Land-cover impacts on streamflow: a change-detection modelling approach that incorporates parameter uncertainty

Jan Seibert[1,2] & Jeffrey J. McDonnell[3]

[1] *Department of Geography, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland*
jan.seibert@geo.uzh.ch

[2] *Department of Physical Geography and Quaternary Geology, Stockholm University, Sweden*
jan.seibert@natgeo.su.se

[3] *Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis, Oregon 97331-5706, USA*

Received 19 December 2008; accepted 12 October 2009; open for discussion until 1 October 2010

Citation Seibert, J. & McDonnell, J. J. (2010) Land-cover impacts on streamflow: a change-detection modelling approach that incorporates parameter uncertainty. *Hydrol. Sci. J.* **55**(3), 316–332.

Abstract The effect of land-use or land-cover change on stream runoff dynamics is not fully understood. In many parts of the world, forest management is the major land-cover change agent. While the paired catchment approach has been the primary methodology used to quantify such effects, it is only possible for small headwater catchments where there is uniformity in precipitation inputs and catchment characteristics between the treatment and control catchments. This paper presents a model-based change-detection approach that includes model and parameter uncertainty as an alternative to the traditional paired-catchment method for larger catchments. We use the HBV model and data from the HJ Andrews Experimental Forest in Oregon, USA, to develop and test the approach on two small (<1 km^2) headwater catchments (a 100% clear-cut and a control) and then apply the technique to the larger 62 km^2 Lookout catchment. Three different approaches are used to detect changes in stream peak flows using: (a) calibration for a period before (or after) change and simulation of runoff that would have been observed without land-cover changes (reconstruction of runoff series); (b) comparison of calibrated parameter values for periods before and after a land-cover change; and (c) comparison of runoff predicted with parameter sets calibrated for periods before and after a land-cover change. Our proof-of-concept change detection modelling showed that peak flows increased in the clear-cut headwater catchment, relative to the headwater control catchment, and several parameter values in the model changed after the clear-cutting. Some minor changes were also detected in the control, illustrating the problem of false detections. For the larger Lookout catchment, moderately increased peak flows were detected. Monte Carlo techniques used to quantify parameter uncertainty and compute confidence intervals in model results and parameter ranges showed rather wide distributions of model simulations. While this makes change detection more difficult, it also demonstrated the need to explicitly consider parameter uncertainty in the modelling approach to obtain reliable results.

Key words change detection; forest hydrology; forest harvesting; HJ Andrews; HBV model

Impacts de l'occupation du sol sur les écoulements en rivière: une approche de détection de changement par modélisation qui inclut les incertitudes sur les paramètres

Résumé L'effet de la modification de l'occupation et de la couverture du sol sur la dynamique des écoulements reste mal compris. Dans la plupart des cas, la gestion des forêts est le facteur le plus important quant à la modification de l'occupation du sol. L'approche d'analyse des bassins versants par appariement est la principale méthode pour quantifier de tels effets. Cependant, cette méthode n'est adaptée que dans le cas de petits bassins versants de tête, où les précipitations incidentes ainsi que les caractéristiques des bassins sont uniformes entre les deux bassins. Cet article présente une approche de modélisation pour la détection de ces changements, qui inclut des calculs d'incertitudes sur le modèle et sur les paramètres, comme alternative à la méthode traditionnelle de l'appariement pour des bassins plus grands. Le modèle HBV a été appliqué à des données de la forêt expérimentale HJ Andrews, située dans l'Etat de l'Oregon, Etats-Unis, pour développer et tester l'approche sur deux petits bassins versants (<1 km^2) de tête (un bassin ayant fait l'objet d'une coupe rase à 100% et un bassin de contrôle), avant de l'appliquer au grand bassin versant de Lookout de 62 km^2. Trois approches différentes ont été employées pour détecter les changements dans les débits de pointe: (a) calage pour une période antérieure (ou postérieure) à la coupe rase, puis simulation de l'écoulement qui aurait été observé en l'absence de changement de l'occupation du sol (reconstruction d'une série d'écoulement); (b) comparaison des valeurs des paramètres calés pour les périodes pré- et post-changement d'occupation du sol; et (c) comparaison des écoulements prévus avec les paramétrages calés pour les périodes pré- et post-changement d'occupation du sol. La validation des concepts de la détection de changement par

