Response to reviews

Reviewer #1

We would like to thank Lieke Melsen for her constructive comments. We will try to improve on the raised issues.

The first thing that struck me when getting introduced to the catchments that were used in this study (Table 1) is that the water balances are not closing. For the Hubbard catchments this is hard to check since only PET is given and AET will be lower, for the HJ Andrews catchments, on the other hand, water is 'lost'. Of course it is not a big surprise that a water balance is not closing, given the uncertainty in the observations, but it becomes tricky when the water balance is used to determine the moisture storage capacity (although you could say that this is also the case for hydrological models that are based on the water balance and that are calibrated on such data). The potential 'disinformation' in observations might influence your estimation of Su,max. I would at least expect a discussion of this potential source of uncertainty, and an estimate of the influence on the results.

This is a very valid point. We relate the fact that the water balance does not close mainly to the calculation of the potential evaporation, which here, due to data availability, was estimated from temperature only. We will add a paragraph in the discussion on the consequences of these uncertainties for the estimation of S_R.

Lines 7-18 on page 10 show a difficulty of the water-balance method to identify Su,max; you have to assume no storage change. The Introduction describes the importance of flexible Su,max for changing conditions; e.g. land-use change and climate change. And this is where it becomes difficult; under a changing climate (no steady state conditions) you can no longer assume that there is no storage change. In other words; to me it seems that the method to identify Su,max based on the water balance is not applicable in a changing climate.

We agree with the statement that under changing conditions storage may change. Nevertheless, in the applied method the water balance is merely used to derive an estimate of average transpiration rates. Therefore, we argue that under changing conditions, this estimate is an upper limit of the actual transpiration, whereas in reality it may be lower.

In addition, a long-term water balance would not reflect the yearly variations in climate, whereas rather short term water balances may be influenced by storage changes. This is also why, in a trade-off and to keep the effects of storage change as low as possible, the water balances over 2-year periods were used. To substantiate this, to put into context and to assess the effect of storage change, please see Figure 1 below, where, for comparative reasons, we additionally estimated Su,max using a 5-year window to further reduce the influence of storage changes. It can be noted here that the green shaded area, representing the water balance-based estimates, is flatter compared to the results obtained with the 2-year water balance (maximum 500mm compared to 600mm in Figure 4 of the manuscript). This is due to more averaging by taking a longer period for the water balance estimation. In spite of that, the general patterns hold, and in our opinion supports our results.

Eventually, we would like to point at the results obtained in the undisturbed reference (or control) watersheds, in Figure S8 of the Supplementary Material. These results are obtained in absence of any land use change, and thus reflect only the changes due to climatic variability (and are thus a proxy for climate influenced inter-annual storage changes). The different pattern compared to the deforested catchments then indicates the isolated effects of storage change due to deforestation and thus transpiration (under the assumption that both control and deforested catchments were subject to the same climate variability). Thus, we would argue that the changes in storage that may occur, are relatively small compared to the annual fluxes of precipitation and discharge.

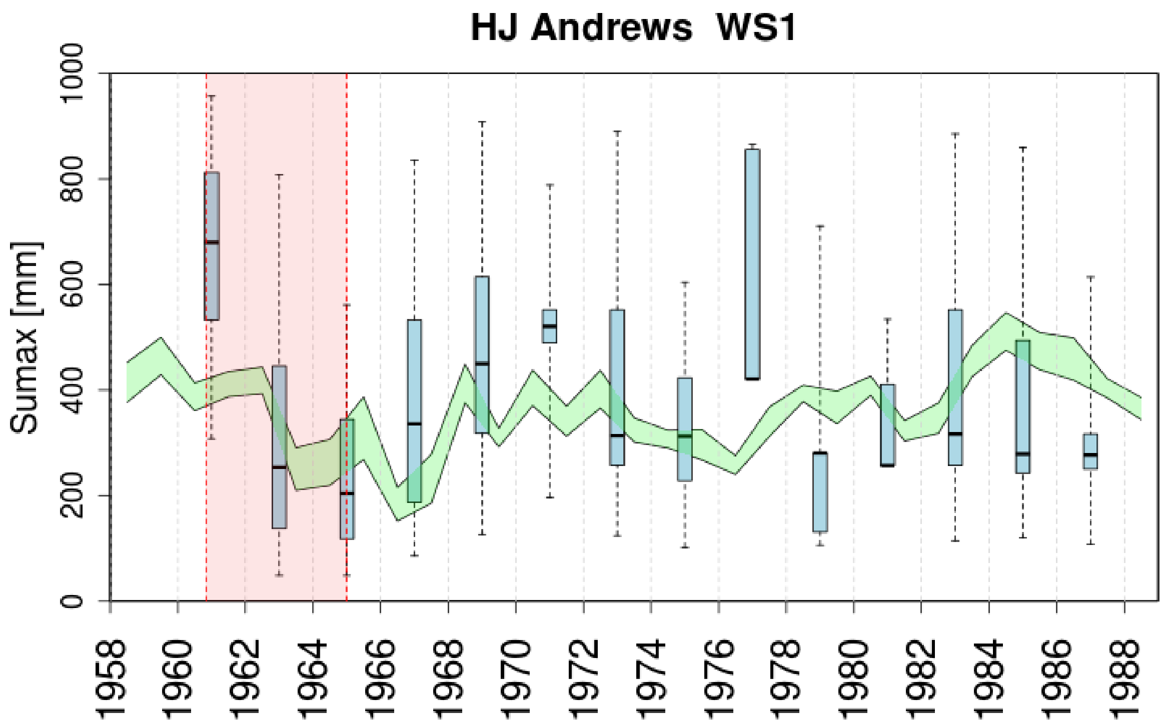

Figure 1. Evolution of root zone storage capacity $S_{R,1yr}$ from a 5-year water balance-based estimation (green shaded area, a range of solutions due to the sampling of the unknown interception capacity) compared with $S_{u,max,2yr}$ estimates obtained from the calibration of FLEX (blue boxplots) for HJ Andrews WS1. Red shaded areas are periods of deforestation.

As a proof of concept, a model was included with a dynamic Su,max, which was calibrated by expert-eye to fit the SR1yr-values that were obtained by the water balance method. I agree that a proof of concept is a first step in increasing the process representation in hydrological models. I would, however, appreciate it if the authors would provide the reader with some suggestions on how to incorporate a dynamic Su,max 'more correctly' in hydrological models. Generally, I am in favor in improving realism in hydrological models, but, extra parameters imply extra uncertainty and the uncertainty should not overwhelm the (hopefully) improved model efficiency. The water balance method seems not feasible in non-steady-state conditions. Do the authors have any suggestions on how to include a dynamic Su,max, or suggestions on observations that could help in this respect?

We would like to suggest simple conceptual formulations of growth dynamics, similar to the growth function applied in this case. This would lead to the addition of, at most, three new parameters. These could be free calibration parameters, but we agree that this may lead to additional uncertainty. And even though the water balance method may only give an estimation of the dynamics of the root zone storage capacity, this method may prove valuable to derive at least some information about the *shape* of the growth curve.

It can also be noted that transpiration estimates are derived from the water balance in this case, but there are also (remote-sensed) products available to estimate the transpiration. In this way, issues with water balances that may not close are fully avoided.

Based on the remarks above, I would suggest to add a separate section to place the results in context (a sort of Discussion, but then different from the one that is included now in the Results section).

We will add a separate section in the discussion about the uncertainties that are introduced by 1) data used in the water balance, 2) storage changes affecting the water balance. In addition, we will elaborate in Section 4.4 on how to explicitly apply our findings in conceptual modelling.

I know that in the work op Gao and de Boer-Eusink it is shown that climate mainly dictates Su,max rather than the soil. It is, however, maybe valuable to have a look at some of the work of Ilja van Meerveld, who investigated the effect of land use change on soil properties, where it is discussed that the hydraulic conductivity changes as a result of land use change. Could it be possible that the changes in Su,max that you find could actually be assigned to the wrong assumption that the Ksat does not change after land-use change? There are, of course, more parameters in a hydrological model besides a constant moisture storage capacity, that might actually not be completely constant. How can you be sure that the effect you find can only be assigned to the root zone storage and not other parameters?

Indeed, there is no absolute certainty that other parameters are not affected by the land use change. Nevertheless, when vegetation is removed, it is not inconceivable to assume that the vegetation-related parameters are considerably affected. This can also be seen from the posterior-distributions of the other parameters, see the Supplementary Material. In the 2-year window calibration, all parameters were left for calibration, and they all had the freedom to change over time. Nevertheless, the root zone storage capacity showed the most dynamical character, whereas others remained more constant in time. In addition, we would expect that changes in hydraulic conductivity are tightly linked to changes in porosity. In other words, an increase of porosity is not unlikely to decrease the flow resistances and thus increase Ksat, while simultaneously reducing the storage capacity. It must also be noted that hydraulic conductivity Ksat cannot be compared directly to any of the catchment scale conceptual model parameters applied here.

In the calibration of the four hydrological models, two Kling-Gupta terms and the Volumetric Efficiency are used as objective function. As far as I can see, the volume error is already included in the KGE by means of the bias (Beta-term), which would mean that in your calibration strategy, you put extra emphasize on the volume error by explicitly including this term twice (or actually, three times since you use KGE twice). Why is that justified?

This is a valid point; we will compare the outcomes with a calibration based on a combination of KGE and logKGE to test how much this influences our results.

In your dynamic model, you included extra parameters to describe Su,max, and concluded that it improved the model performance for several indicators. How can you make sure that this improvement is caused by including this process in the model? I would say that for many models you can obtain a (marginal) improvement in model performance by including an extra degree of freedom (an extra parameter), independent of the process that this parameter describes or the realism of the parameterization.

To avoid this, both model approaches were given the same number of degrees of freedom. In other words, both models had the same number of free calibration parameters. This is why the growth functions were fixed, and not left for calibration.

I think the research questions in the summary do not exactly reflect the research question in the manuscript (Line 1-5 on page 6).

We will rephrase it to be more consistent throughout the manuscript.

Review #2

We would like to thank Anonymous Referee #2 for his/her feedback. We will try to improve on the raised issues.

General comments

In general, I find the paper too long. Maybe some details of the methodology can be moved into the Supplementary Material.

Agreed. We will shorten some parts of the manuscript.

I suggest to be more precise in the title. First, ending the title by "under change" seems quite strange to me. Is it still land use change, or climate change or other ? (same remark at line 10 of page 2). Then, "predictions" is too vague because it can be applied to many processes (prediction of discharge, of flood, of vegetation dynamics...). In addition, more discussion on the potential applications with this kind of method is needed in the conclusion and perspectives.

We rephrased the title to: "The evolution of root zone moisture capacities after deforestation: a step towards hydrological predictions under land use change?". In addition, we will add a discussion on practical applications of the method in conceptual modelling (also suggested by Referee #1).

The results and the figures, which include many hydrological signatures, are not always simple to read and to analyze. Then, the interest of the discussion can be lost during the reading of Section 4. Thus, I would recommend to split this section in 2 sections to distinguish Results and Discussion.

We decided to merge the results and discussion in order to avoid repetition and to make the article more concise. We still prefer to keep it like this, also with regard to the first comment (the paper is still rather long). Nevertheless, we will have a critical look at the figures and discussion, and will try to clarify wherever we can.

Specific comments

Abstract

1/ "long-term data" => you can be more precise

2/ line 24 of page 2: "better representations of high flows and peak flows" => what about the low flows ?

1./ We changed it to "long-term data (30-40 years of observations)"

2./ The low flows improved for the Hubbard Brook catchments, whereas the low flows did not show improvements in the HJ Andrews catchment. See also page 24, line 13-26.

Introduction

3/ To be more precise, the vegetation partitions first precipitation into interception, stemflow and throughfall. Then, the fraction of rainfall that reaches the surface is partitioned into evapotranspiration, drainage and also surface runoff.

4/ <u>line 28 of page 3</u>: the year is missing for Vose et al. and also in the References section.

5/ <u>line 10 of page 4</u>: interception/soil evaporation/transpiration and surface runoff/drainage

6/ <u>line 21 of page 4</u>: "system" is unclear. Please reformulate.

7/ <u>lines 30-32 of page 4</u>: The sentence is difficult to read. Please rewrite.

8/ lines 6-7 of page 5: SR has already been defined in page 3, line 15. The best is to combine "sometimes also referred to as plant available water holding capacity" with the text in line 15 of page 3.

9/ lines 18-21 of page 5: the sentences are very unclear. Please reformulate.

10/ lines 3-4 of page 6: words are missing in the 2nd hypothesis formulation, please check.

3/ We fully agree, and we will rephrase the first sentence to be more correct.

4/ We corrected this.

5/ We rephrased it into "runoff components and evaporation", as we tried to lump the terms together that you refer to.

6/ We changed it to "hydrological system"

7/ We rephrased this.

8/ We changed this and placed the text at page 3, line 15.

9/ We rephrased this.

10/ We checked and rephrased the sentence.

Section 2

11/ In each sub-sections, the references to Table 1 for watershed characteristics should be merged and written once in the section, just before sub-section 2.1. Then, the references at lines 12, 19-20 of page 6 and lines 1-2 of page 7 can be removed.

11/ We agree with the suggestion and changed this.

Section 3

12/ lines 14-17 of page 9: For long-term mean variables: Et => Et. The same for Q and Ep.

13/ line 5 of page 10: "obtained by equation 6" => "obtained by equation 7"

14/ lines 7-9 of page 10: this is a strong assumption, especially under climate change where the water storage changes. This point should be more discussed when the method based on the water balance is applied.

15/ line 11 of page 11: "FLEX-based model" => "The FLEX-based model"

16/ line 1 of page 12: this process is not represented in Figure S2.

17/ line 9 of page 12: what are the fluxes ? Moreover, transpiration is indicated in the text but "Evaporation" is written in Figure S3. Please, check the coherency between the text and the Figure.

18/ line 11 of page 13: what is n ?

19/ line 4 page 14: Z95 should be Zp95

20/ line 2 page 16: "Table 2" => "Table 3"

12/ We changed this.

13/ We changed this.

14/ We agree with this and will add a discussion on this (see also the Response to Referee #1)

15/ We changed this.

16/ This is not the case for the current set-up. We will remove the sentence.

17/ We rephrased it to make it more consistent.

18/ n is a weighing exponent. We will clarify this in the text.

19/ We changed this.

20/ We changed this.

Section 4

21/ lines 23-24 of page 17: this is not particularly obvious in Figure 2f.

22/ lines 20-21 of page 24: I do not see this improvement on Figure 10, maybe due to the scale of the plots.

21/ We do agree that the pattern is rather variable over time, but comparing the highest peaks before deforestation with the peaks after deforestation show that the values were higher before deforestation. The same applies to the lower values. Calculating the mean autocorrelation before deforestation and after also confirm this; 0.65 before deforestation and 0.58 after deforestation.

22/ More specifically, we are referring to the parts of the hydrograph at the end of June until August. Please note the white space between observation and model in the case of a constant root zone storage capacity, whereas for the dynamic model they overlap.

Table/Figures

23/ Table 1:

- *I would add a column for the abbreviations of each catchment, as used in figure 9 (see my comment hereafter for the whole text).*

- *“Precip” should be “Precipitation”.*
- *what is “Pot.” ? It is the potential evaporation?*
- *remove “%” from 87% in the last line.*

24/ Table 3: the reference for Jothityangkoon et al. (2001) is missing in the References section.

25/ Figure 1: in the label of y-axis, “P” should be “PE”

23/ We agree with the suggestions/corrections and will change it. “Pot” refers indeed to potential evaporation.

24/ We corrected this.

25/ We corrected this.

Supplementary material

26/ Table S1: please check the Imax values (Min=Max=0 !)

27/ Figure S2:

- *replace “Snow” term in the figure by “S”.*
- *Peff and interception are not represented in the Figure.*
- *q3 should be replaced by q2 in the figure.*

28/ Table S2: the wilting point cannot be higher than the field capacity. Please check the max values.

29/ Figure S3:

- *replace “Snow” term in the figure by “S”.*
- *q3 should be replace by q2 in the figure.*
- *Q should be replace by Qf.*
- *what is dq ?*

30/ Figure S4: the surface runoff is missing.

26/ This should be 0 – 5 mm

27/ We changed this.

28/ These percentages should be added up (they do not represent the actual wilting point and field capacity). Thus, when wcep is 0.2, and wcfc 0.5, the wilting point is at 0.2 of the soil depth and the field capacity at 0.7 (0.2+0.5).

29/ We corrected this and added the missing description of dq.

30/ Correct, this model structure does not take overland flow into account.

In the whole text

–choose between "parameterization" and "parametrization"

We changed it throughout the whole manuscript to "parameterization"

–I suggest to use the abbreviations of the catchments in the text, as used in figure 9. It will facilitate the reading of the paper.

We will consider this, though this is just a matter of taste. Personally, a text with too many abbreviations may also become harder to read.

–there is a confusion all along the text when the term "evaporation" is used. The term "Evapotranspiration", which is the sum of soil evaporation, interception evaporation and transpiration, is more adequate.

We tried to be consistent throughout the manuscript and refer to evaporation when we mean all the evaporative fluxes. We actually believe that the term "evapotranspiration" should not be used and we would like to refer to Savenije (2004) for arguments to not use this term. Briefly, transpiration is a bio-physical process, with different timescales and characteristics thereby being distinct to all other evaporative fluxes, which are purely physical processes. The term "evapotranspiration" is therefore a misleading definition, adding up different kinds of processes.

References

Savenije, H. H. G.: The importance of interception and why we should delete the term evapotranspiration from our vocabulary, Hydrological Processes, 18, 1507-1511, 10.1002/hyp.5563, 2004

Review #3

We would like Dr. Ducharne for her feedback on the manuscript. We will try to improve on the comments and raised issues.

1. We lack a lot of information regarding the models and their use. *The main idea is to propose evolutions of the root zone moisture capacity (RZMC) at a yearly time step by a kind of inverse modelling using the observed river discharge of the perturbed and unperturbed catchments as input.*

1.a) The simple "water balance model" allows a direct inversion of the RZMC, given parameters describing the canopy interception processes and the vegetation recovery time, and restricting the water balance to only 5 months between May and October, to get rid from the influence of snow (the experimental catchments are located in Oregon and New Hampshire):

- *Unless vegetation growth is really restricted to these 5 months, this tends to underestimate the RZMC, and could explain why the Hubbard Brook estimates are so small for forested sites (23 mm on Figure 1)*

We agree with it that vegetation growth is not restricted to these 5 months, but we argue that droughts are restricted to these 5 months. Changing the approach to the full year will indeed result in higher values, but only because water will be stored in the root zone (the simple method does not account for snow), whereas it is actually snow storage. Nevertheless, the actual dry periods are generally in July – August for these catchments. Thus, the deficit of E-P, which actually controls the storage capacity in the root zone, will be the largest in these periods. We would like to clarify here, that for the estimation of the mean E_t the full two year period is considered, only the calculation of daily deficits of E_t – P was taken over the 5 month summer period.

- *The total evaporation seems to comprise only transpiration and interception loss, and neglect soil evaporation: is it justified?*

It is correct that we do not treat soil evaporation as individual process. Rather, we lump the physical process of evaporation using one interception storage. This will without doubt introduce some uncertainty, but separating the processes is not really warranted by the available data and will result in increased parameter equifinality and thus considerable additional uncertainty. In addition, we argue that our transpiration estimates represent upper limits of transpiration, assuming a negligible amount of soil evaporation. In reality, the transpiration will indeed be lower due to soil evaporation. We will add a paragraph about this in the discussion.

- *Transpiration depends on a potential evaporation, which is not explained in the paper: does potential evaporation depend on the development of the canopy, as could be quantified by the Leaf Area Index (LAI)? This dependence is well known fact, and can be described for instance by the crop coefficient when following the FAO guidelines of Allen et al. (1986), or as a function of LAI like in the SVAT (Soil-Vegetation-Atmosphere Transfers) models. If such dependence exists in the experimental catchments, it should lead transpiration to decrease after deforestation, and recover with vegetation regrowth, with opposite effects on runoff, in agreement with Figure 2(a-c). In this case, if the model overlooks the positive link between vegetation development and the magnitude of transpiration, it should lead to underestimate the decrease of transpiration after deforestation, and to overestimate the decrease of the RZCM to match the increased observed runoff.*

The potential evaporation was determined based on a temperature based method (Hargreaves equation), and thus did not depend on vegetation. We will add this information in the Methodology. Also, the water balance based model used transpiration estimates, which were exclusively based on the observed water balance. Here, potential evaporation is thus <u>not needed</u> to determine the mean transpiration and was only used to <u>scale</u> the long-term mean value of transpiration to a daily time series.

- *A Monte-Carlo approach is used to assess the effect of the 3 parameters involved in the model (see Table 2) and this allows deriving a very useful uncertainty range around the estimated RZCM. Yet, no justification is given regarding the selected range for these parameters, which is a strong constrain to the uncertainty.*

We would like to refer to Figures S9-S26 in the Supplement. Here, all posterior distributions of the parameters are shown. It can be seen that none of the parameters has an extremely narrow posterior distribution close to one of the bounds of the prior distributions (i.e. upper and lower limits), which would point towards too narrow prior distributions. Only in a few instances, the distributions are close to values of zero, but negative values are not possible for these parameters (e.g. Figure S9b and S9f.) Thus, in general the applied parameter ranges were sufficient for the calibration.

1.b) The other four models are published conceptual hydrological models, and are calibrated over consecutive 2-yr windows to match the observed water discharge. These models seem to describe the full hydrological year, including the periods of snow, which is a significant difference with the previous approach. Even if some information is given in the Supplementary (but not at the same level for all the models), the reader should find in the main text if the snow is explicitly described, and how the evapotranspiration is calculated (in particular how it depends on the vegetation development, for the same reasons as explained above).

The conceptual models applied here all use similar functions as originally proposed by Feddes et al. (1978), with the resistance for transpiration as a part of the model (see equations in model descriptions in supplementary material S2). Thus, the models reflect the vegetation influence on transpiration, whereas the potential evaporation exclusively reflects the total energy available for evaporation, which is common practice in the vast majority of hydrological models. All models also used a snow module, as we described in the manuscript (p11,line 12 ; p11, line 27; p12, line 8). Nevertheless, we will try to state more clearly in the model descriptions how evaporation and snow are determined.

