so it would be useful to know more about any possible disadvantages. It would be particularly useful to comment on how this might impact the validity of other model assumptions (e.g. L413) and the sensitivity to these assumptions (e.g. L360-2)

You do touch on this at L413 "might also preclude…" however, you say might rather than would and I am not clear what you mean by "implicit redistribution of subsurface storage". Do you mean that A1 would not be consistent with non-exponential transmissivity functions? Kirkby (1997) seems to argue that the choice of an exponential transmissivity function is required to satisfy the integration (though I could have misunderstood Kirkby here). Do the authors of this paper find that argument convincing? If so what does it mean for the alternative profiles (e.g. Ambrose et al., 1996)? If not then where do you differ from Kirkby (1997)?

Connectivity and run on
The argument that small channels may connect apparently disconnected saturated areas (L350) is not clear to me. In particular mention of small channels at the start of the sentence seems to contradict the end of the sentence. If I understand what you mean here, I think it might be clearer to talk about geomorphic / landscape evolution controls on where channels begin (e.g. Montgomery and Dietrich, 1988). The places where this will break down are those where some other landscape property gets in the way e.g. lithology and rock strength in parts of the Yorkshire Dales. If instead this is a suggestion that the majority of run-on passes from patch to patch and reaches the river as overland flow, then I think more support for the argument is needed. I haven't seen anyone demonstrate this.

Minor comments
L29: parsimonious which – word missing
L35: suggesting ... constant - has this observation now been more widely reproduced? A comment on this would be useful here.
L50: transmissivity - i.e. the depth integrated permeability? It might help to say this explicitly.
L50: represented … units - This is not clear to me: I think that K is a profile averaged permeability. However, for the following equations to work I think that K must be independent of S and therefore independent of depth. I think this is worth saying here.
Eqn 2: S appears on both the RHS and the LHS. I think that is a typo.
L57: just saturated - just seems a strange word here. Do you mean at the onset of saturation (implying that the timing matters) or the position of the water table (i.e. fully saturated or saturated so that the water table is at the surface)?
Eqn 3: If the soil is "just saturated" shouldn't S=S_0? Also, S_0 hasn't been defined.
Eqn 4: Kbar and A are not defined yet. I think that Kbar is catchment averaged permeability i.e. 1/A*integral(K,dA). If that is the case it seems strange to write it this way because Kbar in lambda cancels the Kbar in the denominator of the equation for Sbar.
L50: K is defined as a permeability but I think for the dimensions to work it should be hydraulic conductivity here and elsewhere.
L68: these equations – this would be clearer as 'equations 3 and 4'.
L89: storage changes - comma needed after changes
L90: wet, … moderate slopes - an indication of the range for these might be helpful if it were possible.
L91: where soil permeabilities increase with saturation - do you mean transmissivities due to hydraulic conductivity decreasing with depth or that permeability (hydraulic conductivity) itself is changing?
L100: just saturated – as for L57.
L102: as – I suggest 'saturation now is:' or 'saturation as:'
L113: Qb … along the channel - It is not clear from this whether Qb is a single value assumed to be distributed along the channel but without explicitly accounting for this distribution or whether it is spatially variable.
L114: see Beven - I think it is fine to point to this more detailed derivation, though it is probably worth adding a page reference (p214). The one difficulty I have there is in Equation B6.1.16 where upslope area per unit contour length is used but this cannot be true for anything other than the outlet cell. I think that a in B6.1.16 should instead be defined as the unique upslope area (that does not overlap with upslope area for any other channel length). This is very intuitive when a channel is viewed as a line (i.e. inputs from the banks of a river) but not as an area with an upslope boundary defined by a contour (when inputs from upstream as well as from the banks should be included).
L135: calculated outputs - Is there a reference associated with this? Fine if not but useful if so.
L149: both showed changes over time - I'm not clear, do you mean that the components have changed over time or that the represented time variable processes?
L151: later these stores were integrated – this would benefit from a reference.
L163: later versions of Topmodel – this would benefit from a reference.
L217: The BK79 paper … flashy catchment – this sentence doesn't make sense to me.
L234: Whilst… - this sentence reads as a fragment.
L235: but were always clearly identified – I think this relates to the assumptions but it is not clear, perhaps there is a comma missing.
L256: including – should be included
L285: but t more – typo, remove t
L332-3: as the catchment wets and dries – An additional sentence explaining why would be helpful here
L336-7: This should not be a surprise at Tarrawarra … impermeable subsoil – this would benefit from a reference
L346: connectivity of both surface and subsurface flows - It seems as though there are two issues being discussed here, disconnection and response timescale. 1) Even if the signal can propagate rapidly in the presence of a water