ISSN 0262-6667 print/ISSN 2150-3435 online

doi: 10.1080/02626661003683264
http://www.informaworld.com

modélisation montre que les débits de pointe ont augmenté dans le bassin versant de tête ayant subi la coupe rase, par rapport au bassin versant de contrôle, et que plusieurs valeurs de paramètres du modèle ont changé après la coupe rase. Des changements mineurs ont également été détectés dans le bassin de contrôle, illustrant le problème des fausses détections. Pour le bassin versant principal de Lookout, une augmentation modérée des débits de pointe a été mise en évidence. Des méthodes de simulation Monte Carlo, utilisées pour quantifier les incertitudes sur les paramètres et pour calculer les intervalles de confiance des résultats modélisés des fourchettes de paramétrage, ont montré des distributions assez larges pour les simulations effectuées. Cela complique la détection de changement et démontre la nécessité de considérer explicitement les incertitudes sur les paramètres dans la démarche de modélisation pour obtenir des résultats fiables.

Mots clefs détection de changement; hydrologie forestière; récolte forestière; HJ Andrews; modèle HBV

INTRODUCTION

Our ability to assess the effect of land-use or land-cover change on streamflow is limited. Bruijnzeel (2004) notes that with human population increasing rapidly in some parts of the world (with associated increases in living standards), per capita demand for water, timber and other forest products is increasing, and pressure on the world's remaining forests is growing steadily. Quantifying the hydrological impacts of various forestry operations (thinning, selective harvesting, clear-cutting with and without roads, and removal of understory or riparian vegetation) is still an important activity (Eisenbies *et al*., 2007; Wei *et al*., 2008). Schnorbus & Alila (2004) note that (until only very recently) forest hydrologists have relied almost exclusively on a single technique to pursue these research questions: the paired catchment approach. While paired catchment studies have helped answer many fundamental questions in forest management (Hewlett, 1982; Hamilton & King, 1983; Andréassian, 2004; Bonell & Bruijnzeel, 2005), paired catchment studies are only possible in small headwater catchments (typically <1 km^2) where precipitation inputs, soil and geology conditions, topography, and other variables may be more uniform between the treatment and control catchment. For larger catchments, precipitation inputs and catchment conditions may vary greatly in time and space. Bowling *et al*. (2000) noted this issue in a paired catchment study of larger catchments where about 25% of the precipitation events occurred only in one of the paired catchments. For larger snow-dominated catchments this limitation might be less severe (Troendle *et al*., 2001; Moore & Scott, 2005). Land-use or land-cover (LULC) changes are typically more gradual for larger catchments and often occur over only a portion of the total catchment area. As a result it is usually impossible to find suitable control catchments beyond the headwater scale (Siriwardena *et al*., 2006).

So how might we quantify the effect of forest harvesting on streamflow at the larger catchment scale? Change detection modelling (Kundzewicz & Robson, 2004) may be a way to deal with LULC change detection in larger catchments where suitable control catchments are difficult to find, thus making the paired catchment approach impossible. Change detection modelling is a straightforward use of a precipitation–runoff model (Kuczera, 1987), but surprisingly few studies have used models in this way. Andréassian *et al*. (2003) used a rainfall runoff model to detect gradual changes in catchment behaviour. Others have used the well-known HBV model (the model used in our analyses and described in detail later in the paper) to investigate the effects of clear-cutting on streamflow response (Brandt *et al*., 1988).

While some LULC change detection modelling studies have already been completed, none have examined model and parameter uncertainty in a change detection modelling context. It is generally accepted that different parameter sets might perform equally well for a certain simulation period, but might give varying predictions when used for a different period. One notable exception is Siriwardena *et al*. (2006), who examined eight different parameter sets derived for a conceptual runoff model applied in Australia using different calibration strategies. Clearly, it is important to consider more than just one single "optimal" parameter set. In this study we use a Monte Carlo approach to tackle this issue. We use the well-known Jones & Grant (1996) data set, which includes two headwater catchments and a 62 km^2 catchment at the HJ Andrews Experimental Forest in Oregon, USA (HJA). We first test the change detection modelling approach for land cover changes in two headwater catchments as a proof of concept. Here we use two catchments from the Jones & Grant data set: a control and a treatment (100% clear-cut). We evaluate changes by examining model residuals, model parameters, and comparison of model simulations using the HBV model. Our work is different to previous work in that a control catchment is seldom used to test the possibility of falsely detecting a change. After demonstrating proof of concept at the headwater scale with control and treatment, we examine the effect of land-cover changes at the larger 62-km^2 scale with the same approach, perform runoff reconstruction for analysis of model residuals, and characterize and compare the runoff dynamics through analyses of model parameters calibrated for different periods and model simulations with those parameters. The objective of this paper was to propose a modelling approach for change detection, which also considers model parameter uncertainty, and to test this approach for detection of land-cover -change effects on floods for catchments in HJA, Oregon.