Some details should also be given regarding the calibration itself: How many parameters are calibrated in addition to RZCM (Su,max) for each model? Can all of them change in each 2-yr window, or does only Su,max change? How many tested parameter sets? How many parameter sets are kept at the end of the calibration (equifinality) and what are the corresponding performances to fit the observed discharge? There is a long paragraph from p12L27 to p13L14 which is rather hard to follow for non-specialists of optimization, and *could usefully be replaced by objective information regarding the qualities and weakness of the resulting calibration.*

We will add the number of free parameters for calibration in the model descriptions. Generally, almost all parameters were left as free calibration parameters. All parameters in HYMOD (8 parameters) and TUW (15) were free calibration parameters. The 9 parameters of FLEX were all free for calibration, only the slow reservoir coefficient K_s was sampled between narrower bounds, which were based on a recession analysis. The HYPE model used 15 parameters for calibration. We will also add information about the number of initial model runs (100,000 runs) and the number of final feasible parameter sets. The performances for three calibration objective functions (KGE, logKGE and VE) are summarized in Figures S5-S7, for each sub-period of calibration.

1.c) Another model is used, and presented in 3.5. It's an adaptation of FLEX, one of the above four models, in which an a priori rule for RZCM recovery with time after deforestation is added. First, it would probably be clearer if this model was presented just after the others. Second, much information, again, is lacking:

- How is the evolution Imax described since it also varies with time (p15L17-18)?

We will clarify how Imax changes in time in that model. We applied the same growth function (Equation 11), with growth parameters a and b set to respectively 0.001 [day^{-1}] and 1 [-].

- How are the parameters a and b of Eq. 11 selected? The resulting values are only given in the caption of Fig8, but don't they deserve some analysis? Do they relate logically to the recovery times that are discussed in section 4.3?

We will clarify this, but we would also like to refer to lines 12-16 of page 15. The parameters were determined based on a qualitative judgement (thus, just with the 'expert-eye') as it was just meant as a proof-of-concept. We fully acknowledge (p.15, l.20-27) that this is a mere exploratory analysis and a more thorough analysis, which may also include explicit and more detailed process understanding on root development, may be needed to have more adequate values for the growth parameters.

- How is decided when is RZCM minimum, and which is the minimum value, since Eq. 11 only describes the increasing part of the variations shown on Figure 8?

The minimum and constant values are determined in the same way as the shape of the curve, with qualitative judgement.

- Fig 8 shows performance criteria with and without the dynamic formulation of Su,max: to which period do they correspond? We must assume that the period is the full observed period for each catchment, but does it make sense for HB5, where half of the full period is before deforestation? Couldn't it be interesting to test the proposed function over the recovery period only?

The performance criteria in Fig. 8 correspond to the period just before the treatment until 15 years after the treatment. Therefore, it was not for the full observation period, also for Hubbard Brook WS5. To be more precise, HJ Andrews WS1 was evaluated from 01-10-1960 untill 30-09-1981, Hubbard Brook WS2 from 01-10-1962 untill 30-09-1983, Hubbard Brook WS5 was evaluated from 01-10-1982 untill 30-09-1999. In this way, we tried to 'zoom in' on the recovery period, just as you suggested, see also page 14, lines 22-25. We will make this clearer in the revision.

2. The conclusions are too frequently not supported by the Figures. *Examples:*

- p17,L3-4: "the three deforested catchments in the two research forests show generally similar response dynamics after the logging of the catchments (Fig.2)." No, for each of the rows/signatures, you can find one outlier over the three catchments.

This is why we stated it as ‘generally similar response dynamics’. We never claim the responses are exactly the same for all the catchments. We will rephrase this to ‘on balance similar response dynamics’.

- p18, L24-26 (regarding Figure 4): “Comparing the water balance and model-derived estimates of root zone storage capacity SR and Su,max, respectively, then showed that they exhibit very similar patterns in the study catchments.” This is abusive since TUW and HYMOD completely miss the difference between HJA and HB, and HB5 doesn’t show a clear response to deforestation against inter-annual variability for most models. When discussing Figure 4, the focus is put on the differences in RZCM due to deforestation and recovery. Yet, these differences are much smaller than the ones between the sites, and have a similar magnitude as the inter-annual variability for the two Hubbard Brook catchments. This should be taken in consideration in the discussion.

We would like to point out that we discuss the pattern, thus the dynamics, not the absolute values. Especially TUW and HYMOD show a bias (mostly due to the absence of an interception storage) compared with the water-balance method, but still show similar dynamics (decreasing during deforestation and a gradual increase afterwards). We discussed the possible reasons for the difference between the HJ Andrews and Hubbard Brook catchments (p19, line 5-11 and p20 line 16-18), but we will elaborate more on this in the revision. Briefly, HJ Andrews has a strong seasonal regime, whereas in Hubbard Brook the precipitation is more equally spread throughout the years. Therefore, HJ Andrews has a high need of large root zone storage capacities to allow access to sufficient water throughout the relatively long dry summer period, whereas the Hubbard Brook catchments can survive with much smaller storage volumes, due to significant summer rainfall and thus shorter dry periods that need to be bridged. We agree that inter-annual variability is high, but this is also the reason why we carried out the trend analysis with the undisturbed reference watersheds. In this way, the influence of inter-annual climatic variabilities should be filtered out.

- p20, L23-26: “It can be argued, that a combination of a relatively long period of low rainfall amounts and high potential evaporation, as can be noted by the relatively high mean

annual potential evaporation on top of Figure 4b, led to a high demand in1985". But the top three plots on Fig 4 are so small we can't see much!

We will make the plots bigger for clarity.

- *p21, L3-4: "Generally, the models applied in Hubbard Brook WS2 show similar behavior as in the HJ Andrews catchment." It's far from being obvious for HB5.*

This is absolutely correct and therefore, we do not state this.

- *p22, L16-17: "The results shown in Figure 4 indicate that these catchments had a rather stable root zone storage capacity during deforestation" (for HJA and HB2). Deforestation is indicated by a red band, and we clearly show a decreasing, not stable, RZCM during deforestation in HJA; for HB2, we don't see anything because the y-axis range is too large.*

We will rephrase this; we basically meant from more or less halfway the period of deforestation (for HJ Andrews just after 1964, and Hubbard Brook WS2 1967). We will try to make the plots clearer as well.

- *p23, L24-28: "Evaluating a set of hydrological signatures suggests that the dynamic formulation of Su,max allows the model to have a higher probability to better reproduce most of the signatures tested here (54% of all signatures in the three catchments) as shown in Figure 9a. A similar pattern is obtained for the more quantitative SRP (Figure 9b), where in 52% of the cases improvements are observed." This is abusive because your get degradation of the performance for 46% of the signatures in Fig9a, and 48% in Fig 9b, which is far from being negligible. If you look at HB5 only, the degraded signatures dominate, which contradicts the conclusion at p24, L27-29.*

We only stated what we found and never deny that 46% and 48% of the signatures show a decrease in performance for the two metrics. Moreover, it is also for these decreasing performances that we added the discussion starting from p24, line 13 until p25, line3, where we explained the origins of these decreases. The statement on p24, line 27-29, also refers to the rather light colors of red and blue, which indicate probabilities around 0.5 and S_{RP} values around 0, thus not a strong preference for one of the two models. We will further clarify this in the revision.

- *p24, L6-7: "In addition, a dynamic formulation of Su,max permits a more plausible representation of the variability in land-atmosphere exchange following land use change". Where does this come from? Provided that no signature in Fig 9 and Table 3 addresses the variability of land-atmosphere exchanges (all the signatures describe elements of the streamflow time series).*

We will remove this sentence.

- *p24, L9-10: "Fulfilling its function as a storage reservoir for plant available water, modelled transpiration is significantly reduced post-deforestation, which in turn results in increased runoff coefficients": if I see well on the very small Fig 2c, the results show exactly the opposite for HB5.*

We agree with this, but please note that in the line referred to in this comment to (p24, line9-10) we exclusively discuss the results for HJ Andrews. The two Hubbard Brook catchments are discussed in the following paragraphs.

- *p24, L19-21: "This can also be clearly seen from the hydrographs (Figure 10), where the later part of the recession in the late summer months is much better captured by the time-dynamic model." Personally, I see exactly the opposite, as the time-varying RZCM model in Fig 10b overestimates the peaks, which is not the case of the constant RZCM model in Fig 10a.*

We are confused by this comment, as we clearly see the same considering the peaks in Figure 10b, which we also discuss at page 21, line21-26. We agree that the improvement in the lower parts of the recession (thus not the peaks), is hard to see in Figure 10b, but we still believe this statement is supported by the figure. Please note the additional white space between observed and modelled discharge in the recession of July – August in Figure 10a (time constant model) compared to Figure 10b (time-varying model). To clarify, we will add insets into figs.10a and b, zooming in to a selected low flow period.

- Finally, the conclusion relies on a selection of the results that support the assumption of the authors, without considering the results that contradict it, and without a hint of doubt. The limits of the approach (including the model dependency, the small sample of observations which are not perfectly consistent) are not all discussed, nor any alternative frameworks. The authors could for instance consider the possibility that the RZCM could remain unchanged but not fully exploited by the vegetation. This is typically what helps some types of vegetation to resist to drought conditions.

We tried to keep the discussion brief and stated here the general findings. We believe there are good reasons the results in Hubbard Brook WS5 were less clear, which we also discussed (e.g. p21, line 14 until p22, line 3). Nevertheless, we will add in the discussion and conclusion sections more on several shortcomings and limitations, additional to what we already state in the discussion. We find the remark that root zone storage capacity could remain unchanged very interesting, and we use exactly this argument in our discussion on p19, line 29 until p20, line 6. We will make this clearer in the revision.

3. Abstract:

The abstract is not very clear regarding the methods (the proposed method is not solely based on climate data as written at L8-9, but it requires information on the deforestation, based on inverting the discharge observation in the present case). Like the conclusion, it builds too much on overstatement, but there is also an annoying circular reasoning, since the main conclusion comes from the beginning (L5-7: "Often this parameter [RZCM] is considered to *remain constant in time. This is not only conceptually problematic, it is also a potential source of error under the influence of land use and climate change.")*

We will clarify the abstract with the remarks made here. Again, we tried to generalize, which is unfortunately interpreted as an overstatement. Nevertheless, we will add more on the methods and try to clarify.

4. Other comments:

- Trend analysis (method in 3.4, results in 4.3): is it really about trends or about variability? Can we really speak of "trends" on sub-periods as short as those highlighted in blue and green in Fig 7o and 7r? Couldn't these two periods be lumped together? Some references should be given where to find more details on the extraction and interpretation of the 95%-confidence ellipses. Finally, Fig 7 is much too small.

We agree, at first the method is applied to detect a trend. In the second step, it is used to detect homogeneous sub-periods without a clear trend. We applied the differentiation between sub-periods as objectively as possible, based on the break points in Figures 7d-f. For the construction of the 95%-confidence ellipse, we refer to Equations 9 and 10, and the FAO-guidelines (Allen et al., 1998).

- Some sentences I did not find clear, although the paper is generally well written:

- p3, L13-15: "By extracting plant available water between field capacity and wilting point, roots create moisture storage volumes within their range of influence."

- p 4, L7-8: "other species with different water demands may be more in favor in the competition for resources"

- p4, L15: "These studies found that deforestation often leads to higher seasonal flows". Do you mean higher peak flows?

- p4, 30-31: "More systematic approaches, thus incorporation the change in the model formulation itself"

- *p14, L28-29: "the calibration was run with a series temporally evolving root zone storage capacities"*

- *p26, L27: I suggest using attributed to rather than caused by, unless a clear causality can be demonstrated.*

We will rephrase the sentences mentioned here.

References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, FAO, Rome, 300, D05109, 1998.

Feddes, R. A., Kowalik, P. J., and Zaradny, H.: Simulation of field water use and crop yield, Centre for Agricultural Publishing and Documentation., 1978

List of changes

-Change in title to: "The evolution of root zone moisture capacities after deforestation: a step towards hydrological predictions under change?"

-Abstract: The methodology is more extensively described, as suggested by reviewer #3.

-Introduction: Several textual changes based on comments of the three reviewers

-Study sites: Information on the potential evaporation is added, just as descriptions of the data and the references to it.

-Methodology:

- several textual changes as suggested by the reviewers
- Change in calibration from KGE, logKGE and VE to KGE and logKGE, as suggested by reviewer #1.
- Model descriptions updated with the number of free parameters and descriptions of snow and evaporation calculations.

-Results and Discussion is split into two different sections.

-Additional paragraph in the discussion about "General limitations".

-Conclusions: The conclusions are made less general, and are more about the results per catchment. The reasons for the less obvious results in Hubbard Brook WS5 are also added.

-Table 3: The signatures are renamed, with one symbol and a subscript.

-Figures 4-10 are replaced as the calibration changed slightly. We also tried to make Figures 4 and 7 clearer and added insets in Figure 10.

-Supplement: Additional table with the number feasible parameter sets. A figure was added with the method applied with a 5-year period for the water balance (instead of 2 years), for HJ Andrews WS1 in comparison with the FLEX model. This figure was originally made for Review #1.

-Replacement of Figures 1 and 8, based on Editor Comments of 12-11-2016

The evolution of root zone moisture capacities after ~~land use change~~deforestation: a step towards hydrological predictions under change?

Remko Nijzink[1], Christopher Hutton[2], Ilias Pechlivanidis[4], René Capell[4], Berit Arheimer[4], Jim Freer[2,3], Dawei Han[2], Thorsten Wagener[2,3], Kevin McGuire[5], Hubert Savenije[1], Markus Hrachowitz[1]

[1] Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, The Netherlands

[2] Department of Civil Engineering, University of Bristol, Bristol, UK

[3] Cabot Institute, University of Bristol, BS8 1UJ, Bristol, UK

[4] Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden

[5] Virginia Water Resources Research Center and Department of Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, VA, USA

Correspondence to: R. C. Nijzink (r.c.nijzink@tudelft.nl)

Abstract

The core component of many hydrological systems, the moisture storage capacity available to vegetation, is impossible to observe directly at the catchment scale and is typically treated as a calibration parameter or obtained from a priori available soil characteristics combined with estimates of rooting depth. Often this parameter is considered to remain constant in time. Using long-term data (30-40 years) from three experimental catchments that underwent significant land cover change, we tested the hypotheses that: (1) the root zone storage capacity significantly changes after deforestation, (2) changes in the root zone storage capacity can to a large extent explain post-treatment changes to the hydrological regimes and that (3) a time-dynamic formulation of the root zone storage can improve the performance of a hydrological model.

~~This is not only conceptually problematic, it is also a potential source of error under the influence of land use and climate change. In this paper we test the potential of a~~A recently introduced method to ~~robustly~~ estimate catchment-scale root zone storage capacities ~~exclusively~~ based on climate data (i.e. observed rainfall ~~distribution~~ and an estimate of ~~evaporation~~ transpiration) was used to reproduce the temporal evolution of root zone storage capacity under change. Briefly, the maximum deficit that arises from the difference between cumulative daily precipitation and transpiration can be considered as a proxy for root zone storage capacity. This value was compared to the value obtained from four different conceptual hydrological models that were calibrated for consecutive 2-year windows.~~-~~ ~~Using long-term data from three experimental catchments that underwent significant land use change, we tested the hypotheses that: (1) root zone moisture storage capacities are essentially controlled by land cover and climate, (2) root zone moisture storage capacities are dynamically adapting to changing environmental conditions, and (3) simple conceptual yet dynamic parametrization, mimicking changes in root zone storage capacities, can improve a model's skill to reproduce observed hydrological response dynamics.~~

It was found that water-balance derived root zone storage capacities were similar to the values obtained from calibration of ~~four different conceptual~~the hydrological models. A sharp decline in root zone storage capacity was observed after deforestation, followed by a gradual recovery, for two of the three catchments. Trend analysis suggested hydrological recovery periods between 5 and 13 years after deforestation. In a proof-of-concept analysis, one of the

hydrological models was adapted to allow dynamically changing root zone storage capacities, following the observed changes due to deforestation. Although the overall performance of the modified model did not considerably change, it provided significantly better representations of high flows and peak flows, underlining the potential of the approach. Iin 514% of all the evaluated hydrological signatures, considering all three catchments, improvements were observed when adding a time-variant representation of the root zone storage to the model.

In summary, it is shown that root zone moisture storage capacities can be highly affected by deforestation and climatic influences and that a simple method exclusively based on climate-data can not only provide robust, catchment-scale estimates of this critical parameter, but also reflect its time-dynamic behavior after deforestation.crucial and dynamic parameter.

1 Introduction

Vegetation a~~i~~s a core component of the water cycle, ~~it~~ shapes the partitioning of water fluxes on the catchment scale into ~~drainage~~ runoff components and evaporation, thereby controlling fundamental processes in ecosystem functioning (Rodriguez-Iturbe, 2000; Laio et al., 2001; Kleidon, 2004), such as flood generation (Donohue et al., 2012), drought dynamics (Seneviratne et al., 2010; Teuling et al., 2013), groundwater recharge (Allison et al., 1990; Jobbágy and Jackson, 2004) and land-atmosphere feedback (Milly and Dunne, 1994; Seneviratne et al., 2013; Cassiani et al., 2015). Besides increasing interception storage available for evaporation (Gerrits et al., 2010), vegetation critically interacts with the hydrological system in a co-evolutionary way by root water uptake for transpiration, towards a dynamic equilibrium with the available soil moisture to avoid water shortage (Donohue et al., 2007; Eagleson, 1978, 1982; Gentine et al., 2012; Liancourt et al., 2012) and related adverse effects on carbon exchange and assimilation rates (Porporato et al., 2004; Seneviratne et al., 2010). ~~By extracting plant available water between field capacity and wilting point,~~ R~~r~~oots create moisture storage volumes within their range of influence, from which they extract water that is stored between field capacity and wilting point. This ~~water holding or~~ root zone storage capacity~~,~~ S_R, sometimes also referred to as plant available water holding capacity, in the unsaturated soil is therefore the key component of many hydrological systems (Milly and Dunne, 1994; Rodriguez-Iturbe et al., 2007).

There is increasing theoretical and experimental evidence that vegetation dynamically adapts its root system, and thus S_R , to environmental conditions, ~~balancing between~~to secure, on the one hand, ~~securing~~ access to sufficient moisture to meet the canopy water demand and, on the other hand, to minimize~~ing~~ the carbon investment for sub-surface growth and maintenance of the root system (Brunner et al., 2015; Schymanski et al., 2008; Tron et al., 2015). In other words, the hydrologically active root zone is optimized to guarantee productivity and transpiration of vegetation, given the climatic circumstances (Kleidon, 2004). Several studies ~~already~~ previously showed the strong influence of climate on this hydrologically active root zone (e.g. Reynolds et al., 2000; Laio et al., 2001; Schenk and Jackson, 2002). Moreover, droughts are often identified as critical situations that can affect ecosystem functioning evolution (e.g. Allen et al., 2010; McDowell et al., 2008; Vose et al.).

In addition to the general adaption to environmental conditions, vegetation has some potential to adapt roots to such periods of water shortage (Sperry et al., 2002; Mencuccini, 2003; Bréda et al., 2006). In the short term, stomatal closure and reduction of leaf area will lead to reduced transpiration. In several case studies for specific plants, it was also shown that plants may even shrink their roots and reduce soil-root conductivity during droughts, while recovering after re-wetting (Nobel and Cui, 1992; North and Nobel, 1992). In the longer term, and more importantly, trees can improve their internal hydraulic system, for example by recovering damaged xylem or by allocating more biomass for roots (Sperry et al., 2002; Rood et al., 2003; Bréda et al., 2006). Similarly, Tron et al. (2015) argued that roots follow groundwater fluctuations, which may lead to increased rooting depths when water tables drop. ~~In addition, as circumstances change, other~~ Such changing environmental conditions may also provide other plant species with different water demand~~s~~, than the ones present under given conditions, with an ~~may be more in favor~~ advantage in the competition for resources, as for example shown by Li et al. (2007).

The hydrological functioning of catchments (Black, 1997; Wagener et al., 2007) and thus the partitioning of water ~~fluxes~~ into ~~evaporation/transpiration~~evaporative fluxes and ~~drainage~~runoff components is not only affected by the continuous adaption of vegetation to changing climatic conditions. Rather, it is well understood that anthropogenic changes to land cover, such as deforestation, can considerably alter hydrological regimes. This has been shown historically through many paired watershed studies (e.g. Bosch and Hewlett, 1982; Andréassian, 2004; Brown et al., 2005; Alila et al., 2009). These studies found that deforestation often leads to generally higher seasonal flows and/or an increased frequency of high flows in streams, while decreasing evaporative fluxes. The time scales of hydrological recovery after such land ~~use~~ cover disturbances were shown to be highly sensitive to climatic conditions and the growth dynamics of the regenerating species (e.g. Jones and Post, 2004; Brown et al., 2005) .

Although land-use change effects on hydrological functioning are widely acknowledged, it is less well understood, which parts of the hydrological system are affected in which way and over which time scales. As a consequence, most catchment-scale models were originally not developed to deal with such changes in the system, but rather for 'stationary' ~~situations~~ conditions (Ehret et al., 2014). This is ~~valid~~ true for both top-down hydrological models, such as~~e.g.~~ HBV (Bergström, 1992) or GR4J (Perrin et al., 2003), and bottom-up models, such as~~e.g.~~ MIKE-SHE (Refsgaard and Storm, 1995) or HydroGeoSphere (Brunner and Simmons, 2012). Several modelling studies have in the past incorporated temporal effects of land use change to some degree (Andersson and Arheimer, 2001; Bathurst et al., 2004; Brath et al., 2006), but they mostly rely on ad hoc assumptions about how hydrological parameters are affected (Legesse et al., 2003; Mahe et al., 2005; Onstad and Jamieson, 1970; Fenicia et al., 2009). ~~More systematic approaches, thus~~Approaches which incorporate~~ion~~ the change in the model formulation itself, are rare and have only recently gained momentum (e.g. Du et al., 2016; Fatichi et al., 2016; Zhang et al., 2016). This is of critical importance as on-going land ~~use~~ cover and climate change dictates the need for a better understanding of their effects on hydrological functioning (Troch et al., 2015) and their explicit consideration in hydrological models for more reliable predictions under change (Hrachowitz et al., 2013; Montanari et al., 2013).

