Manuscript: hess-10-13191-2013: Do land parameters matter in large-scale terrestrial water dynamics? – Toward new paradigms in modelling strategies

Major remarks

The authors present an interesting study on simulating monthly water dynamics over small catchments in Europe. They compared results from land surface models and global hydrology models (both are referred to as GHMs in the following) to a so-called data driven approach where they used Random Forest method, a machine learning tool, to calculate runoff. Their results highlight that substantial parts of terrestrial water dynamics are controlled by atmospheric forcing, and it is suggested that land parameters play a negligible role.

Some major clarifications with regard to the characteristics of the Random Forest method used to create the CLPH model are necessary. How I understand the method from what is written in the paper is that it basically creates a discharge solution Q for the point x from forcing variables and discharge observation from the surrounding 9 EWA catchments (n=10). Thus, this means that the method creates an impulse response function h that yields discharge (or river runoff) from the forcing variables. This is rather similar to what has been done in rainfall-runoff modelling for many years. The new point in the present study is that this function is created from surrounding areas of x and not at x itself such as it is commonly being done in rainfall-runoff modelling. In both cases, any parameter dependencies are hidden in the 'black box' of the impulse response function. In rainfall-runoff modelling, it is no surprise that the calibrated function h usually yields better runoff estimates than an explicit model. In the present study, the CPLH model or the respective functions h also mostly yield better results than the GHMs, but here h is based on surrounding areas. This can be interpreted that the hidden parameters are similar for neighbouring catchments, which is a concept that has been previously applied in regionalization studies (Some of those studies and results should be cited and discussed in this respect). It is very interesting to see that this concept actually works quite well on the European scale for small catchments. But it does not mean that parameters are not important. This would only be the case if you train/derive a function h in one area and apply it in a totally different area, e.g. deriving the function in the Mediterranean and applying it in Northern Finland, or by using one function h for all catchments. I doubt that in this case the GHMs would still be generally worse than the CPLH model.

The issue of scales (spatially and temporally) on which the study results are valid also need to be better highlighted. Given that my interpretation of the RF method is correct, the results indicate that the "local" variations of parameters are not important (thus, regionalization is applicable). But this does not mean that "regional" parameter variations are not important, e.g. variations from the semi-arid Mediterranean region to the snowmelt dominated areas of Northern Finland. Here, the definition of "local" and "regional" should be made clear and quantified if possible. If regional variations would not be important either, then the application of globally constant parameters would be sufficient, but many studies in climate and land surface modelling have shown that moving from globally constant to heterogeneous (regionally varying) parameters led to improvements in corresponding model simulations. It also should be noted that GHMs are usually made to be applicable for larger catchments and not necessarily adequate for small scale catchments with the size of one (or a few) 0.5° model grid box (boxes). Thus, there is already an inherent scale limitation of the GHMs itself.

The use of the terms runoff and discharge is not fully consistent and partially misleading. Especially it should be pointed out that the lateral transport of water (which transforms runoff into discharge) is not considered in the present study. For discharge it is known that the slope is important for the travel times of water. As only small catchments are considered on monthly time scales, travel times do not play a role in the present study. But it needs to be highlighted that the results of the present study with regard to the non-importance of locally varying parameters cannot be transferred to large catchments as here the travel times of water in the river network become important, and those strongly depend on locally varying parameters such as the slope and on the presence of wetlands, lakes and artificial reservoirs in the river network. Consequently the CPLH-RFM does not provide a reliable basis for estimating river discharge from Pan-European rivers, such as noted on p. 12304 – lines 11-12. It can only be used to estimate the total runoff for these catchments as the lateral transport is not accounted for. The difference between runoff and discharge also needs to be properly taken into account in Fig. 3, Appendix C and Table C1.

In summary, I suggest accepting the paper for publication after major revisions have been made.

Minor Comments

In the following suggestions for editorial corrections are marked in *Italic*.

Title

The second part of the title, 'towards new paradigms in modelling strategies', may lead to the impression that new modelling strategies are introduced in the paper, which is not the case. Thus, I suggest modifying the title.

p. 13193 – line 14

… for *discrete* land …

p. 13195 – line 8

… *where features* of the atmospheric …

p. 13196 – line 10

Note, *however, that Skøien* et …

p. 13197 – line 2

It is written:

… using Morans / ….

What does that mean?

p. 13201 – line 22

… catchments *were* first …

p. 13202 – line 13

… inclusion *of* locally …

p. 13210 – line 11-12

It is written:

“However, catchment-scale hydrological modelling is usually based on precipitation (the sum of rainfall and snowfall) and temperature only,”

This statement is too general, as the scale and model type need to be clearly defined. For rainfall-runoff modelling types this is true, but with regard to the WATCH GHMs, it is wrong, as here most of the GHMs use more forcing variables than precipitation and temperature. Here, it should be noted that even though PT forced GHMs may yield reasonable results for today’s climate, they potentially fail for future climate conditions, especially over tropical and sub-tropical areas as pointed out by Hagemann et al. (2011).

Hagemann, S., Chen, C, Haerter, J. O., Gerten, D., Heinke, J., and Piani, C.: Impact of a statistical bias correction on the projected hydrological changes obtained from three GCMs and two hydrology models. J. Hydrometeorol. 12, 10.1175/2011JHM1336.1, 556-578, 2011.

<u>p. 13210 – line 22-23</u>
It is written:
“The fact that BIAS and BIASlog computed for the RFM are hardly distinguishable from the LSMs/GHMs shows …”

It should be noted that Fig. B1 shows that one RFM model (unfortunately I can’t read the legend) is clearly worse than all (or almost all) GHMs.

<u>p. 13212 – line 4-5</u>
This is an important result that should not be hidden within the appendix. It also supports findings of Haddeland et al. (2012) who investigated the impact of bias correcting other forcing variables than precipitation and temperature on the simulated hydrology. In this way it adds to the discussion whether the bias correction of these other forcing variables is really necessary for hydrological applications that has also be taken up in Hagemann et al. (2013).

Haddeland, I., Heinke, J., Voß, F., Eisner, S., Chen, C., Hagemann, S., and Ludwig, F.: Effects of climate model radiation, humidity and wind estimates on hydrological simulations. Hydrol. Earth Syst. Sci.. 16, doi:10.5194/hess-16-305-2012, 305-318, 2012.

Hagemann, S., C. Chen, D.B. Clark, S. Folwell, S.N. Gosling, I. Haddeland, N. Hanasaki, J. Heinke, F. Ludwig, F. Voss, and A.J. Wiltshire, 2013: Climate change impact on available water resources obtained using multiple global climate and hydrology models. Earth Syst. Dyn., 4, 129-144, doi:10.5194/esd-4-129-2013

<u>Table C1</u>
Why you do not show the multi-model mean of the GHMs?

<u>Fig. 2, 3,4, 6, B1</u>
Many figures are too small so that it is very hard to read the legends.