METHODS

The HBV model

The HBV model (Bergström, 1976; Lindström *et al.*, 1997) is a conceptual precipitation–runoff model that simulates discharge using a daily time step. Driving variables are precipitation and temperature as well as estimates of long-term averages of monthly potential evaporation. The model consists of different routines in which snowmelt is computed by a degree-day method, groundwater recharge and actual evaporation are functions of actual water storage in a soil box, runoff formation is represented by three linear reservoir equations, and channel routing is simulated by a triangular weighting function (see Table 1 for a list of the 14 model parameters). For both the snow and the soil routines, calculations are performed for each different elevation zone, while the lower box of the groundwater routine is a lumped representation of the catchment. Further descriptions of the model can be found in the appendix and elsewhere (e.g. Bergström, 1992; Lindström *et al.*, 1997; Seibert, 1997). The version of the model used in this study, "HBV light", corresponds in general to the original version described by Bergström (1992) with the exception that, while the upper box of the groundwater routine is treated as lumped for the entire catchment in the original version, it is computed individually for each elevation zone here (see also Uhlenbrook *et al.*, 1999). The parameters in the HBV model each have a physical meaning, but they are not measurable since they represent effective values at the catchment scale.

Study catchments

The study catchments are located within the HJ Andrews Experimental Forest (HJA) in the central western Cascade Mountains of Oregon, USA (44.2°N, 122.2°W). The main drainage within the HJA is Lookout Creek (LOOK, 62 km^2). Past process-based hydrological investigations at HJA have focused on runoff generation (Harr, 1977), snowmelt and snow accumulation (Harr, 1986; Berris & Harr, 1987), catchment nutrient budgets (Sollins *et al.*, 1980), and water residence time (McGuire *et al.*,

Table 1 Model parameters and feasible ranges.

Parameter	Explanation	Unit	Lower bound	Upper bound
Snow routine				
P_{TT}	Threshold temperature	°C	−1.5	2.5
P_{CFMAX}	Degree-day factor	mm °C^{-1} d^{-1}	1	10
P_{SFCF}	Snowfall correction factor*	-	0.5	1.2
P_{CWH}	Water holding capacity	-	0	0.2
P_{CFR}	Refreezing coefficient	-	0	0.1
Soil routine				
P_{FC}	Maximum of S_{SOIL} (storage in the soil)	mm	50	500
P_{LP}	Threshold for reduction of evaporation (S_{SOIL}/P_{FC})	-	0.3	1
P_{BETA}	Shape coefficient	-	1	6
P_{CET}	Factor for correction of long-term evaporation rates based on temperature	-	0	0.3
Response routine				
P_{K0}	Recession coefficient (upper storage)	d^{-1}	0.1	0.5
P_{K1}	Recession coefficient (upper storage)	d^{-1}	0.05	0.3
P_{K2}	Recession coefficient (lower storage)	d^{-1}	0.001	0.1
P_{UZL}	Threshold for the P_{K0}-outflow	mm	0	50
P_{PERC}	Maximal flow from upper to lower box	mm d^{-1}	0	4
P_{MAXBAS}	Routing, length of weighting function	d	1	7

*This parameter also compensates for evaporation from the snow storage which is not simulated explicitly in the model.

2005). Much of the work at HJA has examined the effects of forest management activities on water yield (Rothacher, 1965; Harr & McCorison, 1979) and sediment transport (Grant & Wolff, 1991). HJA is also the location for much of the analysis from Jones & Grant (1996) and the papers that followed (Beschta *et al.*, 2000; Jones & Grant, 2001; Thomas & Megahan, 1998; 2001) debating the interpretation of results of statistical analyses of paired-catchment data. Detailed site descriptions of the overall HJA and the small basins can be found in Rothacher *et al.* (1967), Jones & Grant (1996) and Jones (2000). Our study focuses on the small catchments WS1 (treatment, 0.96 km^2) and WS2 (control, 0.60 km^2) and the larger Lookout Creek catchment (LOOK, 62 km^2). Elevations range from about 450 to 1000 m for the small catchments and up to 1600 m for the Lookout Creek catchment.