As a step towards such an improved understanding and the development of time-dynamic models, we argue that the root zone storage capacity S_R~~-~~, ~~sometimes also referred to as plant available water holding capacity,~~ is a core component determining the hydrological response, and needs to be treated as dynamically evolving parameter in hydrological modelling as a function of climate and vegetation. Gao et al. (2014) recently demonstrated that catchment-scale S_R can be robustly estimated exclusively based on long-term water balance considerations. Wang-Erlandsson et al. (2016) derived global estimates of S_R using remote-sensing based precipitation and evaporation products, which demonstrated considerable spatial variability of S_R in response to climatic drivers. In traditional approaches, S_R is typically determined either by the calibration of a hydrological model (e.g. Seibert and McDonnell, 2010; Seibert et al., 2010) or based on soil characteristics and sparse, averaged estimates of root depths, often obtained from literature (e.g. Breuer et al., 2003; Ivanov et al., 2008). This does neither reflect the dynamic nature of the root system nor does it consider to a sufficient extent the actual function of the root zone: providing plants with continuous and efficient access to water. ~~The main reason for this is that due to the lack of detailed estimates of root depths and their evolution over time, some average values obtained from literature are typically used.~~ This leads to the situation that soil porosity often effectively controls the values of S_R~~.~~ used in a model. Consider, as a thought experiment, two plants of the same species growing on different soils. They will, with the same average root depth, then have access to different volumes of water, which will merely reflect the differences in soil porosity. This is in strong contradiction to the expectation that these plants would design root systems that provide access to similar water volumes, given the evidence for efficient carbon investment in root growth (Milly, 1994; Schymanski et al., 2008; Troch et al., 2009) and posing that plants of the same species have common limits of operation. This argument is supported by a recent study, in which was shown that water balance derived estimates of S_R are at least as plausible as soil derived estimates (de Boer-Euser et al., 2016) in many environments and that the maximum root depth controls evaporative fluxes and drainage (Camporese et al., 2015).

Therefore, using water balance based estimates of S_R in several deforested as well as in untreated reference sites in two experimental forests, we test the hypotheses that (1) the root zone storage capacity S_R significantly changes after deforestation, (2) ~~changes~~ the evolution in S_R can ~~to a large extent~~ explain post-treatment changes to the hydrological regimes and that (3) a time-dynamic formulation of S_R can improve the performance of a hydrological model.

2 Study sites

The catchments under consideration are part of the H.J. Andrews Experimental Forest and the Hubbard Brook Experimental Forest. A summary of the main catchment characteristics can be found in Table 1. Daily discharge (Campbell, 2014a; Johnson and Rothacher, 2016), precipitation (Campbell, 2014b; Daly and McKee, 2016) and temperature time series (Campbell, 2014c, 2014d; Daly and McKee, 2016) were obtained from the databases of the Hubbard Brook Experimental Forest and the HJ Andrews Experimental Forest. Potential evaporation was estimated by the Hargreaves equation (Hargreaves and Samani, 1985).

~~1.1~~2.1 H.J. Andrews Experimental Forest

The H.J. Andrews Experimental Forest is located in Oregon, USA (44.2°N, 122.2°W) and was established in 1948. The catchments at H.J. Andrews are described in many studies (e.g. Rothacher, 1965; Dyrness, 1969; Harr et al., 1975; Jones and Grant, 1996; Waichler et al., 2005) ~~and an overview of the site is presented in Table 1~~.

Before vegetation removal and at lower elevations the forest generally consisted of 100- to 500-year old coniferous species, such as Douglas-fir (*Pseudotsuga menziesii*), western hemlock (*Tsuga heterophylla*) and western redcedar (*Thuja plicata*), whereas upper elevations were characterized by noble fir (*Abies procera*), Pacific silver fir (*Abies amabilis*), Douglas-fir, and western hemlock. Most of the precipitation falls from November to April (about 80% of the annual precipitation), whereas the summers are generally drier, leading to signals of precipitation and potential evaporation that are out of phase. ~~The catchment characteristics of the watersheds in H.J. Andrews (WS) are provided in Table 1~~.

Deforestation of H.J. Andrews WS1 started in August 1962 (Rothacher, 1970). Most of the timber was removed with skyline yarding. After finishing the logging in October 1966, the remaining debris was burned and the site was left for natural regrowth. WS2 is the reference catchment, which was not harvested.

~~1.2~~2.2 Hubbard Brook Experimental Forest

The Hubbard Brook Experimental Forest is a research site established in 1955 and located in New Hampshire, USA (43.9°N, 71.8°W). The Hubbard Brook experimental catchments are described in a many publications (e.g. Hornbeck et al., 1970; Hornbeck, 1973; Dahlgren and Driscoll, 1994; Hornbeck et al., 1997; Likens, 2013). ~~An overview of the site and catchments used in this study are given in Table1.~~

Prior to vegetation removal, the forest was dominated by northern hardwood forest composed of sugar maple (*Acer saccharum*), American beech (*Fagus grandifolia*) and yellow birch (*Betula alleghaniensis*) with conifer species such as red spruce (*Picea rubens*) and balsam fir (*Abies balsamea*) occurring at higher elevations and on steeper slopes with shallow soils. The forest was selectively harvested from 1870 to 1920, damaged by a hurricane in 1938, and is currently not accumulating biomass (Campbell et al., 2013; Likens, 2013). The annual precipitation and runoff is less than in H.J. Andrews (Table 1). Precipitation is rather uniformly spread throughout the year without distinct dry and wet periods, but with snowmelt dominated peak flows occurring around April and distinct low-flows during the summer months due to increased evaporation rates (Federer et al., 1990). Vegetation removal occurred in the catchment of WS2 between 1965-1968 and in WS5 between 1983-1984. Hubbard Brook WS3 is the undisturbed reference catchment.

Hubbard Brook WS2 was completely deforested in November and December 1965 (Likens et al., 1970). To minimize disturbance, no roads were constructed and all timber was left in the catchment. On June 23, 1966, herbicides were sprayed from a helicopter to prevent regrowth. Additional herbicides were sprayed in the summers of 1967 and 1968 from the ground.

In Hubbard Brook WS5, all trees were removed between October 18, 1983 and May 21, 1984, except for a 2 ha buffer near an adjacent reference catchment (Hornbeck et al., 1997). WS5 was harvested as a whole-tree mechanical clearcut with removal of 93% of the above-ground biomass (Hornbeck et al., 1997; Martin et al., 2000); thus, including smaller branches and debris. Approximately 12% of the catchment area was developed as the skid trail network. Afterwards, no treatment was applied and the site was left for regrowth.

23 Methodology

To assure reproducibility and repeatability, the executional steps in the experiment were defined in a detailed protocol, following Ceola et al. (2015), which is provided as supplementary material in Section S1.

2.13.1 Water balance-derived root zone moisture capacities S_R

The root zone moisture storage capacities S_R and their change over time were determined according to the methods suggested by Gao et al. (2014), and subsequently succesfully tested by de Boer-Euser et al. (2016) and Wang-Erlandsson et al. (2016). Briefly, the long-term water balance provides information on actual mean transpiration. In a first step, the interception capacity has to be assumed, in order to determine the effective precipitation P_e [L T^{-1}], following the water balance equation for interception storage:

$$\frac{dS_i}{dt} = P - E_i - P_e \quad , \qquad (1)$$

With S_i [L] interception storage, P the precipitation [L T^{-1}], E_i the interception evaporation [L T^{-1}]. This is solved with the constitutive relations:

$$E_i = \begin{cases} E_p & if\ E_p dt < S_i \\ \frac{S_i}{dt} & if\ E_p dt \geq S_i \end{cases} \qquad (2)$$

$$P_e = \begin{cases} 0 & if\ S_i \leq I_{max} \\ \frac{S_i - I_{max}}{dt} & if\ S_i > I_{max} \end{cases} \qquad (3)$$

With, additionally, E_p the potential evaporation [L T^{-1}] and I_{max} [L] the interception capacity. NeverthelessAs, I_{max} will also be affected by land use cover change,. tThis was addressed by introducing the three parameters $I_{max,eq}$ (long-term equilibrium interception capacity) [L], $I_{max,change}$ (post-treatment interception capacity) [L] and T_r (recovery time) [T], leading to a time-dynamic formulation of I_{max}:

$$I_{max} = \begin{cases} I_{max,eq} & for\ \ t < t_{change}\ , t > t_{change,end} + T_r \\ I_{max,eq} - \dfrac{I_{max,eq} - I_{max,change}}{t_{change,end} - t_{change,start}}\left(t - t_{change,start}\right) & for\ t_{change,start} < t < t_{change,end} \\ I_{max,change} + \dfrac{I_{max,eq} - I_{max,change}}{T_r}\left(t - t_{change,end}\right) & for\ t_{change,end} < t < t_{change,end} + T_r \end{cases} \quad (4)$$

with $t_{change,start}$ the time that deforestation started and $t_{start,end}$ the time deforestation finished.

Following a Monte-Carlo sampling approach, upper and lower bounds of E_i were then estimated based on 1000 random samples of these parameters, eventually leading to upper and lower bounds for P_e. The interception capacity was assumed to increase after deforestation for Hubbard Brook WS2, as the debris was left at the site. For Hubbard Brook WS5 and HJ Andrews WS1 the interception capacity was assumed to decrease after deforestation, as here the debris was respectively burned and removed. Furthermore, in the absence of more detailed information, it was assumed that the interception capacities changed linearly during deforestation towards $I_{max,change}$ and linearly recovered to I_{max} over the period T_r as well. See Table 2 for the applied parameter ranges.

Hereafter, the long term mean transpiration can be estimated with the remaining components of the long term water balance, assuming no additional gains/losses, storage changes and/or data errors:

$$\overline{E}_t = \overline{P}_e - \overline{Q}, \quad (5)$$

where $\overline{E}_t$ ~~E_t~~ [L T^{-1}] is the long-term mean actual transpiration, $\overline{P}_e$ ~~P_e~~ [L T^{-1}] is the long-term mean effective precipitation and $\overline{Q}$ ~~Q~~ [L T^{-1}] is the long-term mean catchment runoff. Taking into account seasonality, the actual mean transpiration is scaled with the ratio of long-term mean daily potential evaporation E_p over the mean annual potential evaporation E_p:

$$E_t(t) = \frac{E_p(t)}{\overline{E_p}} * \overline{E_t} \qquad (6)$$

Based on this, the cumulative deficit between actual transpiration and precipitation over time can be estimated by means of an 'infinite-reservoir'. In other words, the cumulative sum of daily water deficits, i.e. evaporation minus precipitation, is calculated between T_0, which is the time the deficit equals zero, and T_1, which is the time the total deficit returned to zero. The maximum deficit of this period then represents the volume of water that needs to be stored to provide vegetation continuous access to water throughout that time:

$$S_R = \max \int_{T0}^{T1} (E_t - P_e)\, dt, \qquad (7)$$

where S_R [L] is the maximum root zone storage capacity over the time period between T_0 and T_1. See also Figure 1 for a graphical example of the calculation for the Hubbard Brook catchment for one specific realization of the parameter sampling. The $S_{R,20yr}$ for drought return periods of 20 years was estimated using the Gumbel extreme value distribution (Gumbel, 1941) as previous work suggested that vegetation designs S_R to satisfy deficits caused by dry periods with return periods of approximately 10-20 years (Gao et al., 2014; de Boer-Euser et al., 2016). Thus, the ~~yearly~~ maximum values of S_R~~,~~ for each year, as obtained by equation 7~~6~~, were fitted to the extreme value distribution of Gumbel, and subsequently, the $S_{R,20yr}$ was determined.

For the study catchments that experienced logging and subsequent reforestation, it was assumed that the root system converges towards a dynamic equilibrium approximately 10 years after reforestation. Thus, the equilibrium $S_{R,20yr}$ was estimated using only data over a period that started at least 10 years after the treatment. For the growing root systems during the years after reforesting, the storage capacity does not yet reach its dynamic equilibrium $S_{R,20yr}$. Instead of determining an equilibrium value, the maximum occurring deficit for each year was in that case considered as the maximum demand and thus as the maximum required storage $S_{R,1yr}$ for that year. To make these yearly estimates, the mean transpiration was determined in a similar ~~fashion~~ way as stated by Equation 5. However, the assumption of no storage change may not be valid for 1-year periods. In a trade-off~~,~~ to limit the potential bias introduced by inter-annual storage changes in the catchments, the mean transpiration was determined based on the 2-year water balance, thus assuming ~~no~~ negligible storage change over these years.

The deficits in the months October-April are highly affected by snowfall, as estimates of the effective precipitation are estimated without accounting for snow, leading to soil moisture changes that spread out over an unknown longer period due to the melt process. Therefore, to avoid this influence of snow, only deficits as defined by Equation 7, in the period of May – September are taken into consideration, which is also the period where deficits are ~~caused~~ significantly increasing due to ~~by~~ relatively low rainfall ~~precipitation~~ and high transpiration rates, thus causing soil moisture depletion and drought stress for the vegetation, ~~and~~which in turn, shap~~ing~~es the root zone.

~~2.2~~3.2 Model-derived root zone storage capacity $S_{u,max}$

The water balance derived equilibrium $S_{R,20yr}$ as well as the dynamically changing $S_{R,1yr}$ that reflects regrowth patterns in the years after treatment were compared with estimates of the calibrated parameter $S_{u,max}$, which represents the mean catchment root zone storage capacity in lumped conceptual hydrological models. Due to the lack of direct observations of the changes in the root zone storage capacity, this comparison was used to investigate whether the estimates of the root zone storage capacity $S_{R,1yr}$, ~~and~~ their sensitivity to land ~~use~~ cover change ~~as well as~~and their effect on hydrological functioning, can provide ~~similar~~ plausible results ~~as the model-based root zone storage~~. Model-based estimates of root zone storage capacity may be highly influenced by model formulations and parameterizations. Therefore, four different hydrological models were used to derive the parameter ~~of~~ $S_{u,max}$ in order to obtain a set of different estimates of the catchment scale root zone storage capacity. The major features of the model routines for root-zone moisture tested here are briefly summarized below and detailed descriptions including the relevant equations are provided as supplementary material (Section S2).

~~2.2.1~~3.2.1 FLEX

The~~A~~ FLEX-based model (Fenicia et al., 2008) was applied in a lumped way to the catchments. The model has 9 parameters, 8 of which are free calibration parameters, sampled from relatively wide, uniform prior distributions. In contrast, based on the estimation of a Master Recession Curve (e.g. Fenicia et al., 2006), an informed prior distribution between narrow bounds could be used for determining the slow reservoir coefficient K_S.

-The model~~It~~ consists of five storage components. First, a snow routine has to be run, which is a simple degree-day module, similar as used in, for example, HBV (Bergström, 1976). After the snow routine, ~~before~~ the precipitation enters the interception reservoir. Here, water evaporates at potential rates or, when exceeding a threshold, ~~continues~~ directly reaches~~to~~ the soil moisture reservoir. The soil moisture routine is modelled in a similar way as the Xinanjiang model (Zhao, 1992). Briefly, it contains a distribution function that determines the fraction of the catchment where the storage deficit in the root zone is satisfied and that is therefore hydrologically connected to the stream and generating storm runoff. From the soil moisture reservoir, water can further vertically percolate down to recharge the groundwater or leave the reservoir through transpiration. Transpiration is a function of maximum root zone storage $S_{u,max}$ and the actual root zone storage, similar to the functions described by Feddes et al. (1978).

Water that cannot be stored in the soil moisture storage then is split into preferential percolation to the groundwater and runoff generating fluxes that enter a fast reservoir, which represents fast responding system components such as shallow subsurface and overland flow.

~~(Fenicia et al., 2006; Fenicia et al., 2008)~~

~~2.2.2~~3.2.2 HYPE

The HYPE model (Lindström et al., 2010) estimates soil moisture for Hydrological Response Units (HRU), which is the finest calculation unit in this catchment model. In the current set-up, 15 parameters were left free for calibration. Each HRU consists of a unique combination of soil and land-use classes with assigned soil depths. Water input is estimated from precipitation after interception and a snow module at the catchment scale, after which the water enters the three defined soil layers in each HRU. Evaporation and transpiration ~~takes place~~occurs in ~~from~~ the first two layers and fast surface runoff is produced when these layers are fully saturated or when rainfall rates exceeds the maximum infiltration capacities. Water can move between the layers through percolation or laterally via fast flow pathways. ~~The catchment can also receive input of lateral flow from upper sub-catchments.~~ The groundwater table is fluctuating between the soil layers with the lowest soil layer normally reflecting the base flow component in the hydrograph. The water balance of each HRU is calculated independently and the runoff is then aggregated in a local stream with routing before entering the main stream.

~~2.2.3~~3.2.3 TUW

The TUW model (Parajka et al., 2007) is a conceptual model with a structure similar to that of HBV (Bergström, 1976) and has 15 free calibration parameters. After a snow module, based on a degree-day approach, water enters a soil moisture routine. From this soil moisture routine, water is partitioned into runoff generating fluxes and ~~transpiration~~evaporation. Here, transpiration is determined as a function of maximum root zone storage $S_{u,max}$ and actual root zone storage as well. The runoff generating fluxes percolate into two series of reservoirs. A fast responding reservoir with overflow outlet represents shallow subsurface and overland flow, while the slower responding reservoir represents the groundwater.

~~2.2.4~~3.2.4 HYMOD

HYMOD (Boyle, 2001) is similar to the applied model structure for FLEX, but only has 8 parameters. ~~b~~ Besides that, the interception module and percolation from soil moisture to the groundwater are missing. Nevertheless, the model accounts similarly for the partitioning of transpiration and runoff generation in a soil moisture routine. Also for this model, transpiration is a function of maximum storage and actual storage in the root zone. The runoff generating fluxes are ~~then~~ eventually divided over a slow reservoir, representing groundwater, and a fast reservoir, representing the fast processes.

~~2.3~~3.3 Model calibration

Each model was calibrated using a Monte-Carlo strategy within consecutive two year windows in order to obtain a time series of root zone moisture capacities $S_{u,max}$. FLEX, TUW and HYMOD were all run 100,000 times, whereas HYPE was run 10,000 times and 20,000 times for HJ Andrews WS1 and the Hubbard Brook catchments respectively, due to the required runtimes. The Kling-Gupta efficiency for flows (Gupta et al., 2009)~~,~~and the Kling-Gupta efficiency for the logarithm of the flows ~~and the Volume Error (Criss and Winston, 2008)~~ were simultaneously used as objective functions in a multi-objective calibration approach to evaluate the model performance for each window. These were selected in order to obtain rather balanced solutions that enable a sufficient representation of peak flows, low flows and the water balance. The unweighted Euclidian Distance D_E of the three objective functions served as an informal measure to obtain these balanced solutions (e.g. Hrachowitz et al., 2014; Schoups et al., 2005):

~~$$L(\theta) = 1 - \sqrt{(1 - E_{KG})^2 + (1 - E_{logKG})^2 + (1 - E_{VE})^2}$$~~

$$L(\theta) = 1 - \sqrt{(1 - E_{KG})^2 + (1 - E_{logKG})^2} \qquad (8)$$

where L(θ) is the conditional probability for parameter set θ [-], E_{KG} the Kling-Gupta efficiency [-], E_{logKG} the Kling-Gupta efficiency for the log of the flows [-], and E_{VE} the volume error [-].

Eventually, a weighing method based on the GLUE-approach of Freer et al. (1996) was applied. To estimate posterior parameter distributions all solutions with Euclidian Distances smaller than 1 were maintained as feasible. The posterior distributions were then determined with the Bayes rule (cf. Freer et al., 1996):

$$L_2(\theta) = L(\theta)^n * L_0(\theta)/C \qquad (9)$$

where $L_0(\theta)$ is the ~~uninformed~~ prior parameter distribution [-], $L_2(\theta)$ is the posterior conditional probability [-] , n is a weighing factor (set to 5) [-], and C a normalizing constant [-]. 5/95th model uncertainty intervals were then constructed based on the posterior conditional probabilities.

~~2.4~~3.4 Trend analysis

To test if $S_{R,1yr}$ significantly changes following de- and subsequent reforestation, which would also indicate shifts in distinct hydrological regimes, a trend analysis, as suggested by Allen et al. (1998), was applied to the $S_{R,1yr}$ values obtained from the water balance-based method. As the sampling of interception capacities (Eq. 4) leads to $S_{R,1yr}$ values for each point in time, which are all equally likely in absence of any further knowledge, the mean of this range was assumed as an approximation of the time-dynamic character of $S_{R,1yr}$.

Briefly, a linear regression between the full series of the cumulative sums of $S_{R,1yr}$ in the deforested catchment and the unaffected control catchment is established and the residuals and the cumulative residuals are plotted in time. A 95%-confidence ellipse is then constructed from the residuals:

$$X = \frac{n}{2}\cos(\alpha) \qquad (10)$$

$$Y = \frac{n}{\sqrt{n-1}}\, Z_{p95}\, \sigma_r\, \sin(\alpha) \qquad (11)$$

where X presents the x-coordinates of the ellipse [T], Y represents the y-coordinates of the ellipse [L], n is the length of the time series [T], α is the angle defining the ellipse (0 - 2π) between the diagonal of the ellipse and the x-axis [-], Z_{p95} is the value belonging to a probability of 95% of the standard student t-distribution [-] and σ_r is the standard deviation of the residuals (assuming a normal distribution) [L].

When the cumulative sums of the residuals plot outside the 95%-confidence interval defined by the ellipse, the null-hypothesis that the time series are homogeneous is rejected. In that case, the residuals from this linear regression where residual values change from either solely increasing to decreasing or vice versa, can then be used to identify different sub-periods in time.