table configured as assumed by Topmodel it may not propagate if that assumed configuration breaks down resulting partial or total disconnection of some upslope area. 2) Even if the water table is configured as assumed in topmodel the signal propagation may not be sufficient to generate steady state like response to rainfall. I thought Barling's modification was largely focussed on the second problem.
L348: This paragraph, and particularly this sentence contains two ideas that might be more easily understood if separated: subsurface and overland flow connectivity.
L348: always connectivity – do you mean of subsurface flow?
L353: can be represented in… - It has also been represented in the network index, a modification to the topographic index to identify areas of disconnected saturation within Topmodel, the runoff from which is then assumed to entirely reinfiltrate as run-on (Lane et al., 2004, 2009, Lane and Milledge, 2013).
L355: wetter vegetation patterns – perhaps a reference here?
L361: particularly – this was not clear to me. Do you mean that Kirkby shows it for the particular case of exponential transmissivity or that it is particularly true for the exponential transmissivity profile case? I think it is the latter. If so, what are the implications for other assumed transmissivity profiles?
L365: parallel to the surface - How much empirical support is there for assumption A2? You provide a statement on this for A1 and a similar statement here would be useful.
L366: important - It might be helpful to say important to what. This may be tricky given the wide range of applications of Topmodel and its derivatives and I don't think an exhaustive list of 'important to...' is necessary here. However, an indication that both the scale over which the assumption is not valid and the question being addressed contribute to whether or not violation of A2 will be important. For example, I could imagine bedrock exfiltration being very important to local patterns of saturation deficit on the scale of metres to tens of metres (and therefore to pore pressure at scales associated with landslide triggering) but much less important over scales of hundreds to thousands of metres and therefore for runoff generation.
L370: (it might… values) – Has anyone done this? If not it might be worth flagging that more strongly. One of the valuable things about this paper is the opinions of the original authors on what has been done and what is yet to be done from their perspectives.
L374: values tend to be high - Is that a commensurability issue associated with scale dependence in the measurement? If so it might be useful to comment on that here.
L388: A further criticism… unsaturated zone – a reference is needed here. Also, this doesn't seem to get the detailed response that other criticisms received. Is the following sentence a general conclusion to the section or specific to this criticism?
L425: small – adding the catchment area would be useful here.
L429-30: This was … number of storms – mixed tenses in this sentence
L431: would predict – should this be: would predict (i.e. it is possible) or predicted (i.e. this was the prediction)?
L463: predicted modelled – I think only one or the other of predicted or modelled is needed here.
L463: storage deficits to water tables - You don't give this much treatment. Can you point to a reference and give an indication of the magnitude of the issue, as written it appears a straightforward exercise.
L466: represent different quantities - An example here might be helpful. E.g. modelled water table depth at 2 m resolution compared with observations from a single well within that 2 m cell.
L467: this is a particular problem - Is it commensurability between predictions and observations that is a particular problem in this case or commensurability between the input parameter (spatially uniform T0) and the local conditions (heterogeneous hydraulic conductivity in 3-D)?
L470: would not expect the predictions to match… - Has anyone provided an indication of the likely magnitude of the commensurability errors? If so how do these compare to the misfit observed in these studies?
L505: do not explicitly consider convergence… - I disagree: the indices of Barling, Woods, explicitly consider convergence etc through upslope area as the topographic index does, that of Berne uses a hillslope width function and that of Hjerdt was conceived as a replacement for local slope which has subsequently been used within a modified a/tanB topographic index.
L521: a rumour - this moves away from the authors' own experiences and is difficult to support. Consider rephrasing or removing the last two or three sentences since I think you can make your point without this statement.
L528: that contributing – should be that the contributing
L601-5: One of the… A second result… - A reference that you associate with each of these results would be helpful at the end of each sentence.
L605: longer time steps – could you quantify longer timesteps?
L626: risk - Could you expand this to replace 'risk'? You haven't previously defined risk and it may mean different things to different people.

L646: A further step - Should this be 'an alternative step'? Dynamic Topmodel predates the distributed version and does not adopt its surface flow treatment.
L696-7: some interest - references to this interest would be useful here, unless the following sentence is an example.
L729: topographic index - this actually uses the network index, a modified version of the topographic index that accounts for downslope connectivity of saturation excess overland flow (see: Lane et al., 2004; 2009).
L761: some of the other assumptions – it would be useful to be explicit about which assumptions here.
L762: hydrological important – should be: hydrologically important.
L767: cease to allow – I think there is a comma missing here.
L780: This approach … much meaning – it is not clear to me what this comment means.

References

Borga, M., Dalla Fontana, G. and Cazorzi, F., 2002. Analysis of topographic and climatic control on rainfall-triggered shallow landsliding using a quasi-dynamic wetness index. *Journal of Hydrology*, *268*(1-4), pp.56-71.
Chirico, G.B., Grayson, R.B. and Western, A.W., 2003. On the computation of the quasi-dynamic wetness index with multiple-flow-direction algorithms. *Water resources research*, *39*(5).
Lane, S.N., Brookes, C.J., Kirkby, M.J. and Holden, J., 2004. A network-index-based version of TOPMODEL for use with high-resolution digital topographic data. *Hydrological processes*, *18*(1), pp.191-201.
Lane, S.N., Reaney, S.M. and Heathwaite, A.L., 2009. Representation of landscape hydrological connectivity using a topographically driven surface flow index. *Water Resources Research*, *45*(8).
Lane, S.N. and Milledge, D.G., 2013. Impacts of upland open drains upon runoff generation: a numerical assessment of catchment-scale impacts. *Hydrological Processes*, *27*(12), pp.1701-1726.
Montgomery, D.R. and Dietrich, W.E., 1988. Where do channels begin?. *Nature*, *336*(6196), pp.232-234.
Montgomery, D.R. and Dietrich, W.E., 1994. A physically based model for the topographic control on shallow landsliding. *Water resources research*, *30*(4), pp.1153-1171.
Montgomery, D.R. and Dietrich, W.E., 2002. Runoff generation in a steep, soil-mantled landscape. *Water Resources Research*, *38*(9), pp.7-1.