For each catchment a series of runoff peaks was derived. Our rule for including an event was based on a threshold; flow rates had to exceed two times the long-term mean runoff. Only the highest peak within any 10-day period was included to avoid counting multiple peaks from the same event. On average there were six such events per year. The events were grouped into large, medium and small events. The threshold between large and medium events was set to a specific discharge of 50 mm d^{-1}, which corresponds approximately to a two-year return-period peak-runoff value. A value of 25 mm d^{-1} was used to separate the medium and small events.

The long-term mean annual precipitation varies from about 2300 mm at lower elevations to 3550 mm at upper elevations. Most of the precipitation (~80%) falls between November and April, typically during long-duration frontal storms of low to moderate intensity. In the small catchments precipitation falls mainly as rainfall with snow more common at higher elevations of LOOK. While winters are generally wet and mild (average January temperature of 1°C at 430 m), summers are dry and rather cool (average of 18°C in July).

The longest climate record for HJA is available from the station CS2MET (44.12.54°N, 122.14.57°W, 485 m). Since precipitation measurements at this station did not start before 1957, we extended the precipitation time series using a station outside the HJA. For this we selected the McKenzie Bridge RS station (NWS station 5362, 44.11°N, 122.07°W, 451 m, distance ~10 km) and a correction factor of 1.26, which was based on comparison of the overlapping observation period (20 years). Temperature has been measured at CS2MET since 1958. For extending the temperature data record, we used data from the Cascadia State Park station (NWS station 1433, 44.24°N, 122.29°W, 262 m, distance ~25 km) in addition to the McKenzie Bridge RS station data. In order to represent conditions at CS2MET these data were adjusted by adding a correction constant to the observed values at the two respective stations. Data from the overlapping observation periods (20 years) strongly suggested the use of seasonally varying correction constants. The correction constant added to the temperature from the two stations varied between –2.3°C and –0.2°C and –2.5°C and +1.2°C, respectively, with higher values during the summer months. When data from both stations were available the average value was used.

For evaporation, long-term average monthly values based on evaporation pan (Class A, four-foot diameter) measurements at the station Detroit dam (NWS station 2922, 44.43°N, 122.15°W, 372 m, distance ~50 km) were used. These long-term values were modified based on daily temperature anomalies (Lindström & Bergström, 1992) (see equation (A7)).

Jones & Grant (1996) described the land-cover change history at HJA. We refer the reader to that publication for a complete treatment of this time series. Briefly, the small basin treatments at HJA occurred in the late 1950s and early 1960s. WS1 was 100% clear-cut from 1962 to 1966 and broadcast burned in 1966. Clear-cutting in the Lookout Creek catchment was more gradual (between 1 and 10% during a 10-year period), with the greatest harvesting rate during the 1960s (Fig. 1) and only slight clear-cutting after 1970.

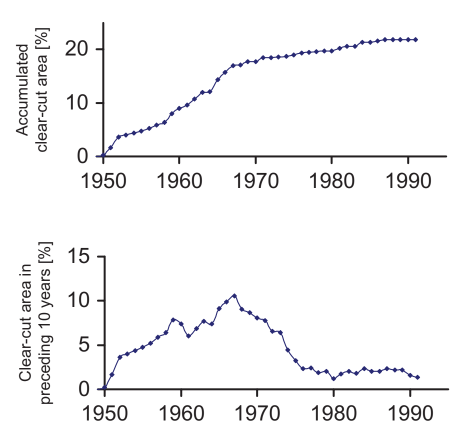

Fig. 1 Harvesting history for the Lookout Creek (LOOK) catchment.

Model application

The HBV model was applied to the different catchments using daily precipitation and temperature series as well as long-term mean monthly potential evaporation. The catchments were divided into different elevation zones (one per 100 m) and lapse rates were used for temperature (–0.5°C per 100 m) and precipitation (+5 % per 100 m). These lapse rates were based on previous work in the region (Daly *et al.*, 1994; Lookingbill & Urban, 2003) and analysis of data from stations in or neighbouring HJA, which were available for shorter periods. Time series of observed runoff were used for model calibration. The total time series of about 45 years was divided into 8-year periods to compromise between the need for a representative period for calibration and a suitable resolution for the detection of land-cover-change effects. We recognize that this choice of modelling intervals is somewhat arbitrary. Too short an interval would mean that we would not have enough data for calibration and therefore problems constraining the model. With increasing interval length there is less resolution to detect changes over time and treatment effects might change during the interval due to forest re-growth. Our 8-year segment choice was further motivated by the rule of the thumb that one needs 5–10 years of data to calibrate models like the HBV. The beginning of the first period was chosen according to data availability and, in the case of the small experimental catchments, to assure that the pre- and post-treatment periods fell into distinct periods. In each case at least one year was used as a warming-up period.