Thus, in a second step, for each identified sub-period a new regression, with new (cumulative) residuals, can be used to check homogeneity for these sub-periods. In a similar way as before, when the cumulative residuals of these sub-periods now plot within the accompanying newly created 95%-confidence ellipse, the two series are homogeneous for these sub-periods. In other words, the two time series show a consistent behavior over this particular period.

3.5 Model with time-dynamic formulation of $S_{u,max}$

In a last step, the FLEX model was reformulated to allow for a time-dynamic representation of the parameter $S_{u,max}$, reflecting the root zone storage capacity.

As a reference, the long-term water balance derived root zone storage capacity $S_{R,20yr}$ was used as a static formulation of $S_{u,max}$ in the model, and thus kept constant in time. The remaining parameters were calibrated using the calibration strategy outlined above over a period starting with the treatment in the individual catchments until at least 15 years after the end of the treatment. This was done to focus on the period under change (i.e. vegetation removal and recovery), during which the differences between static and dynamic formulations of $S_{u,max}$ are assumed to be most pronounced.

To test the effect of a dynamic formulation of $S_{u,max}$ as a function of forest regrowth, the calibration was run with a ~~series~~ temporally evolving series of root zone storage capacit~~yies, similar to formulations of leaf area index and overstore height for the DHSVM model by Waichler et al. (2005)~~. The time-dynamic series of $S_{u,max}$ were obtained from a relatively simple growth function, the Weibull function (Weibull, 1951):

~~$S_{u,max}(t) = S_{R,20y} \left(1 - e^{-a*t^b}\right)$~~

$$S_{u,max}(t) = S_{R,20yr} \left(1 - e^{-a\,t^b}\right)$$ ~~$S_{u,max}(t) = S_{R,20yr} \left(1 - e^{-at^b}\right)$~~

~~$S_{u,max}(t) = S_{R,20yr} \left(1 - e^{-a*t^b}\right)$~~, (11)

where $S_{u,max}$ (t) is the root zone storage capacity *t* time steps after reforestation [L], $S_{R,20yr}$ is the equilibrium value [L], and *a* $[T^{-1}]$ and *b* [-] are shape parameters. In the absence of more information, this equation was selected as a first, simple way of incorporating the time-dynamic character of the root zone storage capacity in a conceptual hydrological model. In this way, root growth is exclusively determined dependent on time, whereas the shape-parameters *a* and *b* merely implicitly reflect the influence of other factors, such as climatic forcing in a lumped way. These parameters were estimated based on qualitative judgement so that $S_{u,max}(t)$ coincides well with the suite of S_{R1yr} values after logging. In other words, the values were chosen by trial and error in such a way, that the time-dynamic formulation of $S_{u,max}(t)$ shows a visually good correspondence with the S_{R1yr} values. This approach was followed to filter out the short term fluctuations in the S_{R1yr} values, which is not warranted by this equation. ~~In addition, it should be noted~~Note that this rather simple approach is merely meant as a proof-of-concept for a dynamic formulation of $S_{u,max}$.

In addition, the remaining parameter directly related to vegetation, the interception capacity (*Imax*), was also assigned a time-dynamic formulation. Here, the same growth function was applied (Eq. 11), but the shape of the growth function was assumed fixed (i.e. growth parameters *a* and *b* were fixed to values of 0.001 $[day^{-1}]$ and 1 [-]) loosely based on the posterior ranges of the window calibrations, with qualitative judgement as well. This growth function was used to ensure the degrees of freedom for both the time-variant and the time-invariant models, leaving the equilibrium value of the interception capacity as the only free calibration parameter for this process. Note that the empirically parameterized growth functions can be readily extended and/or replaced by more mechanistic, process-based descriptions of vegetation growth if warranted by the available data and was here merely used to test the effect of considering changes in vegetation on the skill of models to reproduce hydrological response dynamics.

To assess the performance of the dynamic model compared to the time-invariant formulation, beyond the calibration objective functions, model skill in reproducing 28 hydrological signatures was evaluated (Sivapalan et al., 2003). Even though the signatures are not always fully independent of each other, this larger set of measures allows a more complete evaluation of the model skill as, ideally, the model should be able to ~~perfectly and~~ simultaneously reproduce ~~each~~ all signatures. An overview of the signatures is given in Table 3~~2~~. The results of the comparison were quantified on the basis of the probability of improvement for each signature (Nijzink et al., 2016):

$$P_{I,S} = P\left(S_{dyn} > S_{stat}\right) = \sum_{i=1}^{n} P\left(S_{dyn} > S_{stat} \mid S_{dyn} = r_i\right) P\left(S_{dyn} = r_i\right) \quad (12)$$

where S_{dyn} and S_{stat} are the distributions of the signature performance metrics of the dynamic and static model, respectively, for the set of all feasible solutions retained from calibration, r_i is a single realization from the distribution of S_{dyn} and n is the total number of realizations of the S_{dyn} distribution. For $P_{I,S} > 0.5$ it is then more likely that the dynamic model outperforms the static model with respect to the signature under consideration, and vice versa for $P_{I,S} < 0.5$. The signature performance metrics that were used are the relative error for single-valued signatures and the Nash-Sutcliffe efficiency (Nash and Sutcliffe, 1970) for signatures that represent a time series.

In addition, as a more quantitative measure, the Ranked Probability Score, giving information on the magnitude of model improvement or deterioration, was calculated (Wilks, 2005):

$$S_{RP} = \frac{1}{M-1} \sum_{m=1}^{M} \left[\left(\sum_{k=1}^{m} p_k \right) - \left(\sum_{k=1}^{m} o_k \right) \right]^2 \quad (13)$$

where M is the number of feasible solutions, p_k the probability of a certain signature performance to occur and o_k the probability of the observation to occur (either 1 or 0, as there is only a single observation). Briefly, the S_{RP} represents the area enclosed between the cumulative probability distribution obtained by model results and the cumulative probability distribution of the observations. Thus, when modelled and observed cumulative probabilities are identical, the enclosed area goes to zero. Therefore, the difference between the S_{RP} for the feasible set of solutions for the time-variant and time-invariant model formulation was used in the comparison, identifying which model is quantitatively closer to the observation.

~~3~~4 Results ~~and Discussion~~

~~3.1~~4.1 Deforestation and changes in hydrological response dynamics

We found that the three deforested catchments in the two research forests show on balance ~~generally~~ similar response dynamics after the logging of the catchments (Fig.2). This supports the findings from previous studies of these catchments (Andréassian, 2004; Bosch and Hewlett, 1982; Hornbeck et al., 1997; Rothacher et al., 1967). More specifically, it was found that the observed annual runoff coefficients for HJ Andrews WS1 and Hubbard Brook WS2 (Fig. 2a,b) change after logging of the catchments, also in comparison with the adjacent, undisturbed reference watersheds. Right after deforestation, runoff coefficients increase, ~~but~~ which is ~~are~~ followed by a gradual decrease. ~~This change in runoff behavior points towards shifts in the yearly sums of transpiration, which can, except for climatic variation, be linked to the regrowth of vegetation that takes place at a similar pace as the changes in hydrological dynamics. This coincidence of regrowth dynamics and evolution of runoff coefficients was not only noticed by Hornbeck et al. (2014) for the Hubbard Brook, but was also previously acknowledged for example by Swift and Swank (1981) in the Coweeta experiment or Kuczera (1987) for eucalypt regrowth after forest fires. The key role of vegetation in this partitioning between runoff and transpiration (Donohue et al., 2012), or more specifically root zones (Gentine et al., 2012), necessarily leads to a change in runoff coefficients when vegetation is removed. Similarly, Gao et al. (2014) found a strong correlation between root zone storage capacities and runoff coefficients in more than 300 US catchments, which lends further support to the hypothesis that root zone storage capacities may have decreased in deforested catchments right after removal of the vegetation.~~

The annual autocorrelation coefficients with a 1-day lag time are generally lower after logging than in the years before the change, which can be seen in particular from Figures 2e and 2f as here a long pre-treatment time series record is available. Nevertheless, the climatic influence cannot be ignored here, as the reference watershed shows a similar pattern. Only for Hubbard Brook WS5 (Fig. 2f), the autocorrelation shows reduced values in the first years after logging. Thus, the flows at any time *t+1* are less dependent on the flows at *t*, which points towards less memory and thus less storage in the system (i.e. reduced S_R), leading to increased peak flows, similar to the reports of, for example, Patric and Reinhart (1971) for one of the Fernow experiments.

The declining limb density for HJ Andrews WS1 (Fig. 2g) shows increased values right after deforestation, whereas longer after deforestation the values seem to plot closer to the values obtained from the reference watershed. This indicates that for the same number of peaks less time was needed for the recession in the hydrograph in the early years after logging. In contrast, the rising limb density shows increased values during and right after deforestation for Hubbard Brook WS2 and WS5 (Fig 2k-2l), compared to the reference watershed. Here, less time was needed for the rising part of the hydrograph in the more early years after logging. Thus, the recession seems to be affected in HJ Andrews WS1, whereas the Hubbard Brook watersheds exhibits a quicker rise of the hydrograph.

Eventually, the flow duration curves, as shown in Figures 2m-2o, indicate a higher variability of flows, as the years following deforestation plot with an increased steepness of the flow duration curve, i.e. a higher flashiness. This increased flashiness of the catchments after deforestation can also be noted from the hydrographs shown in Figure 3. The peaks in the hydrographs are generally higher, and the flows return faster to the baseflow values in the years right after deforestation than some years l later after some forest regrowth, all with similar values for the yearly sums of precipitation and potential evaporation.

~~3.2~~4.2 Temporal evolution of S_R and $S_{u,max}$

The observed changes in the hydrological response of the study catchments (as discussed above) were also clearly reflected in the temporal evolution of the root zone storage capacities as described by the catchment models (Fig. 4). The models all exhibited Kling-Gupta efficiencies ranging between 0.5 and 0.8 and Kling-Gupta efficiencies of the log of the flows between 0.2 and 0.8 (see the supplementary material Figures S5-7, with all posterior parameter distributions in Figures S10~~9~~-S2~~6~~7, and the number of feasible solutions in Tables S5-S7). Comparing the water balance and model-derived estimates of root zone storage capacity S_R and $S_{u,max}$, respectively, then showed that they exhibit very similar patterns in the study catchments. Especially for HJ Andrews WS1 and Hubbard Brook WS2, ~~In general,~~ ~~r~~root zone storage capacities sharply decreased after deforestation and~~, when regrowth occurred,~~ gradually recovered during regrowth towards a dynamic equilibrium of climate and vegetation, whereas the undisturbed reference catchments of HJ Andrews WS2 and Hubbard Brook WS3 showed a rather constant signal over the full period (see the supplementary material Figure S8).

~~This in agreement with Mahe et al. (2005), who found in a modelling exercise that water holding capacities needed to be lowered after a reduction in vegetation.~~

The HJ Andrews WS1 shows the clearest signal when looking at the water balance derived S_R, as can be seen by the green shaded area in Figure 4a. Before deforestation, the root zone storage capacity $S_{R,1yr}$ was found to be around 400mm. ~~In spite of the high annual precipitation volumes, such comparatively high $S_{R,1yr}$ is plausible given the marked seasonality of the precipitation in the Mediterranean climate (Koeppen-Geiger class Csb) and the approximately 6 months phase shift between precipitation and potential evaporation peaks in the study catchment, which dictates that the storage capacities need to be large enough to store precipitation falling mostly during winter throughout the extended dry periods with higher energy supply throughout the rest of the year (Gao et al., 2014).~~ During deforestation, the $S_{R,1yr}$ required to provide the remaining vegetation with sufficient and continuous access to water decreased from around 400 mm to 200 mm. For the first 4-6 years after deforestation the $S_{R,1yr}$ increased again, reflecting the increased water demand of vegetation with the regrowth of the forest. In addition, it was observed that in the period 1971- 1978 $S_{R,1yr}$ slowly decreased again in HJ Andrews.

The four models show a similar pronounced decrease of the calibrated, feasible set of $S_{u,max}$ during deforestation and a subsequent gradual increase over the first years after deforestation. The model concepts, thus our assumptions about nature, can therefore only account for the changes in hydrological response dynamics of a catchment, when calibrated in a window calibration approach with different parameterizations for each time frame. The absolute values of $S_{u,max}$ obtained from the most parsimonious HYMOD and FLEX models (both 8 free calibration parameters) show a somewhat higher similarity to $S_{R,1yr}$ and its temporal evolution than the values from the other two models. In spite of similar general patterns in $S_{u,max}$, the higher number of parameters in TUW (i.e. 15) result, due to compensation effects between individual parameters, in wider uncertainty bounds which are less sensitive to change. It was also observed that in particular TUW overestimates $S_{u,max}$ compared to $S_{R,1yr}$, which ~~is caused by~~can be attributed to the absence of an interception reservoir, leading to a root zone that has to satisfy not only transpiration but all evaporative fluxes.

~~It was observed that in the period 1971- 1978 $S_{R,1yr}$ slowly decreased again in HJ Andrews. This pattern indicates that the storage demand in these years was lower as more rainfall reduced the need for storage in the system, which can be seen from the rainfall chart on top of Figure 4a. This reduced demand for storage could potentially indicate a contracting root system during that period, as an effort of vegetation to optimize its subsurface energy and carbon allocation for root maintenance in a trade-off for increased above surface growth. However, this conclusion is at this point not warranted by the available data and it can also be argued that the system is in a state of over-capacity for that period, still maintaining the root systems for the dryer years to come. The hydrograph for the years 1978-1979 (Figure 5) rather support the latter. Even though the FLEX model calibrated for this period tended towards larger values of $S_{u,max}$ (Figure 4a), still the modelled peaks are relatively high compared to the observed peaks. This suggests that the model requires a higher buffer in the root zone to reduce the peak flows rather than that root zones should have contracted in this time of reduced need. Thus, from 1980 and onwards the system can rather easily survive the period of growing demand caused by the relatively dry and warm years.~~

Hubbard Brook WS2 exhibits a similarly clear decrease in root zone storage capacity as a response to deforestation, as shown in Figure 4b. The water balance-based $S_{R,1yr}$ estimates approach values of zero during and right after deforestation. In these years the catchment was treated with herbicides, removing effectively any vegetation, thereby minimizing transpiration. ~~Low $S_{R,1yr}$ values are highly plausible in this catchment because the relatively humid climate and the absence of pronounced rainfall seasonality strongly reduces storage requirements (Gao et al., 2014).~~ In this catchment a more gradual regrowth pattern occurred, which continued after logging started in 1966 until around 1983. ~~However, the marked increase in $S_{R,1yr}$ at that time rather points towards an exceptional year, in terms of~~

~~climatological factors, than a sudden expansion of the root zone. It can also be observed from Figure 3a that the runoff coefficient was relatively low for 1985, suggesting either increased evaporation or a storage change. It can be argued, that a combination of a relatively long period of low rainfall amounts and high potential evaporation, as can be noted by the relatively high mean annual potential evaporation on top of Figure 4b, led to a high demand in 1985. Parts of the vegetation may not have survived these high-demand conditions due to insufficient access to water, which in turn can explain the dip in $S_{R,1yr}$ for the following year, which is in agreement with reduced growth rates of trees after droughts as observed by for example Bréda et al. (2006).~~

~~The hydrographs of 1984-1985 (Figure 6a) and 1986-1987 (Figure 6b) also show that July-August 1985 was exceptionally dry, whereas the next year in August 1986 the catchment seems to have increased peak flows. This either points towards an actual low storage capacity due to contraction of the roots during the dry summer or a low need of the system to use the existing capacity, for instance to recover other vital aspects of the system.~~

Generally, the models applied in Hubbard Brook WS2 show similar behavior as in the HJ Andrews catchment. The calibrated $S_{u,max}$ clearly follows the temporal pattern of $S_{R,1yr}$, reflecting the pronounced effects of de- and reforestation. It can, however, also be observed that the absolute values of $S_{u,max}$ exceed the $S_{R,1yr}$ estimates. While FLEX on balance exhibits the closest resemblance between the two values, in particular the TUW model exhibits wide uncertainty bounds with elevated $S_{u,max}$ values. Besides the role of interception evaporation, which is only explicitly accounted for in FLEX, the results are also linked to the fact that the humid climatic conditions with little seasonality reduces the importance of the model parameter $S_{u,max}$, and makes it thereby more difficult to identify by calibration. The parameter is most important for lengthy dry periods when vegetation needs enough storage to ensure continuous access to water.

The temporal variation in S_R in Hubbard Brook WS5 does not show such a distinct signal as in the other two study catchments (Figure 4c). Moreover, it can be noted that in the summers of 1984 and 1985 the values of $S_{R,1yr}$ are relatively high. Nevertheless, the model based values of $S_{u,max}$ show again similar dynamics as the water balance based $S_{R,1yr}$ values. TUW and HYMOD show again higher model based values, but also FLEX is now overestimating the root zone storage capacity. ~~Here the forest was removed in a whole-tree harvest in winter '83-'84 followed by natural regrowth. The summers of 1984 and 1985 were very dry~~

summers, as also reflected by the high values of $S_{R,1yr}$. The young system had already developed enough roots before these dry periods to have access to a sufficiently large water volume to survive this summer. This is plausible, as the period of the highest deficit occurred in mid-July and lasted until approximately the end of September, thus long after the growing season, allowing enough time for an initial growth and development of young roots from April until mid-July. In addition, the composition of the new forest differed from the old forest with more pin cherry (*Prunus pensylvanica*) and paper birch (*Betula papyrifera*). This supports the statements of a quick regeneration as these species have a high growth rate and reach canopy closure in a few years. Furthermore, the forest was not treated with either herbicides (Hubbard Brook WS2) or burned (HJ Andrews WS1), leaving enough low shrubs and herbs to maintain some level of transpiration (Hughes and Fahey, 1991; Martin, 1988). It can thus be argued, similar to Li et al. (2007), that the remaining vegetation experienced less competition and could increase root water uptake efficiency and transpiration per unit leaf area. This is in agreement with Hughes and Fahey (1991), who also stated that several species benefited from the removal of canopies and newly available resources in this catchment. Lastly, several other authors related the absence of a clear change in hydrological dynamics to the severe soil disturbance in this catchment (Hornbeck et al., 1997; Johnson et al., 1991). These disturbances lead to extra compaction, whereas at the same time species were changing, effectively masking any changes in runoff dynamics.

3.34.3 Process understanding - trend analysis and change in hydrological regimes

The trend analysis for water-balance derived values of $S_{R,1yr}$ suggests that for all three study catchments significantly different hydrological regimes in time can be identified before and after deforestation, linked to changes in $S_{R,1yr}$ (Fig. 7). For all three catchments, the cumulative residuals plot outside the 95%-confidence ellipse, indicating that the time series obtained in the control catchments and the deforested catchments are not homogeneous (Figures 7g-7i).

Rather obvious break points can be identified in the residuals plots for the catchments HJ Andrews WS1 and Hubbard Brook WS2 (Fig. 7d-7e). Splitting up the $S_{R,1yr}$ time series according to these break points into the periods before deforestation, deforestation and recovery resulted in three individually homogenous time series that are significantly different from each other, indicating switches in the hydrological regimes. The results shown in Figure

4 indicate that these catchments developed a~~had a~~ rather stable root zone storage capacity ~~during~~ sometime after the start of deforestation (for HJ Andrews WS1 after 1964, for Hubbard Brook WS2 after 1967). Hence, recovery and deforestation balanced each other, leading to a temporary equilibrium. The recovery signal then becomes more dominant in the years after deforestation. The third homogenous period suggests that the root zone storage capacity reached a dynamic equilibrium without any further systematic changes. This can be interpreted in the way that in the HJ Andrews WS1 hydrological recovery after deforestation due to the recovery of the root zone store capacity took about 6-9 years (Fig. 7p), while Hubbard Brook WS2 required 10-13 years for hydrological recovery (Fig. 7q). This strongly supports the results of Hornbeck et al. (2014), who reported changes in water yield for WS2 for up to year 12 after deforestation.

The identification of different periods is less obvious for Hubbard Brook WS5, but the two time series of control catchment and treated catchment are significantly different (see the cumulative residuals in Figure 7i). Nevertheless, the most obvious break point in residuals can be found in 1989 (Figure 7f). In addition, it can be noted that turning points also exist in 1983 and 1985. These years can be used to split the time series into four groups (leading to the periods of 1964-1982, 1983-1985, 1986-1989 and 1990-2009 for further analysis). The cumulative residuals from the new regressions, based on the grouping, plot within the confidence bounds again, and show a period with deforestation (1983-1985) and recovery (1986-1989). Mou et al. (1993) reported similar findings with the highest biomass accumulation in 1986 and 1988, and slower vegetation growth in the early years. Therefore, full recovery took 5-6 years in Hubbard Brook WS5.

~~The above results do in general suggest similar recovery periods for forest systems as reported in earlier studies, such as Brown et al. (2005) or Hornbeck et al. (2014), who found that catchments reach a new equilibrium with a similar timescale as reported here with the direct link to the parameter describing the catchment-scale root zone storage capacity. The timescales are also in agreement with regression models to predict water yield after logging of Douglass (1983), who assumed a duration of water yield increases of 12 years for coniferous catchments. The timescales found here are around 10 years (here 5-13 years for the catchments under consideration), but will probably depend on climatic factors and vegetation type.~~

3.44.4 Time-variant model formulation

The adjusted model routine for FLEX, which uses a dynamic time series of $S_{u,max}$, generated with the Weibull growth function (Eq.11), resulted in a rather small impact on the overall model performance in terms of the calibration objective function values (Figure 8b, 8d, 8f) compared to the time-invariant formulation of the model. The strongest improvements for calibration were observed for the dynamic formulation of FLEX for HJ Andrews WS1 and Hubbard Brook WS2 (Figures 8b and 8d), which reflects the rather clear signal from deforestation in these catchments.

Evaluating a set of hydrological signatures suggests that the dynamic formulation of $S_{u,max}$ allows the model to have a higher probability to better reproduce most of the signatures tested here (5154% of all signatures in the three catchments) as shown in Figure 9a. A similar pattern is obtained for the more quantitative S_{RP} (Figure 9b), where in 52% of the cases improvements are observed. Most signatures for HJ Andrews WS1 show a high probability of improvement, with a maximum $P_{I,S}$ =0.69 (for $S_{Q95,winter}$) and an average $P_{I,S}$ = 0.55. Considering the large difference between the deforested situation and the new equilibrium situation of about 200 mm, this supports the hypothesis that here a time-variant formulation of $S_{u,max}$ does provide means for an improved process representation and, thus, hydrological signatures. Here, improvements are observed especially in the high flows in summer ($S_{Q5,summer}$, $S_{Q50,summer}$) and peak flows (e.g. S_{Peaks}, $S_{Peaks,summer}$, $S_{Peaks,winter}$), that illustrates that the root zone storage affects mostly the fast responding components of the system. as also suggested previously (e.g. de Boer-Euser et al., 2016; Euser et al., 2015; Oudin et al., 2004), by providing a buffer to storm response. In addition, a dynamic formulation of $S_{u,max}$ permits a more plausible representation of the variability in land-atmosphere exchange following land use change, which is a critical input to climate models (Entekhabi et al., 1996; Seneviratne et al., 2010). Fulfilling its function as a storage reservoir for plant available water, modelled transpiration is significantly reduced post-deforestation, which in turn results in increased runoff coefficients (cf. Gao et al., 2014), which have been frequently reported for post-deforestation periods by earlier studies (e.g. Hornbeck et al., 2014; Rothacher, 1970; Swift and Swank, 1981) .

At Hubbard Brook WS2 a more variable pattern is shown in the ability of the model to reproduce the hydrological signatures. It is interesting to note that the low flows (S_{Q95} ,$S_{Q95,summer}$, $S_{Q50,summer}$) improve, opposed to the expectation raised by the argumentation for

HJ Andrews WS1 that peak flows and high flows should improve. In this case, the peaks are too high for the time-dynamic model. Apparently, the model with a constant, and thus higher, $S_{u,max}$ stores water in the root zone, reducing recharge to the groundwater reservoir that maintains the lower flows and buffering more water, reducing the peaks. This can also be clearly seen from the hydrographs (Figure 10), where the later part of the recession in the late-summer months is much better captured by the time-dynamic model. Nevertheless, the peaks are too high for the time-dynamic model, which here is linked to an insufficient representation of snow-related processes, as can be seen from the hydrograph (April-May) as well, and possibly by an inadequate interception growth function, both leading to too high amounts of effective precipitation entering the root zone. An adjustment of these processes would have resulted in less infiltration and a smaller root zone storage capacity.

The probabilities of improvement for the signatures in Hubbard Brook WS5 show an even less clear signal, the model cannot clearly identify a preference for either a dynamic or static formulation of $S_{u,max}$ (relatively white colors in Fig. 9). This absence of a clear preference can be related to the observed patterns in water balance derived S_R (Figure 4c), which does not show a very clear signal after deforestation as well, indicating that the root zone storage capacity is of less importance in this humid region characterized by limited seasonality. Nevertheless, a similar argument as for the Hubbard Brook WS2 can be made here, as can be noted that the low flow statistics (e.g. Q_{95}, LFR) slightly improve, and some statistics concerning peak flows deteriorate (e.g. Peaks, AC), indicating similar issues regarding the modelling of snow and interception.

Wang-Erlandsson et al. (2016)

5 Discussion

5.1 Deforestation and changes in hydrological response dynamics

The changes found in the runoff behavior of the deforested catchments point towards shifts in the yearly sums of transpiration, which can, except for climatic variation, be linked to the regrowth of vegetation that takes place at a similar pace as the changes in hydrological dynamics. This coincidence of regrowth dynamics and evolution of runoff coefficients was not only noticed by Hornbeck et al. (2014) for the Hubbard Brook, but was also previously acknowledged for example by Swift and Swank (1981) in the Coweeta experiment or Kuczera (1987) for eucalypt regrowth after forest fires.

Therefore, the key role of vegetation in this partitioning between runoff and transpiration (Donohue et al., 2012), or more specifically root zones (Gentine et al., 2012), necessarily leads to a change in runoff coefficients when vegetation is removed. Similarly, Gao et al. (2014) found a strong correlation between root zone storage capacities and runoff coefficients in more than 300 US catchments, which lends further support to the hypothesis that root zone storage capacities may have decreased in deforested catchments right after removal of the vegetation.

5.2 Temporal evolution of S_R and $S_{u,max}$

The differences between the Hubbard Brook catchments and HJ Andrews catchments can be related to climatic conditions. In spite of the high annual precipitation volumes, high $S_{R,1yr}$ values are plausible for HJ Andrews WS1 given the marked seasonality of the precipitation in the Mediterranean climate (Koeppen-Geiger class Csb) and the approximately 6 months phase shift between precipitation and potential evaporation peaks in the study catchment, which dictates that the storage capacities need to be large enough to store precipitation falling mostly during winter throughout the extended dry periods with higher energy supply throughout the rest of the year (Gao et al., 2014). At the same time, low $S_{R,1yr}$ values in Hubbard Brook WS2 can be related to the relatively humid climate and the absence of pronounced rainfall seasonality strongly reduces storage requirements.

It can also be argued that there is a strong influence of the inter-annual climatic variability on the estimated root zone storage capacities. For example, the marked increase in $S_{R,1yr}$ in Hubbard Brook WS2 in 1985 rather points towards an exceptional year, in terms of climatological factors, than a sudden expansion of the root zone. It can also be observed from Figure 3a that the runoff coefficient was relatively low for 1985, suggesting either increased evaporation or a storage change. A combination of a relatively long period of low rainfall amounts and high potential evaporation, as can be noted by the relatively high mean annual potential evaporation on top of Figure 4b, may have led to a high demand in 1985. Parts of the vegetation may not have survived these high-demand conditions due to insufficient access to water, explaining the dip in $S_{R,1yr}$ for the following year, which is also in agreement with reduced growth rates of trees after droughts as observed by for example Bréda et al. (2006). The hydrographs of 1984-1985 (Figure 6a) and 1986-1987 (Figure 6b) also show that July-August 1985 was exceptionally dry, whereas the next year in August 1986 the catchment seems to have increased peak flows. This either points towards an actual low storage capacity due to contraction of the roots during the dry summer or a low need of the system to use the existing capacity, for instance to recover other vital aspects of the system.

Nevertheless, Hubbard Brook WS2 does not show a clear signal of reduced root zone storage, followed by a gradual regrowth. Here, the forest was removed in a whole-tree harvest in winter '83-'84 followed by natural regrowth. The summers of 1984 and 1985 were very dry summers, as also reflected by the high values of $S_{R,1yr}$. The young system had already developed enough roots before these dry periods to have access to a sufficiently large water volume to survive this summer. This is plausible, as the period of the highest deficit occurred in mid-July and lasted until approximately the end of September, thus long after the beginning of the growing season, allowing enough time for an initial growth and development of young roots from April until mid-July. In addition, the composition of the new forest differed from the old forest with more pin cherry (*Prunus pensylvanica*) and paper birch (*Betula papyrifera*). This supports the statements of a quick regeneration as these species have a high growth rate and reach canopy closure in a few years. Furthermore, the forest was not treated with either herbicides (Hubbard Brook WS2) or burned (HJ Andrews WS1), leaving enough low shrubs and herbs to maintain some level of transpiration (Hughes and Fahey, 1991; Martin, 1988). It can thus be argued, similar to Li et al. (2007), that the remaining vegetation experienced less competition and could increase root water uptake efficiency and transpiration per unit leaf area. This is in agreement with Hughes and Fahey (1991), who also stated that several species benefited from the removal of canopies and newly available resources in this catchment. Lastly, several other authors related the absence of a clear change in hydrological dynamics to the severe soil disturbance in this catchment (Hornbeck et al., 1997; Johnson et al., 1991). These disturbances lead to extra compaction, whereas at the same time species were changing, effectively masking any changes in runoff dynamics.

5.3 Process understanding - trend analysis and change in hydrological regimes

The found recovery periods correspond to recovery time scales for forest systems as reported elsewhere (e.g. Brown et al., 2005; Hornbeck et al., 2014; Elliott et al., 2016), who found that catchments reach a new equilibrium with a similar timescale as reported here with the direct link to the parameter describing the catchment-scale root zone storage capacity. The timescales are also in agreement with regression models to predict water yield after logging of Douglass (1983), who assumed a duration of water yield increases of 12 years for coniferous catchments.

The timescales found here are around 10 years (here 5-13 years for the catchments under consideration), but will probably depend on climatic factors and vegetation type. HJ Andrews WS1 has a recovery (6-9 years) slightly shorter compared to Hubbard Brook WS2 (10-13 years), which could depend on the different climatological conditions of the catchments. Nevertheless, it could also be argued that especially the spraying of herbicides had a strong impact on the recovery of vegetation in Hubbard Brook WS2, as the Hubbard Brook WS5 does not show such a distinct recovery signal.

5.4 Time-variant model formulation

It was found that a time dynamic formulation of $S_{u,max}$ merely improved the high and peak flow signatures for HJ Andrews WS1. Other authors also suggested previously (e.g. de Boer-Euser et al., 2016; Euser et al., 2015; Oudin et al., 2004) that that the root zone storage affects mostly the fast responding components of the system, by providing a buffer to storm response. Fulfilling its function as a storage reservoir for plant available water, modelled transpiration is significantly reduced post-deforestation, which in turn results in increased runoff coefficients (cf. Gao et al., 2014), which have been frequently reported for post-deforestation periods by earlier studies (e.g. Hornbeck et al., 2014; Rothacher, 1970; Swift and Swank, 1981)

Nevertheless, signatures considering the peak flows did not improve for the Hubbard Brook catchments. Apparently, the model with a constant, and thus higher, $S_{u,max}$ stored water in the root zone, reducing recharge to the groundwater reservoir that maintains the lower flows and buffering more water, reducing the peaks. This can also be clearly seen from the hydrographs (Figure 10), where the later part of the recession in the late-summer months is much better captured by the time-dynamic model. Nevertheless, the peaks are too high for the time-dynamic model, which here is linked to an insufficient representation of snow-related processes, as can be seen from the hydrograph (April-May) as well, and possibly by an inadequate interception growth function, both leading to too high amounts of effective precipitation entering the root zone. An adjustment of these processes would have resulted in less infiltration and a smaller root zone storage capacity.

It was acknowledged previously by several authors that certain model parameters may need time-dynamic formulations, like Waichler et al. (2005) with time-dynamic formulations of leaf area index and overstore height for the DHSVM model. In addition, Westra et al. (2014) captured long term dynamics in the storage parameter of the GR4J model with a trend correction, in fact leading to a similar model behavior as with the Weibull growth function in this study. Nevertheless, they only hypothesized about the actual hydrological reasons for this, which aimed at the changing number of farmer dams in the catchment. The results presented here indicate that vegetation, and especially root zone dynamics, has a strong impact on the long term non-stationarity of model parameters. The simple Weibull equation can be used as an extra equation in conceptual hydrological models to more closely reflect the dynamics of vegetation. The additional growth parameters may be left for calibration, but can also be estimated from simple water balance-based estimations of the root zone storage. In this way, the extra parameters should not add any uncertainty to the model outcomes.

5.5 General Limitations

The results presented here depend on the quality of the data and several assumptions made in the calculations. A limiting factor is that the potential evaporation is determined from temperature only, leading to values that may be relatively low and water balances that may not close completely. Generally, this would lead to a discrepancy between the modelled $S_{u,max}$, where potential evaporation is directly used, and the water balance-estimates of S_R. The models will probably generate higher root zone storages in order to compensate for the rather low potential evaporation. This can also be noted when looking at Figure 4 for several models.

In addition, the assumption that the water balance closes in the 2-year periods under consideration may in reality be often violated. It can be argued that the estimated transpiration for the calculation of S_R represents an upper boundary, when storage changes are ignored. This would lead to estimates of S_R that may be lower than presented here. Nevertheless, attempts with 5-year water balances to reduce the influence of storage changes (see the Supplementary Material Figure S9), showed that similar patterns were obtained. Values here were slightly lower due to more averaging in the estimation of the transpiration by the longer time period used for the water balance. Nevertheless, still a strong decrease after deforestation and gradual recovery can be observed.

The raised issues here can be fully avoided when, instead of a water balance-based estimation of the transpiration, remote sensing products are used to estimate the transpiration, similar to Wang-Erlandsson et al. (2016). However, water balance-based estimates may provide a rather quick solution.

The transpiration estimates were also only corrected for interception evaporation, thus assuming a negligible amount of soil evaporation. Making this additional separation is typically not warranted by the available data and would result in additional uncertainty. The transpiration estimates presented here merely represent an upper limit of transpiration and will be lower in reality due to soil evaporation. Thus, the values for $S_{R,1yr}$ may expected to be lower in reality as well.

46 Conclusion

In this study, three deforested catchments (HJ Andrews WS1, Hubbard Brook WS2 and WS5) were investigated to assess the dynamic character of root zone storage capacities using water balance, trend analysis, four different hydrological models and one modified model version. Root zone storage capacities were estimated based on a simple water balance approach. Results demonstrate a good correspondence between water-balance derived root zone storage capacities and values obtained by a 2-year moving window calibration of four distinct hydrological models.

~~There are significant changes in root zone storage capacity after deforestation, which were detected by both, a water-balance based method and the calibration of hydrological models. We found a good correspondence between water balance derived root zone storage capacities and values obtained by a 2-year moving window calibration of four distinct hydrological models.~~There are significant changes in root zone storage capacity after deforestation, which were detected by both, a water-balance based method and the calibration of hydrological models in two of the three catchments. –More specifically, root zone storage capacities showed for HJ Andrews WS1 and Hubbard Brook WS2 a sharp decrease in root zone storage capacities immediately after deforestation with a gradual recovery towards a new equilibrium.

This could to a large extent explain post-treatment changes to the hydrological regime. -These signals were however not clearly observed for Hubbard Brook WS5, probably due to soil disturbance, a new vegetation composition and a climatologically exceptional year. Nevertheless, ~~This could to a large extent explain post-treatment changes to the hydrological regime. T~~trend analysis showed significant differences for all three catchments with their corresponding, undisturbed reference watersheds. ~~suggested~~Based on this, recovery times were estimated to be between 5-13 years for the three catchments under consideration.

These findings underline the fact that root zone storage capacities in hydrological models, which are more often than not treated as constant in time, may need time-dynamic formulations with reductions after logging and gradual regrowth afterwards. Therefore, one of the models was subsequently formulated with a time-dynamic description of root zone storage capacity. Particularly under climatic conditions with pronounced seasonality and phase shifts between precipitation and evaporation, this resulted in improvements in model performance as evaluated by 28 hydrological signatures.

Even though this more complex system behavior may lead to extra unknown growth parameters, it has been shown here that a simple equation, reflecting the long-term growth of the system, can already suffice for a time-dynamic estimation of this crucial hydrological parameter. Therefore, this study clearly shows that observed changes in runoff characteristics after land ~~use~~ cover changes can be linked to relatively simple time-dynamic formulations of vegetation related model parameters.

Acknowledgements

We would like to acknowledge the European Commission FP7 funded research project “Sharing Water-related Information to Tackle Changes in the Hydrosphere– for Operational Needs” (SWITCH-ON, grant agreement number 603587), as this study was conducted within the context of SWITCH-ON as an example of scientific potential when using open data for collaborative research in hydrology.

Open Data were provided by the the Hubbard Brook Ecosystem Study (HBES), which is a collaborative effort at the Hubbard Brook Experimental Forest which is operated and maintained by the USDA Forest Service Northern Research Station, Newtown Square, PA, USA.

Open Data were provided by the HJ Andrews Experimental Forest research program, funded by the National Science Foundation's Long-Term Ecological Research Program (DEB 08-23380), US Forest Service Pacific Northwest Research Station, and Oregon State University.

References

Alila, Y., Kuraś, P. K., Schnorbus, M., and Hudson, R.: Forests and floods: A new paradigm sheds light on age-old controversies, Water Resources Research, 45, n/a-n/a, 10.1029/2008WR007207, 2009.

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. H., Gonzalez, P., Fensham, R., Zhang, Z., Castro, J., Demidova, N., Lim, J.-H., Allard, G., Running, S. W., Semerci, A., and Cobb, N.: A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests, Forest Ecology and Management, 259, 660-684, http://dx.doi.org/10.1016/j.foreco.2009.09.001, 2010.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, FAO, Rome, 300, D05109, 1998.

Allison, G. B., Cook, P. G., Barnett, S. R., Walker, G. R., Jolly, I. D., and Hughes, M. W.: Land clearance and river salinisation in the western Murray Basin, Australia, Journal of Hydrology, 119, 1-20, http://dx.doi.org/10.1016/0022-1694(90)90030-2, 1990.

Andersson, L., and Arheimer, B.: Consequences of changed wetness on riverine nitrogen – human impact on retention vs. natural climatic variability, Regional Environmental Change, 2, 93-105, 10.1007/s101130100024, 2001.

Andréassian, V.: Waters and forests: from historical controversy to scientific debate, Journal of Hydrology, 291, 1-27, http://dx.doi.org/10.1016/j.jhydrol.2003.12.015, 2004.

Bathurst, J. C., Ewen, J., Parkin, G., O'Connell, P. E., and Cooper, J. D.: Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses, Journal of Hydrology, 287, 74-94, http://dx.doi.org/10.1016/j.jhydrol.2003.09.021, 2004.

Bergström, S.: Development and application of a conceptual runoff model for Scandinavian catchments. SMHI Reports RHO, Norrköping, 1976.

Bergström, S.: The HBV model: Its structure and applications, Swedish Meteorological and Hydrological Institute, 1992.

Black, P. E.: WATERSHED FUNCTIONS1, JAWRA Journal of the American Water Resources Association, 33, 1-11, 10.1111/j.1752-1688.1997.tb04077.x, 1997.

Bosch, J. M., and Hewlett, J. D.: A review of catchment experiments to determine the effect of vegetation changes on water yield and evapotranspiration, Journal of Hydrology, 55, 3-23, http://dx.doi.org/10.1016/0022-1694(82)90117-2, 1982.

Boyle, D. P.: Multicriteria calibration of hydrologic models, 2001.

Brath, A., Montanari, A., and Moretti, G.: Assessing the effect on flood frequency of land use change via hydrological simulation (with uncertainty), Journal of Hydrology, 324, 141-153, http://dx.doi.org/10.1016/j.jhydrol.2005.10.001, 2006.

Bréda, N., Huc, R., Granier, A., and Dreyer, E.: Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences, Ann. For. Sci., 63, 625-644, 2006.

Breuer, L., Eckhardt, K., and Frede, H.-G.: Plant parameter values for models in temperate climates, Ecological Modelling, 169, 237-293, http://dx.doi.org/10.1016/S0304-3800(03)00274-6, 2003.

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., and Vertessy, R. A.: A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, Journal of Hydrology, 310, 28-61, http://dx.doi.org/10.1016/j.jhydrol.2004.12.010, 2005.

Brunner, I., Herzog, C., Dawes, M. A., Arend, M., and Sperisen, C.: How tree roots respond to drought, Frontiers in Plant Science, 6, 547, 10.3389/fpls.2015.00547, 2015.

Brunner, P., and Simmons, C. T.: HydroGeoSphere: A Fully Integrated, Physically Based Hydrological Model, Ground Water, 50, 170-176, 10.1111/j.1745-6584.2011.00882.x, 2012.

Campbell, J.: Hubbard Brook Experimental Forest (USDA Forest Service): Daily Streamflow by Watershed, 1956 - present, http://www.hubbardbrook.org/data/dataset.php?id=2, 2014a.

Campbell, J.: Hubbard Brook Experimental Forest (US Forest Service): Daily Precipitation Standard Rain Gage Measurements, 1956 - present, http://www.hubbardbrook.org/data/dataset.php?id=13, 2014b.

Campbell, J.: Hubbard Brook Experimental Forest (USDA Forest Service): Daily Maximum and Minimum Temperature Records, 1955 - present, http://www.hubbardbrook.org/data/dataset.php?id=59, 2014c.

Campbell, J.: Hubbard Brook Experimental Forest (USDA Forest Service): Daily Mean Temperature Data, 1955 - present, http://www.hubbardbrook.org/data/dataset.php?id=58, 2014d.

Campbell, J. L., Bailey, A. S., Eagar, C., Green, M. B., and Battles, J. J.: Vegetation treatments and hydrologic responses at the Hubbard Brook Experimental Forest, New Hampshire, in: Long-term silvicultural & ecological studies: Results for science and management,Volume 2, Yale University, Global Institute of Sustainable Forestry, Research Paper 013: 011-019, 2013.

Camporese, M., Daly, E., and Paniconi, C.: Catchment-scale Richards equation-based modeling of evapotranspiration via boundary condition switching and root water uptake schemes, Water Resources Research, 51, 5756-5771, 10.1002/2015WR017139, 2015.

Cassiani, G., Boaga, J., Vanella, D., Perri, M. T., and Consoli, S.: Monitoring and modelling of soil–plant interactions: the joint use of ERT, sap flow and eddy covariance data to characterize the volume of an orange tree root zone, Hydrol. Earth Syst. Sci., 19, 2213-2225, 10.5194/hess-19-2213-2015, 2015.

Ceola, S., Arheimer, B., Baratti, E., Blöschl, G., Capell, R., Castellarin, A., Freer, J., Han, D., Hrachowitz, M., Hundecha, Y., Hutton, C., Lindström, G., Montanari, A., Nijzink, R., Parajka, J., Toth, E., Viglione, A., and Wagener, T.: Virtual laboratories: new opportunities for collaborative water science, Hydrol. Earth Syst. Sci., 19, 2101-2117, 10.5194/hess-19-2101-2015, 2015.

Criss, R. E., and Winston, W. E.: Do Nash values have value? Discussion and alternate proposals, Hydrological Processes, 22, 2723-2725, 10.1002/hyp.7072, 2008.

Dahlgren, R. A., and Driscoll, C. T.: The effects of whole-tree clear-cutting on soil processes at the Hubbard Brook Experimental Forest, New Hampshire, USA, Plant and Soil, 158, 239-262, 10.1007/BF00009499, 1994.

Daly, C., and McKee, W.: Meteorological data from benchmark stations at the Andrews Experimental Forest, 1957 to present. Long-Term Ecological Research. Forest Science Data Bank, http://andrewsforest.oregonstate.edu/data/abstract.cfm?dbcode=MS001 2016.

de Boer-Euser, T., McMillan, H. K., Hrachowitz, M., Winsemius, H. C., and Savenije, H. H. G.: Influence of soil and climate on root zone storage capacity, Water Resources Research, n/a-n/a, 10.1002/2015WR018115, 2016.

Donohue, R. J., Roderick, M. L., and Mcvicar, T. R.: On the importance of including vegetation dynamics in Budyko's hydrological model, Hydrology and Earth System Sciences Discussions, 11, 983-995, 2007.

Donohue, R. J., Roderick, M. L., and McVicar, T. R.: Roots, storms and soil pores: Incorporating key ecohydrological processes into Budyko’s hydrological model, Journal of Hydrology, 436–437, 35-50, http://dx.doi.org/10.1016/j.jhydrol.2012.02.033, 2012.

Douglass, J. E.: THE POTENTIAL FOR WATER YIELD AUGMENTATION FROM FOREST MANAGEMENT IN THE EASTERN UNITED STATES1, JAWRA Journal of the American Water Resources Association, 19, 351-358, 10.1111/j.1752-1688.1983.tb04592.x, 1983.

Du, E., Link, T. E., Wei, L., and Marshall, J. D.: Evaluating hydrologic effects of spatial and temporal patterns of forest canopy change using numerical modelling, Hydrological Processes, 30, 217-231, 10.1002/hyp.10591, 2016.

Dyrness, C.: Hydrologic properties of soils on three small watersheds in the western Cascades of Oregon, Res. Note PNW-111, US Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station: Portland, OR, 17, 1969.

Eagleson, P. S.: Climate, soil, and vegetation: 3. A simplified model of soil moisture movement in the liquid phase, Water Resources Research, 14, 722-730, 10.1029/WR014i005p00722, 1978.

Eagleson, P. S.: Ecological optimality in water-limited natural soil-vegetation systems: 1. Theory and hypothesis, Water Resources Research, 18, 325-340, 10.1029/WR018i002p00325, 1982.

Ehret, U., Gupta, H. V., Sivapalan, M., Weijs, S. V., Schymanski, S. J., Blöschl, G., Gelfan, A. N., Harman, C., Kleidon, A., Bogaard, T. A., Wang, D., Wagener, T., Scherer, U., Zehe, E., Bierkens, M. F. P., Di Baldassarre, G., Parajka, J., van Beek, L. P. H., van Griensven, A., Westhoff, M. C., and Winsemius, H. C.: Advancing catchment hydrology to deal with predictions under change, Hydrol. Earth Syst. Sci., 18, 649-671, 10.5194/hess-18-649-2014, 2014.

Elliott, K. J., Caldwell, P. V., Brantley, S. T., Miniat, C. F., Vose, J. M., and Swank, W. T.: Water yield following forest-grass-forest transitions, Hydrol. Earth Syst. Sci. Discuss., 2016, 1-53, 10.5194/hess-2016-548, 2016.

Entekhabi, D., Rodriguez-Iturbe, I., and Castelli, F.: Soil Moisture Theories and ObservationsMutual interaction of soil moisture state and atmospheric processes, Journal of Hydrology, 184, 3-17, http://dx.doi.org/10.1016/0022-1694(95)02965-6, 1996.

Euser, T., Winsemius, H. C., Hrachowitz, M., Fenicia, F., Uhlenbrook, S., and Savenije, H. H. G.: A framework to assess the realism of model structures using hydrological signatures, Hydrol. Earth Syst. Sci., 17, 1893-1912, 10.5194/hess-17-1893-2013, 2013.

Euser, T., Hrachowitz, M., Winsemius, H. C., and Savenije, H. H. G.: The effect of forcing and landscape distribution on performance and consistency of model structures, Hydrological Processes, 29, 3727-3743, 10.1002/hyp.10445, 2015.

Fatichi, S., Pappas, C., and Ivanov, V. Y.: Modeling plant–water interactions: an ecohydrological overview from the cell to the global scale, Wiley Interdisciplinary Reviews: Water, 3, 327-368, 10.1002/wat2.1125, 2016.

Feddes, R. A., Kowalik, P. J., and Zaradny, H.: Simulation of field water use and crop yield, Centre for Agricultural Publishing and Documentation., 1978.

Federer, A. C., Flynn, L. D., Martin, W. C., Hornbeck, J. W., and Pierce, R. S.: Thirty years of hydrometeorologic data at the Hubbard Brook Experiment Forest, New Hampshire, 1990.

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Is the groundwater reservoir linear? Learning from data in hydrological modelling, Hydrol. Earth Syst. Sci., 10, 139-150, 10.5194/hess-10-139-2006, 2006.

Fenicia, F., Savenije, H. H. G., Matgen, P., and Pfister, L.: Understanding catchment behavior through stepwise model concept improvement, Water Resources Research, 44, n/a-n/a, 10.1029/2006WR005563, 2008.

Fenicia, F., Savenije, H. H. G., and Avdeeva, Y.: Anomaly in the rainfall-runoff behaviour of the Meuse catchment. Climate, land-use, or land-use management?, Hydrol. Earth Syst. Sci., 13, 1727-1737, 10.5194/hess-13-1727-2009, 2009.

Freer, J., Beven, K., and Ambroise, B.: Bayesian Estimation of Uncertainty in Runoff Prediction and the Value of Data: An Application of the GLUE Approach, Water Resources Research, 32, 2161-2173, 10.1029/95WR03723, 1996.

Gao, H., Hrachowitz, M., Schymanski, S. J., Fenicia, F., Sriwongsitanon, N., and Savenije, H. H. G.: Climate controls how ecosystems size the root zone storage capacity at catchment scale, Geophysical Research Letters, 41, 7916-7923, 10.1002/2014GL061668, 2014.

Gentine, P., D'Odorico, P., Lintner, B. R., Sivandran, G., and Salvucci, G.: Interdependence of climate, soil, and vegetation as constrained by the Budyko curve, Geophysical Research Letters, 39, n/a-n/a, 10.1029/2012GL053492, 2012.

Gerrits, A. M. J., Pfister, L., and Savenije, H. H. G.: Spatial and temporal variability of canopy and forest floor interception in a beech forest, Hydrological Processes, 24, 3011-3025, 10.1002/hyp.7712, 2010.

Gumbel, E. J.: The Return Period of Flood Flows, The Annals of Mathematical Statistics, 12, 163-190, 1941.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, Journal of Hydrology, 377, 80-91, http://dx.doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hargreaves, G. H., and Samani, Z. A.: Reference Crop Evapotranspiration from Temperature, Applied Engineering in Agriculture, 1, 96-99, 10.13031/2013.26773, 1985.

Harr, R. D., Harper, W. C., Krygier, J. T., and Hsieh, F. S.: Changes in storm hydrographs after road building and clear-cutting in the Oregon Coast Range, Water Resources Research, 11, 436-444, 10.1029/WR011i003p00436, 1975.

Hornbeck, J. W., Pierce, R. S., and Federer, C. A.: Streamflow Changes after Forest Clearing in New England, Water Resources Research, 6, 1124-1132, 10.1029/WR006i004p01124, 1970.

Hornbeck, J. W.: Storm flow from hardwood-forested and cleared watersheds in New Hampshire, Water Resources Research, 9, 346-354, 10.1029/WR009i002p00346, 1973.

Hornbeck, J. W., Martin, C. W., and Eagar, C.: Summary of water yield experiments at Hubbard Brook Experimental Forest, New Hampshire, Canadian Journal of Forest Research, 27, 2043-2052, 10.1139/x97-173, 1997.

Hornbeck, J. W., Eagar, C., Bailey, A., and Campbell, J. L.: Comparisons with results from the Hubbard Brook Experimental Forest in the Northern Appalachians, Long-Term Response of a Forest Watershed Ecosystem: Clearcutting in the Southern Appalachians, 213, 2014.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E., and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)—a review, Hydrological Sciences Journal, 58, 1198-1255, 10.1080/02626667.2013.803183, 2013.

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H. H. G., and Gascuel-Odoux, C.: Process consistency in models: The importance of system signatures, expert knowledge, and process complexity, Water Resources Research, 50, 7445-7469, 10.1002/2014WR015484, 2014.

Hughes, J. W., and Fahey, T. J.: Colonization Dynamics of Herbs and Shrubs in a Disturbed Northern Hardwood Forest, Journal of Ecology, 79, 605-616, 10.2307/2260656, 1991.

Ivanov, V. Y., Bras, R. L., and Vivoni, E. R.: Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks, Water Resources Research, 44, n/a-n/a, 10.1029/2006WR005588, 2008.

Jobbágy, E. G., and Jackson, R. B.: Groundwater use and salinization with grassland afforestation, Global Change Biology, 10, 1299-1312, 10.1111/j.1365-2486.2004.00806.x, 2004.

Johnson, C. E., Johnson, A. H., Huntington, T. G., and Siccama, T. G.: Whole-Tree Clear-Cutting Effects on Soil Horizons and Organic-Matter Pools, Soil Science Society of America Journal, 55, 497-502, 10.2136/sssaj1991.03615995005500020034x, 1991.

Johnson, S., and Rothacher, J.: Stream discharge in gaged watersheds at the Andrews Experimental Forest, 1949 to present. Long-Term Ecological Research.Forest Science Data Bank, Forest Science Data Bank, http://andrewsforest.oregonstate.edu/data/abstract.cfm?dbcode=HF004 2016.

Jones, J. A., and Grant, G. E.: Peak Flow Responses to Clear-Cutting and Roads in Small and Large Basins, Western Cascades, Oregon, Water Resources Research, 32, 959-974, 10.1029/95WR03493, 1996.

Jones, J. A., and Post, D. A.: Seasonal and successional streamflow response to forest cutting and regrowth in the northwest and eastern United States, Water Resources Research, 40, n/a-n/a, 10.1029/2003WR002952, 2004.

Jothityangkoon, C., Sivapalan, M., and Farmer, D. L.: Process controls of water balance variability in a large semi-arid catchment: downward approach to hydrological model development, Journal of Hydrology, 254, 174-198, http://dx.doi.org/10.1016/S0022-1694(01)00496-6, 2001.

Kleidon, A.: Global Datasets of Rooting Zone Depth Inferred from Inverse Methods, Journal of Climate, 17, 2714-2722, 10.1175/1520-0442(2004)017<2714:GDORZD>2.0.CO;2, 2004.

Kuczera, G.: Prediction of water yield reductions following a bushfire in ash-mixed species eucalypt forest, Journal of Hydrology, 94, 215-236, http://dx.doi.org/10.1016/0022-1694(87)90054-0, 1987.

Laio, F., Porporato, A., Ridolfi, L., and Rodriguez-Iturbe, I.: Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress: II. Probabilistic soil moisture dynamics, Advances in Water Resources, 24, 707-723, http://dx.doi.org/10.1016/S0309-1708(01)00005-7, 2001.

Legesse, D., Vallet-Coulomb, C., and Gasse, F.: Hydrological response of a catchment to climate and land use changes in Tropical Africa: case study South Central Ethiopia, Journal of Hydrology, 275, 67-85, http://dx.doi.org/10.1016/S0022-1694(03)00019-2, 2003.

Li, K. Y., Coe, M. T., Ramankutty, N., and Jong, R. D.: Modeling the hydrological impact of land-use change in West Africa, Journal of Hydrology, 337, 258-268, http://dx.doi.org/10.1016/j.jhydrol.2007.01.038, 2007.

Liancourt, P., Sharkhuu, A., Ariuntsetseg, L., Boldgiv, B., Helliker, B. R., Plante, A. F., Petraitis, P. S., and Casper, B. B.: Temporal and spatial variation in how vegetation alters the soil moisture response to climate manipulation, Plant and Soil, 351, 249-261, 10.1007/s11104-011-0956-y, 2012.

Likens, G. E., Bormann, F. H., Johnson, N. M., Fisher, D. W., and Pierce, R. S.: Effects of Forest Cutting and Herbicide Treatment on Nutrient Budgets in the Hubbard Brook Watershed-Ecosystem, Ecological Monographs, 40, 23-47, 10.2307/1942440, 1970.

Likens, G. E.: Biogeochemistry of a forested ecosystem, Springer, New York, 2013.

Lindström, G., Pers, C., Rosberg, J., Strömqvist, J., and Arheimer, B.: Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales, Hydrology Research, 41, 295-319, 10.2166/nh.2010.007, 2010.

Mahe, G., Paturel, J.-E., Servat, E., Conway, D., and Dezetter, A.: The impact of land use change on soil water holding capacity and river flow modelling in the Nakambe River, Burkina-Faso, Journal of Hydrology, 300, 33-43, http://dx.doi.org/10.1016/j.jhydrol.2004.04.028, 2005.

Martin, C. W.: Soil disturbance by logging in New England--review and management recommendations, Northern Journal of Applied Forestry, 5, 30-34, 1988.

Martin, C. W., Hornbeck, J. W., Likens, G. E., and Buso, D. C.: Impacts of intensive harvesting on hydrology and nutrient dynamics of northern hardwood forests, Canadian Journal of Fisheries and Aquatic Sciences, 57, 19-29, 10.1139/f00-106, 2000.

McDowell, N., Pockman, W. T., Allen, C. D., Breshears, D. D., Cobb, N., Kolb, T., Plaut, J., Sperry, J., West, A., Williams, D. G., and Yepez, E. A.: Mechanisms of plant survival and mortality during drought: why do some plants survive while others succumb to drought?, New Phytologist, 178, 719-739, 10.1111/j.1469-8137.2008.02436.x, 2008.

Mencuccini, M.: The ecological significance of long-distance water transport: short-term regulation, long-term acclimation and the hydraulic costs of stature across plant life forms, Plant, Cell & Environment, 26, 163-182, 10.1046/j.1365-3040.2003.00991.x, 2003.

Milly, P. C. D.: Climate, soil water storage, and the average annual water balance, Water Resources Research, 30, 2143-2156, 10.1029/94WR00586, 1994.

Milly, P. C. D., and Dunne, K. A.: Sensitivity of the Global Water Cycle to the Water-Holding Capacity of Land, Journal of Climate, 7, 506-526, doi:10.1175/1520-0442(1994)007<0506:SOTGWC>2.0.CO;2, 1994.

Montanari, A., and Toth, E.: Calibration of hydrological models in the spectral domain: An opportunity for scarcely gauged basins?, Water Resources Research, 43, n/a-n/a, 10.1029/2006WR005184, 2007.

Montanari, A., Young, G., Savenije, H. H. G., Hughes, D., Wagener, T., Ren, L. L., Koutsoyiannis, D., Cudennec, C., Toth, E., Grimaldi, S., Blöschl, G., Sivapalan, M., Beven, K., Gupta, H., Hipsey, M., Schaefli, B., Arheimer, B., Boegh, E., Schymanski, S. J., Di Baldassarre, G., Yu, B., Hubert, P., Huang, Y., Schumann, A., Post, D. A., Srinivasan, V., Harman, C., Thompson, S., Rogger, M., Viglione, A., McMillan, H., Characklis, G., Pang, Z., and Belyaev, V.: "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022, Hydrological Sciences Journal, 58, 1256-1275, 10.1080/02626667.2013.809088, 2013.

Mou, P., Fahey, T. J., and Hughes, J. W.: Effects of Soil Disturbance on Vegetation Recovery and Nutrient Accumulation Following Whole-Tree Harvest of a Northern Hardwood Ecosystem, Journal of Applied Ecology, 30, 661-675, 10.2307/2404245, 1993.

Nash, J. E., and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282-290, http://dx.doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nijzink, R. C., Samaniego, L., Mai, J., Kumar, R., Thober, S., Zink, M., Schäfer, D., Savenije, H. H. G., and Hrachowitz, M.: The importance of topography-controlled sub-grid process heterogeneity and semi-quantitative prior constraints in distributed hydrological models, Hydrol. Earth Syst. Sci., 20, 1151-1176, 10.5194/hess-20-1151-2016, 2016.

Nobel, P. S., and Cui, M.: Hydraulic Conductances of the Soil, the Root-Soil Air Gap, and the Root: Changes for Desert Succulents in Drying Soil, Journal of Experimental Botany, 43, 319-326, 10.1093/jxb/43.3.319, 1992.

North, G. B., and Nobel, P. S.: Drought-induced changes in hydraulic conductivity and structure in roots of Ferocactus acanthodes and Opuntia ficus-indica, New Phytologist, 120, 9-19, 10.1111/j.1469-8137.1992.tb01053.x, 1992.

Onstad, C. A., and Jamieson, D. G.: Modeling the Effect of Land Use Modifications on Runoff, Water Resources Research, 6, 1287-1295, 10.1029/WR006i005p01287, 1970.

Oudin, L., Andréassian, V., Perrin, C., and Anctil, F.: Locating the sources of low-pass behavior within rainfall-runoff models, Water Resources Research, 40, n/a-n/a, 10.1029/2004WR003291, 2004.

Parajka, J., Merz, R., and Blöschl, G.: Uncertainty and multiple objective calibration in regional water balance modelling: case study in 320 Austrian catchments, Hydrological Processes, 21, 435-446, 10.1002/hyp.6253, 2007.

Patric, J. H., and Reinhart, K. G.: Hydrologic Effects of Deforesting Two Mountain Watersheds in West Virginia, Water Resources Research, 7, 1182-1188, 10.1029/WR007i005p01182, 1971.

Perrin, C., Michel, C., and Andréassian, V.: Improvement of a parsimonious model for streamflow simulation, Journal of Hydrology, 279, 275-289, http://dx.doi.org/10.1016/S0022-1694(03)00225-7, 2003.

Porporato, A., Daly, E., Rodriguez, x, Iturbe, I., and Associate Editor: William, F. F.: Soil Water Balance and Ecosystem Response to Climate Change, The American Naturalist, 164, 625-632, 10.1086/424970, 2004.

Refsgaard, J. C., and Storm, B.: MIKE SHE, in: Computer Models of Watershed Hydrology, edited by: Singh, V. J., Water Resour. Publ., Littleton, Colorado, 1995.

Reynolds, J. F., Kemp, P. R., and Tenhunen, J. D.: Effects of long-term rainfall variability on evapotranspiration and soil water distribution in the Chihuahuan Desert: A modeling analysis, Plant Ecology, 150, 145-159, 10.1023/a:1026530522612, 2000.

Rodriguez-Iturbe, I.: Ecohydrology: A hydrologic perspective of climate-soil-vegetation dynamies, Water Resources Research, 36, 3-9, 10.1029/1999WR900210, 2000.

Rodriguez-Iturbe, I., D'Odorico, P., Laio, F., Ridolfi, L., and Tamea, S.: Challenges in humid land ecohydrology: Interactions of water table and unsaturated zone with climate, soil, and vegetation, Water Resources Research, 43, n/a-n/a, 10.1029/2007WR006073, 2007.

Rood, S. B., Braatne, J. H., and Hughes, F. M. R.: Ecophysiology of riparian cottonwoods: stream flow dependency, water relations and restoration, Tree Physiology, 23, 1113-1124, 10.1093/treephys/23.16.1113, 2003.

Rothacher, J.: Streamflow from small watersheds on the western slope of the Cascade Range of Oregon, Water Resources Research, 1, 125-134, 10.1029/WR001i001p00125, 1965.

Rothacher, J., Dyrness, C. T., Fredriksen, R. L., Forest, P. N., and Station, R. E.: Hydrologic and related characteristics of three small watersheds in the Oregon Cascades, Pacific Northwest Forest and Range Experiment Station, U.S. Dept. of Agriculture, 1967.

Rothacher, J.: Increases in Water Yield Following Clear-Cut Logging in the Pacific Northwest, Water Resources Research, 6, 653-658, 10.1029/WR006i002p00653, 1970.

Schenk, H. J., and Jackson, R. B.: Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems, Journal of Ecology, 90, 480-494, 10.1046/j.1365-2745.2002.00682.x, 2002.

Schoups, G., Hopmans, J. W., Young, C. A., Vrugt, J. A., and Wallender, W. W.: Multi-criteria optimization of a regional spatially-distributed subsurface water flow model, Journal of Hydrology, 311, 20-48, http://dx.doi.org/10.1016/j.jhydrol.2005.01.001, 2005.

Schymanski, S. J., Sivapalan, M., Roderick, M. L., Beringer, J., and Hutley, L. B.: An optimality-based model of the coupled soil moisture and root dynamics, Hydrology and Earth System Sciences Discussions, 12, 913-932, 2008.

Seibert, J., and McDonnell, J. J.: Land-cover impacts on streamflow: a change-detection modelling approach that incorporates parameter uncertainty, Hydrological Sciences Journal, 55, 316-332, 10.1080/02626661003683264, 2010.

Seibert, J., McDonnell, J. J., and Woodsmith, R. D.: Effects of wildfire on catchment runoff response: a modelling approach to detect changes in snow-dominated forested catchments, Hydrology Research, 41, 378-390, 2010.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture–climate interactions in a changing climate: A review, Earth-Science Reviews, 99, 125-161, http://dx.doi.org/10.1016/j.earscirev.2010.02.004, 2010.

Seneviratne, S. I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A., Cheruy, F., Higgins, M. E., Meier, A., Brovkin, V., Claussen, M., Ducharne, A., Dufresne, J.-L., Findell, K. L., Ghattas, J., Lawrence, D. M., Malyshev, S., Rummukainen, M., and Smith, B.: Impact of soil moisture-climate feedbacks on CMIP5 projections: First results from the GLACE-CMIP5 experiment, Geophysical Research Letters, 40, 5212-5217, 10.1002/grl.50956, 2013.

Shamir, E., Imam, B., Morin, E., Gupta, H. V., and Sorooshian, S.: The role of hydrograph indices in parameter estimation of rainfall–runoff models, Hydrological Processes, 19, 2187-2207, 10.1002/hyp.5676, 2005.

Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrological prediction, Hydrological Processes, 17, 2101-2111, 10.1002/hyp.1425, 2003.

Sperry, J. S., Hacke, U. G., Oren, R., and Comstock, J. P.: Water deficits and hydraulic limits to leaf water supply, Plant, Cell & Environment, 25, 251-263, 10.1046/j.0016-8025.2001.00799.x, 2002.

Swift, L. W., and Swank, W. T.: Long term responses of streamflow following clearcutting and regrowth / Réactions à long terme du débit des cours d'eau après coupe et repeuplement, Hydrological Sciences Bulletin, 26, 245-256, 10.1080/02626668109490884, 1981.

Teuling, A. J., Van Loon, A. F., Seneviratne, S. I., Lehner, I., Aubinet, M., Heinesch, B., Bernhofer, C., Grünwald, T., Prasse, H., and Spank, U.: Evapotranspiration amplifies European summer drought, Geophysical Research Letters, 40, 2071-2075, 10.1002/grl.50495, 2013.

Troch, P. A., Martinez, G. F., Pauwels, V. R. N., Durcik, M., Sivapalan, M., Harman, C., Brooks, P. D., Gupta, H., and Huxman, T.: Climate and vegetation water use efficiency at catchment scales, Hydrological Processes, 23, 2409-2414, 10.1002/hyp.7358, 2009.

Troch, P. A., Lahmers, T., Meira, A., Mukherjee, R., Pedersen, J. W., Roy, T., and Valdés-Pineda, R.: Catchment coevolution: A useful framework for improving predictions of hydrological change?, Water Resources Research, 51, 4903-4922, 10.1002/2015WR017032, 2015.

Tron, S., Perona, P., Gorla, L., Schwarz, M., Laio, F., and Ridolfi, L.: The signature of randomness in riparian plant root distributions, Geophysical Research Letters, 42, 7098-7106, 10.1002/2015GL064857, 2015.

Vose, J. M., Miniat, C. F., Luce, C. H., Asbjornsen, H., Caldwell, P. V., Campbell, J. L., Grant, G. E., Isaak, D. J., Loheide Ii, S. P., and Sun, G.: Ecohydrological implications of drought for forests in the United States, Forest Ecology and Management, http://dx.doi.org/10.1016/j.foreco.2016.03.025,

Wagener, T., Sivapalan, M., Troch, P., and Woods, R.: Catchment Classification and Hydrologic Similarity, Geography Compass, 1, 901-931, 10.1111/j.1749-8198.2007.00039.x, 2007.

Waichler, S. R., Wemple, B. C., and Wigmosta, M. S.: Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest, Hydrological Processes, 19, 3177-3199, 10.1002/hyp.5841, 2005.

Wang-Erlandsson, L., Bastiaanssen, W. G. M., Gao, H., Jägermeyr, J., Senay, G. B., van Dijk, A. I. J. M., Guerschman, J. P., Keys, P. W., Gordon, L. J., and Savenije, H. H. G.: Global root zone storage capacity from satellite-based evaporation, Hydrol. Earth Syst. Sci., 20, 1459-1481, 10.5194/hess-20-1459-2016, 2016.

Weibull, W.: A Statistical Distribution Function of Wide Applicability, Journal of applied mechanics, 18, 293-297, 1951.

Westerberg, I. K., Guerrero, J. L., Younger, P. M., Beven, K. J., Seibert, J., Halldin, S., Freer, J. E., and Xu, C. Y.: Calibration of hydrological models using flow-duration curves, Hydrol. Earth Syst. Sci., 15, 2205-2227, 10.5194/hess-15-2205-2011, 2011.

Westra, S., Thyer, M., Leonard, M., Kavetski, D., and Lambert, M.: A strategy for diagnosing and interpreting hydrological model nonstationarity, Water Resources Research, 50, 5090-5113, 10.1002/2013WR014719, 2014.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Elsevier Science, 2005.

Yadav, M., Wagener, T., and Gupta, H.: Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins, Advances in Water Resources, 30, 1756-1774, http://dx.doi.org/10.1016/j.advwatres.2007.01.005, 2007.

Yilmaz, K. K., Gupta, H. V., and Wagener, T.: A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model, Water Resources Research, 44, n/a-n/a, 10.1029/2007WR006716, 2008.

Zhang, S., Yang, H., Yang, D., and Jayawardena, A. W.: Quantifying the effect of vegetation change on the regional water balance within the Budyko framework, Geophysical Research Letters, 43, 1140-1148, 10.1002/2015GL066952, 2016.

Zhao, R.-J.: The Xinanjiang model applied in China, Journal of Hydrology, 135, 371-381, http://dx.doi.org/10.1016/0022-1694(92)90096-E, 1992.

Table 1. Overview of the catchments and their sub-catchments (WS).

Catchment	Deforestation period	Treatment	Area [km^2]	Affected Area [%]	Aridity index [-]	Precipitation [mm/year]	Discharge [mm/year]	Potential evaporation [mm/year]
HJ Andrews WS1	1962 -1966.	Burned 1966	0.956	100	0.39	2305	1361	902
HJ Andrews WS2	-	-	0.603	-	0.39	2305	1251	902
Hubbard Brook WS2	1965-1968	Herbicides	0.156	100	0.57	1471	1059	784
Hubbard Brook WS3	-	-	0.424	-	0.54	1464	951	787
Hubbard Brook WS5	1983-1984	No treatment	0.219	87%	0.51	1518	993	746

Table 2. Applied parameter ranges for root zone storage derivation

Catchment	$I_{max,eq}$ [mm]	$I_{max,change}$ [mm]	T_r [days]
HJ Andrews WS1	1-5	0-5	0-3650
HJ Andrews WS2	1-5	-	-
Hubbard Brook WS2	1-5	5-10	0-3650
Hubbard Brook WS3	1-5	-	-
Hubbard Brook WS5	1-5	0-5	0-3650

Table 3. Overview of the hydrological signatures

Signature	Description	Reference
S~~Q_M~~$_{QMA}$	Mean annual runoff	
S_{AC}	One day autocorrelation coefficient	Montanari and Toth (2007)
$S_{AC,summer}$	One day autocorrelation the summer period	Euser et al. (2013)
$S_{AC,winter}$	One day autocorrelation the winter period	Euser et al. (2013)
S_{RLD}	Rising limb density	Shamir et al. (2005)
S_{DLD}	Declining limb density	Shamir et al. (2005)
S_{Q5}	Flow exceeded in 5% of the time	Jothityangkoon et al. (2001)
S_{Q50}	Flow exceeded in 50% of the time	Jothityangkoon et al. (2001)
S_{Q95}	Flow exceeded in 95% of the time	Jothityangkoon et al. (2001)
$S_{Q5,summer}$	Flow exceeded in 5% of the summer time	Yilmaz et al. (2008)
$S_{Q50,summer}$	Flow exceeded in 50% of the summer time	Yilmaz et al. (2008)
$S_{Q95,summer}$	Flow exceeded in 95% of the summer time	Yilmaz et al. (2008)
$S_{Q5,winter}$	Flow exceeded in 5% of the winter time	Yilmaz et al. (2008)
$S_{Q50,winter}$	Flow exceeded in 50% of the winter time	Yilmaz et al. (2008)
$S_{Q95,winter}$	Flow exceeded in 95% of the winter time	Yilmaz et al. (2008)
S_{Peaks}	Peak distribution	Euser et al. (2013)
$S_{Peaks,summer}$	Peak distribution summer period	Euser et al. (2013)
$S_{Peaks,winter}$	Peak distribution winter period	Euser et al. (2013)
$S_{Qpeak,10}$	Flow exceeded in 10% of the peaks	
$S_{Qpeak,50}$	Flow exceeded in 50% of the peaks	
$S_{Qsummer,peak,10}$	Flow exceeded in 10% of the summer peaks	
$S_{Qsummer,peak,50}$	Flow exceeded in 10% of the summer peaks	
$S_{Qwinter,peak,10}$	Flow exceeded in 10% of the winter peaks	

$S_{Qwinter,peak,50}$	Flow exceeded in 50% of the winter peaks	
S_{SFDC}	Slope flow duration curve	Yadav et al. (2007)
S_{LFR}	Low flow ratio (Q_{90}/Q_{50})	
S_{FDC}	Flow duration curve	Westerberg et al. (2011)
$S_{AC,serie}$	Autocorrelation series (200 days lag time)	Montanari and Toth (2007)

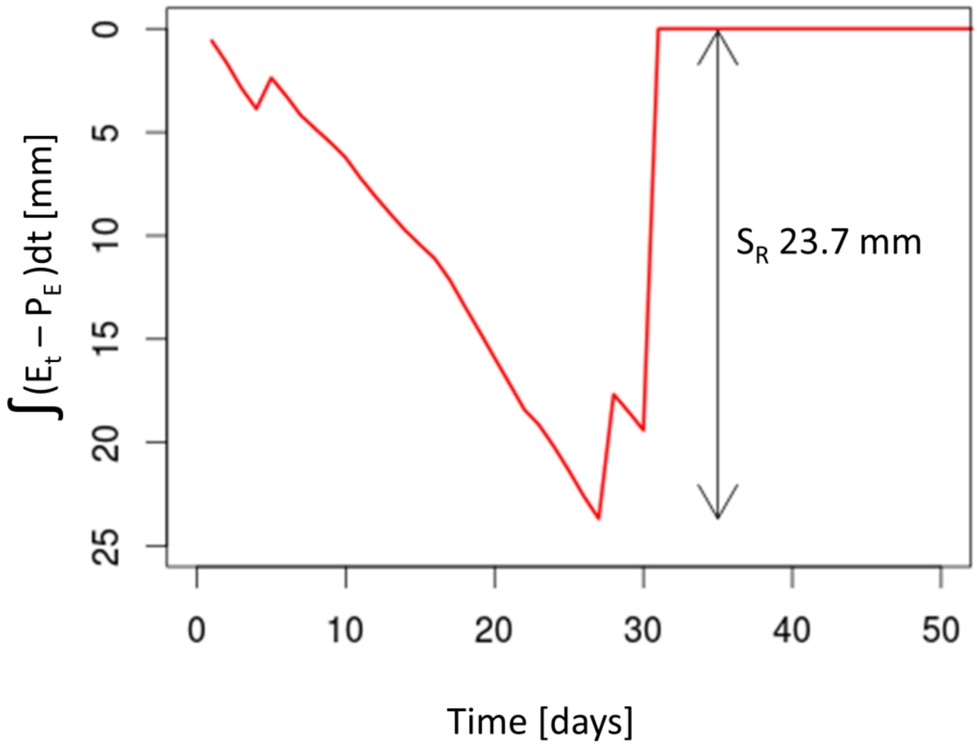

Figure 1. Derivation of root zone storage capacity (S_R) for one specific time period in the Hubbard Brook WS2 catchment as difference between the cumulative transpiration (E_t) and the cumulative effective precipitation (P_E).

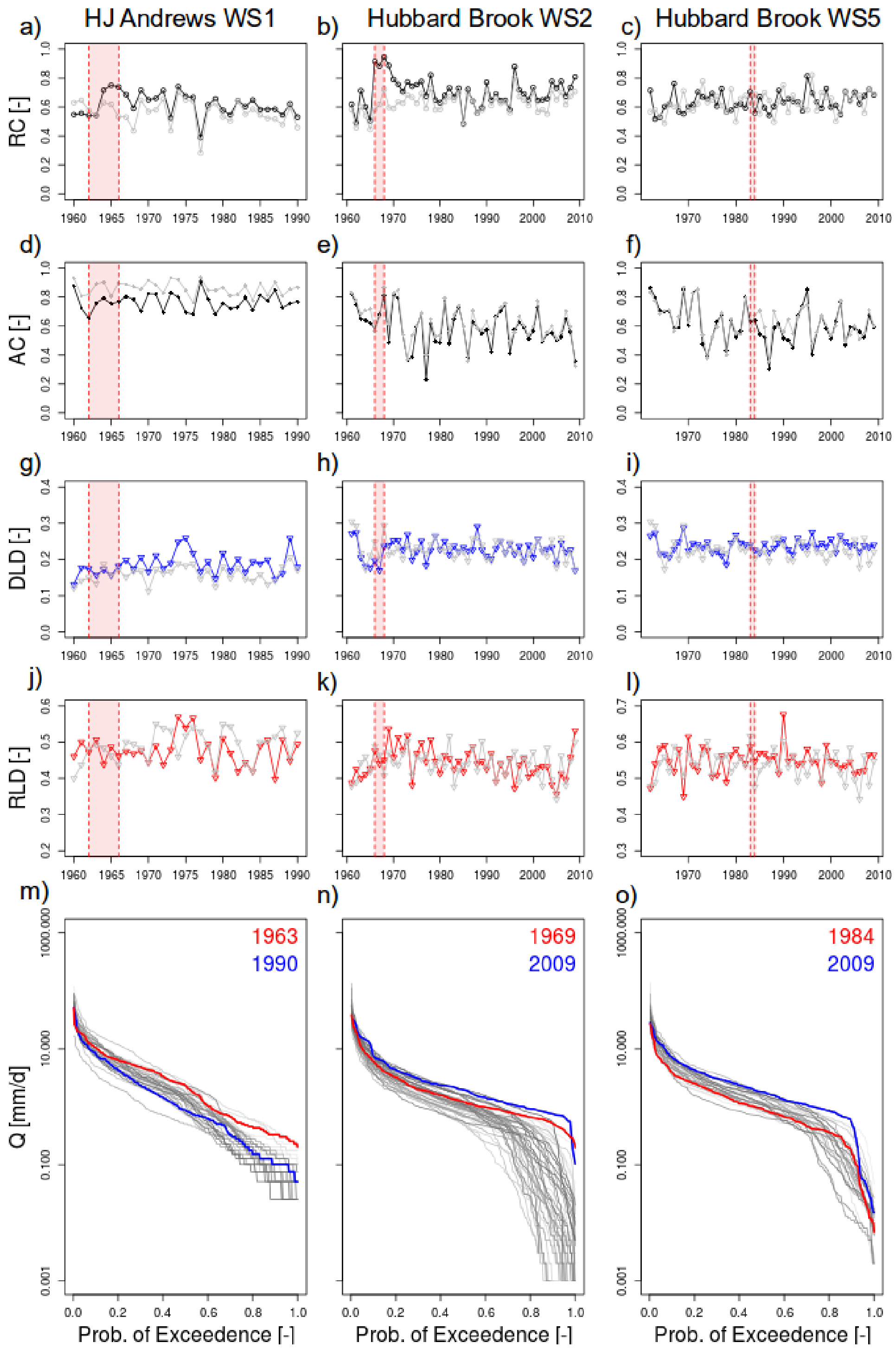

Figure 2. Evolution of signatures in time of a-c) the runoff coefficient, d-f) the 1-day autocorrelation, g-i) the declining limb density, j-l) the rising limb density with the reference watersheds in grey and periods of deforestation in red shading. The flow duration curves for HJ Andrews WS1, Hubbard Brook WS2 and Hubbard Brook WS5 are shown in m-o), where years between the first and last year are colored from lightgray till darkgrey progressively in time.

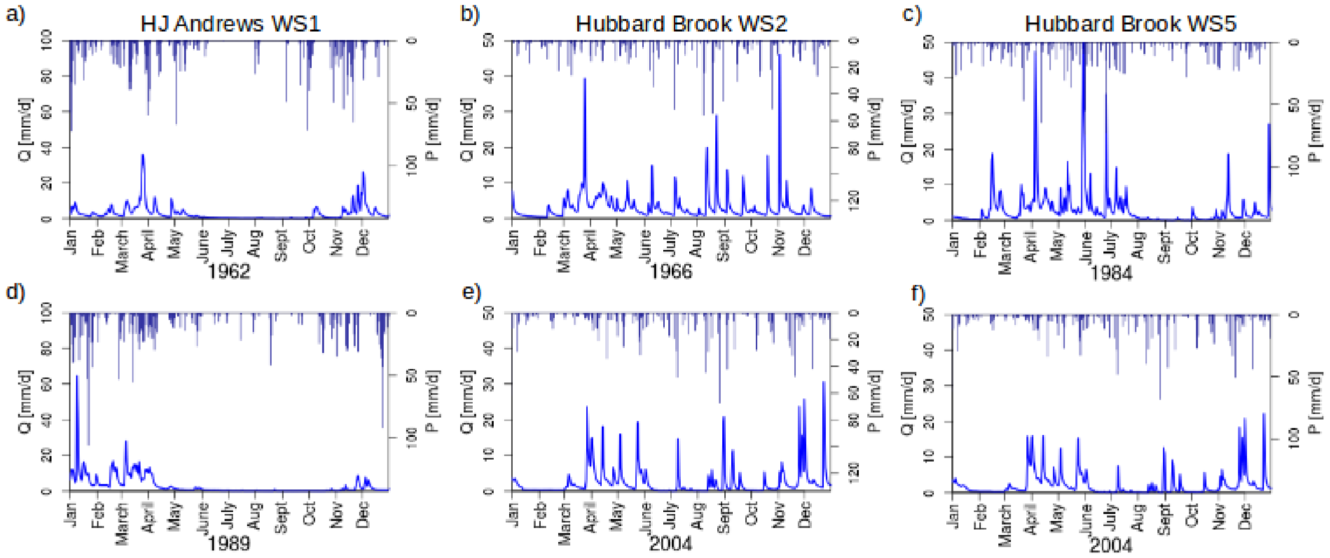

Figure 3. Hydrographs for HJ Andrews WS1 in a) 1963 (annual precipitation P_A=2018 mm yr^{-1}, $E_{p,A}$= 951 mm yr^{-1}) and b) 1989 (P_A= 1752 mm yr^{-1}, $E_{p,A}$= 846 mm yr^{-1}), Hubbard Brook WS2 in c) 1966 (P_A = 1222 mm yr^{-1}, $E_{p,A}$ = 788 mm yr^{-1} and d) 2004 (P_A = 1296 mm yr^{-1}, annual $E_{p,A}$ = 761 mm yr^{-1} and Hubbard Brook WS5 in e) 1984 (P_A=1480 mm yr^{-1}, annual $E_{p,A}$ = 721 mm yr^{-1}) and f) 2004 (P_A= 1311 mm $yr^{-1,}$ $E_{p,A}$ = 731 mm yr^{-1}).

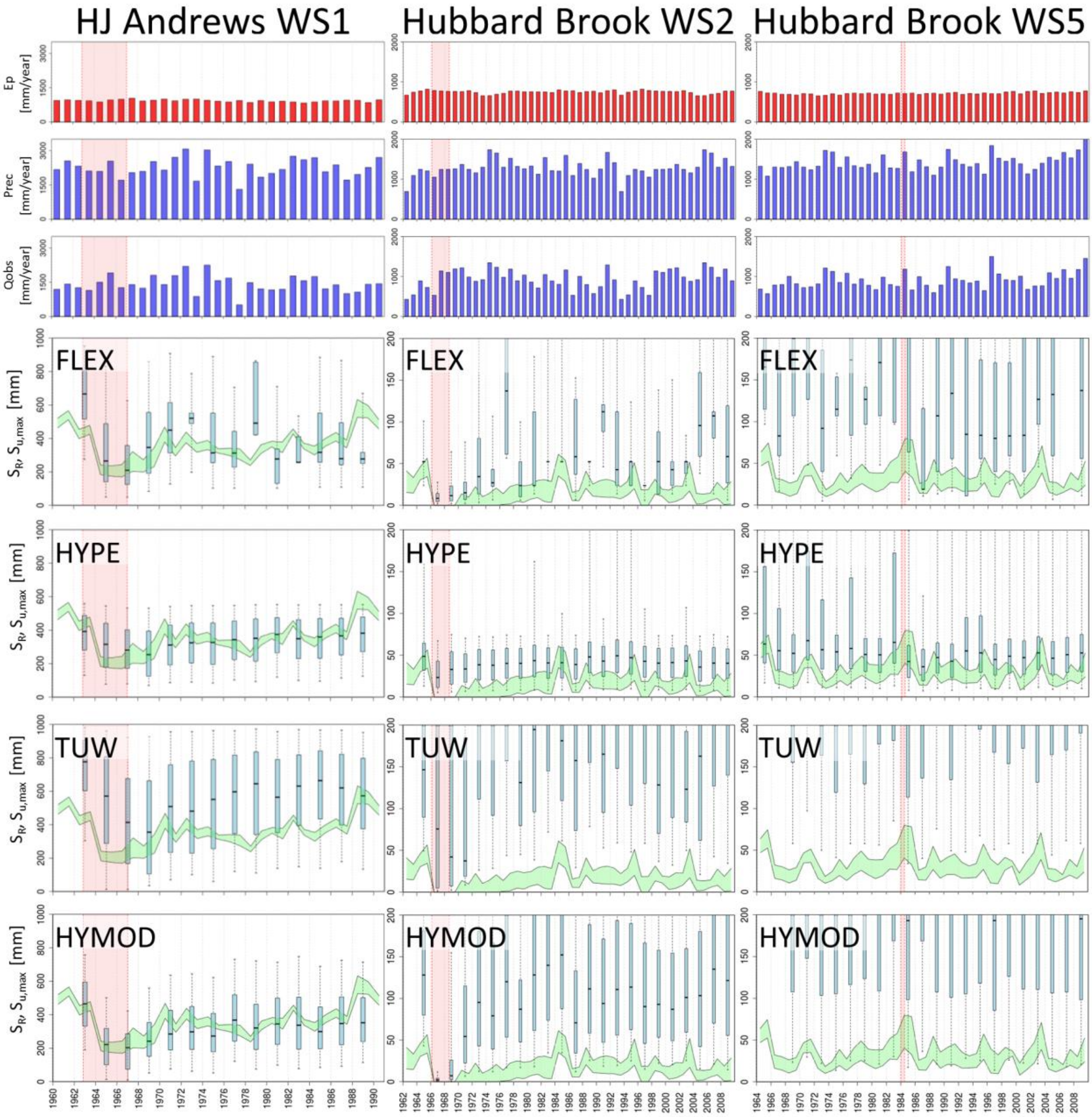

Figure 4. Evolution of root zone storage capacity $S_{R,1yr}$ from water balance-based estimation (green shaded area, a range of solutions due to the sampling of the unknown interception capacity) compared with $S_{u,max,2yr}$ estimates obtained from the calibration of four models (FLEX, HYPE, TUW, HYMOD; blue boxplots) for a) HJ Andrews WS1, b) Hubbard Brook WS2 and c) Hubbard Brook WS5. Red shaded areas are periods of deforestation.

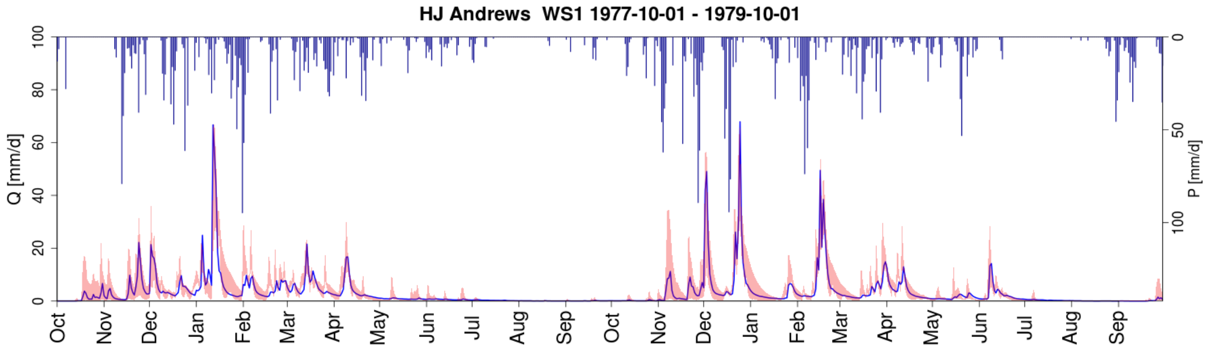

Figure 5. Observed and modelled hydrograph for HJ Andrews WS1 the years of 1978 and 1979, with the red colored area indicating the 5/95% uncertainty intervals of the modelled discharge. Blue bars show daily precipitation.

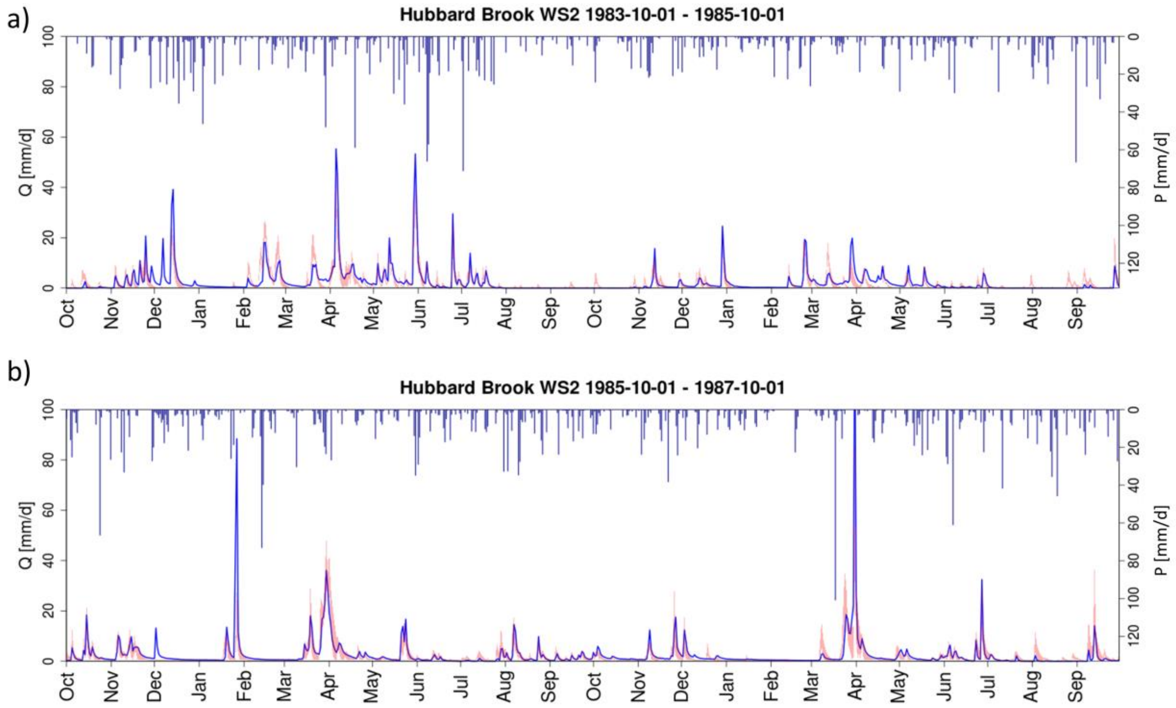

Figure 6. Observed and modelled hydrograph for Hubbard Brook WS2 for a) the years of 1984 and 1985 and b) the years of 1986 and 1987, with the red colored area indicating the 5/95% uncertainty intervals of the modelled discharge. Blue bars show daily precipitation.

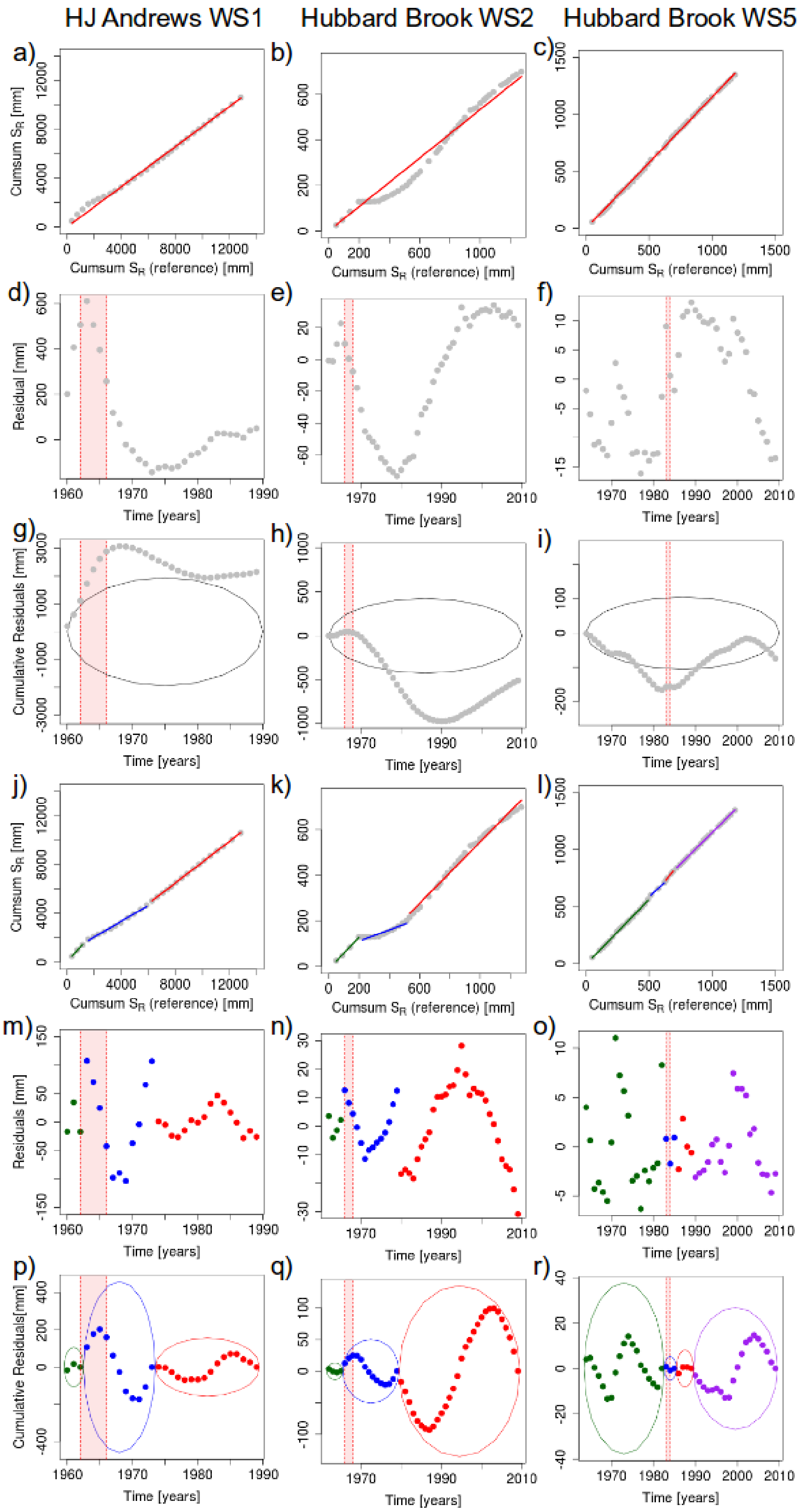
HJ Andrews WS1
Hubbard Brook WS2
Hubbard Brook WS5
a)
b)
c)
d)
e)
f)
g)
h)
i)
j)
k)
l)
m)
n)
o)
p)
q)
r)
Cumsum S_R [mm]
Cumsum S_R (reference) [mm]
Residual [mm]
Cumulative Residuals [mm]
Residuals [mm]
Cumulative Residuals[mm]
Time [years]

Figure 7. Trend analysis for $S_{R,1yr}$ in HJ Andrews WS1, Hubbard Brook WS2 and WS5 based on comparison with the control watersheds with a-c) Cumulative root zone storages ($S_{R,1yr}$) with regression, d-f) residuals of the regression of cumulative root zone storages, g-i) significance test; the cumulative residuals do not plot within the 95%-confidence ellipse, rejecting the null-hypothesis that the two time series are homogeneous, j-l) piecewise linear regression based on break points in residuals plot, m-o) residuals of piecewise linear regression, p-r) significance test based on piecewise linear regression with homogeneous time series of $S_{R,1yr}$. The different colors (green, blue, red, violet) indicate individual homogeneous time periods.

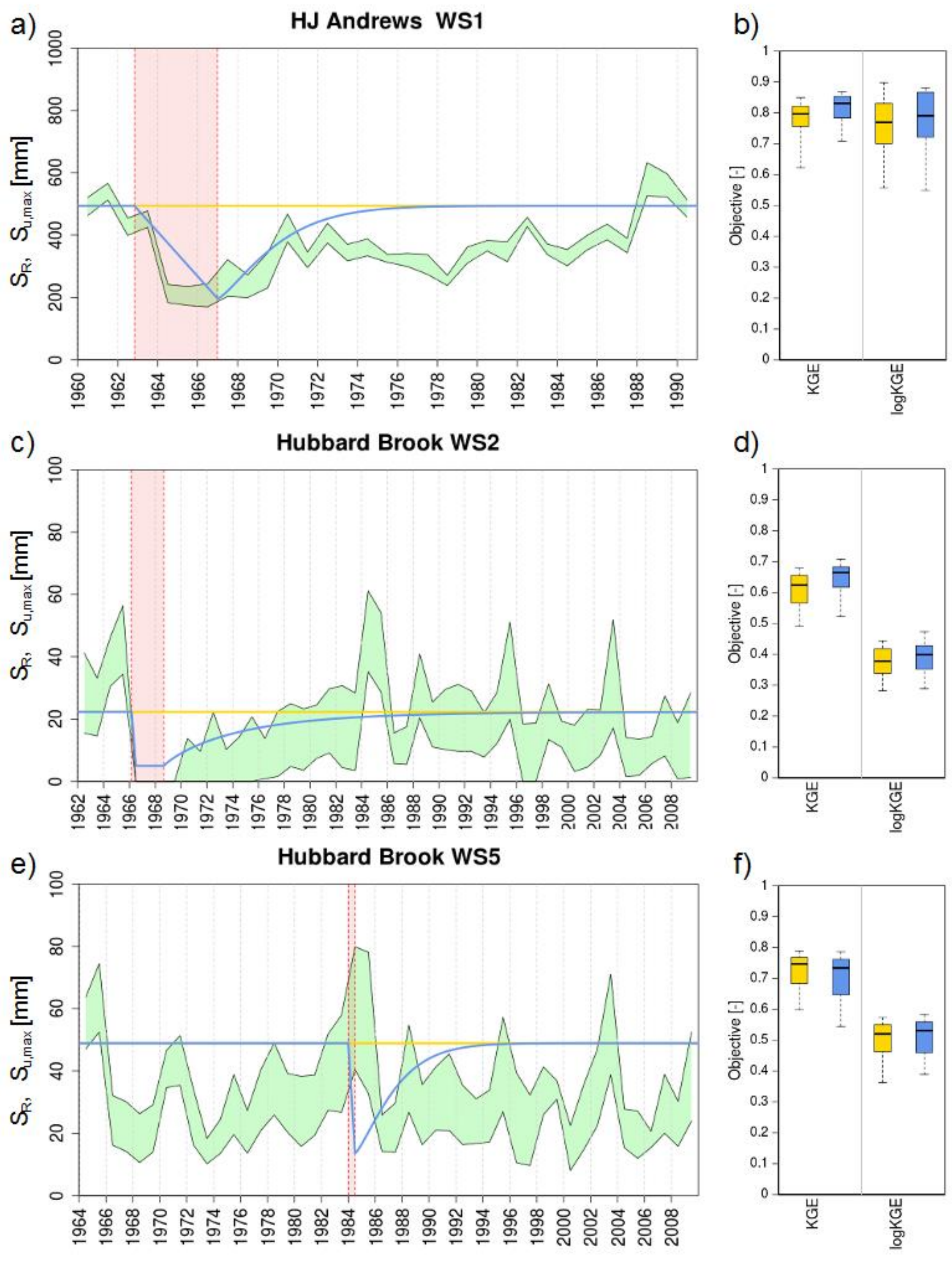

Figure 8. The time invariant $S_{u,max}$ formulation represented by $S_{R,\ 20yr}$ (yellow) and time dynamic $S_{u,max}$ fitted Weibull growth function (blue) with a linear reduction during deforestation (red shaded area) and mean 20-year return period root zone storage capacity $S_{R,\ 20yr}$ as equilibrium value for a) HJ Andrews WS1 with $a=0.0001\ days^{-1}$, $b=1.3$ and $S_{R,\ 20yr} = 494\ mm$ with b) the objective function values, c) Hubbard Brook WS2 with $a=0.001\ days^{-1}$,

b=0.9 and $S_{R,\ 20yr}$ = *22 mm* with d) the objective function values, and e) Hubbard Brook WS5 with *a=0.001* $days^{-1}$, *b=0.9* and $S_{R,\ 20yr}$ = *49 mm* and with f) the objective function values. The green shaded area represents the maximum and minimum boundaries of $S_{R,1yr}$ from the water balance-based estimation, caused by the sampling of interception capacities.

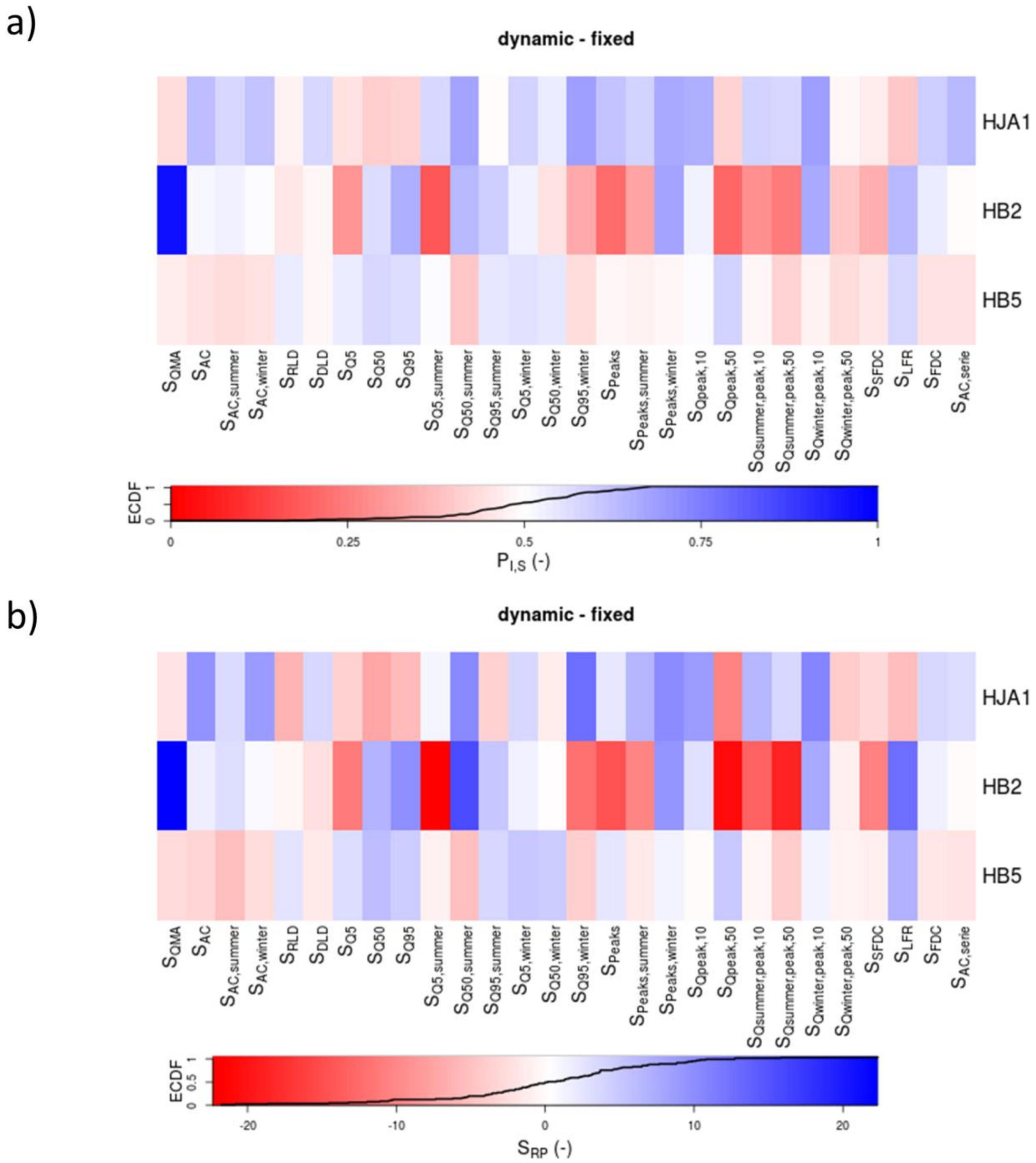

Figure 9. Signature comparison between a time-dynamic and time-invariant formulation of root zone storage capacity in the FLEX model with a) probabilities of improvement and b) Ranked Probability Score for 28 hydrological signatures for HJ Andrews WS1 (HJA1), Hubbard Brook WS2 (HB2) and Hubbard Brook WS5 (HB5). High values are shown in blue, whereas a low values are shown in red.

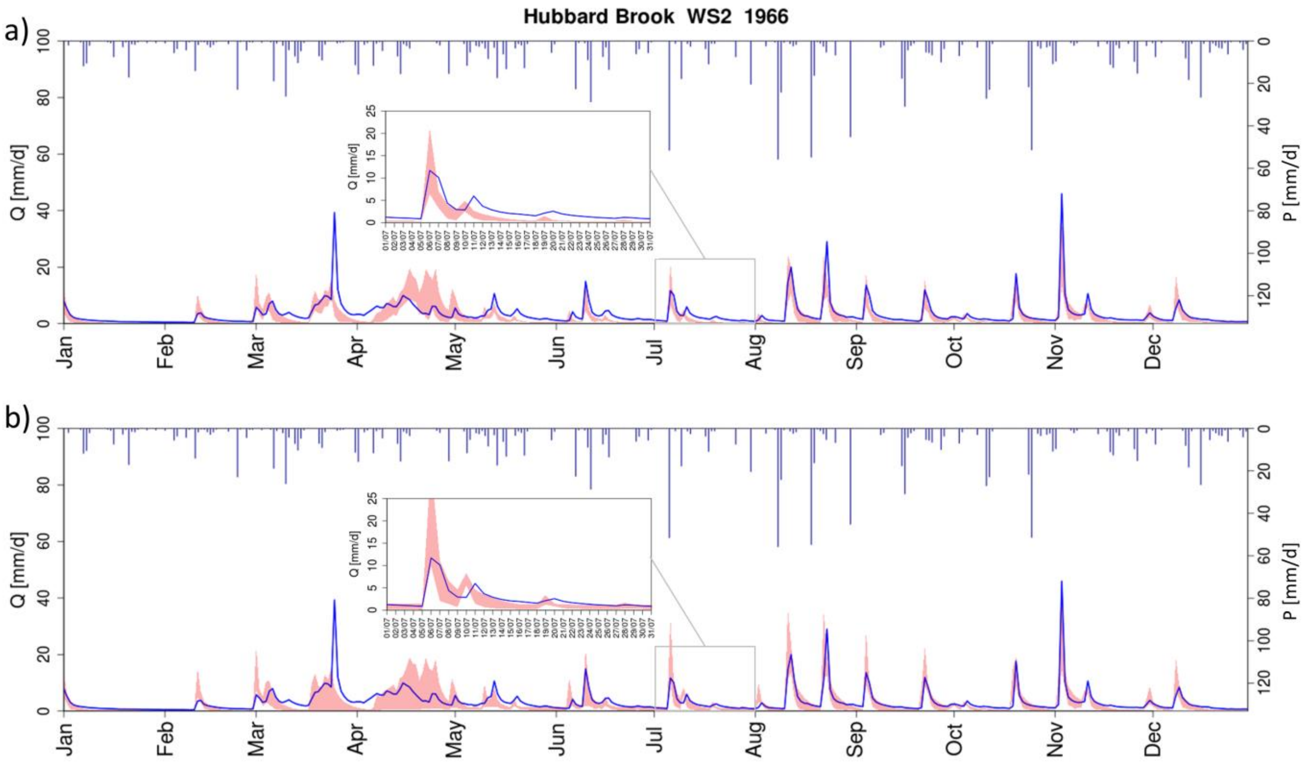

Figure 10. Hydrograph of Hubbard Brook WS2 with the observed discharge (blue) and the modelled discharge represented by the 5/ 95% uncertainty intervals (red), obtained with a) a constant representation of the root zone storage capacity $S_{u,max}$ and b) a time-varying representation of the root zone storage capacity $S_{u,max}$. Blue bars indicate precipitation.