For each catchment, 300 000 parameter sets were generated with parameter values randomly chosen within specified feasible ranges (Table 1). The model was run with each of these parameter sets and the agreement between observed (Q_{obs}) and simulated (Q_{sim}) catchment runoff for the different 8-year periods was evaluated by the model efficiency (Nash & Sutcliffe, 1970), here denoted R_{eff}:

$$R_{eff} = 1 - \frac{\sum (Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - \overline{Q_{obs}})^2} \qquad (1)$$

Collections of the best (i.e. highest efficiency values) 30 parameter sets for each period were determined. Only these collections of parameter sets were used for further analysis. The number 30 was chosen arbitrarily, but results did not vary significantly as long as the number was large enough to capture the variability among the "best" parameter sets and small enough to ensure that only the very best parameter sets were chosen.

It is important to note that we used a model with a daily time step and, thus, examined daily peak flow rates, whereas the studies by Jones & Grant (1996), Thomas & Megahan (1998) and Beschta *et al.* (2000) used instantaneous peak flow rates. Daily and instantaneous peak flows of course differ. For the large events examined in our study the daily peak flow rates we used were on average about 30% lower than the instantaneous peak flow rates. On the other hand, there was a strong correlation between daily and instantaneous peak flows ($r^2 > 0.95$) even for the two small catchments.

Change detection

Three different approaches were used to evaluate potential runoff changes. First we examined time series of model residuals, which means that we compared model simulations with parameters calibrated on a reference period with observed runoff for periods with potential change. We also compared parameter values of the best parameter sets for the different time periods. Finally we compared daily runoff peaks simulated by using the best parameter sets for the different time periods. For WS1 these different time periods could be clearly separated into the time series before and after clear-cutting. For the LOOK basin where harvesting and road construction occurred on a more protracted basis, changes were related to the harvesting history record. We used these three approaches to detect changes in flow and system behaviour for the three catchments WS1, WS2, and LOOK. The catchment WS2, in which there was no harvesting activity at all, was included as a test to detect false change.

Model residuals By relating observed runoff to the runoff simulated using a model and parameter set valid for some reference period, climatic influences on runoff can be filtered out. We assumed that the simulated runoff could be used to reconstruct runoff time series, which would have been observed given a certain meteorological forcing, and that residuals could be interpreted as the effect of changes in the catchment. For all the observed runoff peaks relative residuals, D_i, between observed (Q_{obs}) and simulated (Q_{sim}) peak flows were computed for each event and each parameter set:

$$D_i = \frac{Q_{\mathrm{obs},i} - Q_{\mathrm{sim},i}}{Q_{\mathrm{obs},i}} \quad (2)$$

Values of D_i should scatter around zero for events during the reference period and periods without any LULC change related change in runoff. Values are larger than zero if the model underestimates daily peaks; since the model represents conditions of the reference period, such an underestimation can be interpreted as an effect of LULC changes. To test whether there was a change in runoff between two periods, P_A and P_B, we used the null hypothesis, H_0, that there was no change in the values of D_i between the different periods. The alternative hypothesis, H_1, was that there was a change in either direction (i.e. two-sided test). The independence of the samples could be assumed because of the way the runoff events used to compute D_i were selected. The non-parametric Wilcoxon rank-sum test (also called the Mann-Whitney test) was used because the values of D_i cannot be assumed to be normally distributed. For this test all values for D_i are sorted and then sums are calculated for the rank numbers of the D_i computed from the simulations of the parameter set collections of the respective periods. Based on comparison of these rank sums, p values for differences in the D_i values can be computed and H_0 can be rejected or accepted depending whether the p value is above or below the significance level. Since we had 30 different parameter sets, this test resulted in 30 different p values. In this paper we present only results where the residuals were computed for parameter sets calibrated on the period 1954–1962. For the analysis of the residuals the events were grouped into periods of 8 years, with the exception of the last period which was 1986–1999.

Parameter values Model parameters might differ when the model is calibrated to different time periods. Obviously, this is expected when there has been a LULC change. The analysis of differences in parameter values is not straightforward, since various different parameter sets might be equally possible. Consequently, different values for a certain parameter might be found not only for different time periods but also for the same period. To tackle this problem of parameter uncertainty we compared distributions of parameter values rather than single values for each parameter. Again we used the Wilcoxon rank-sum test and the null hypothesis, H_0, was that both parameter-value distributions come from identical populations of parameter values.

Comparison of model simulations More interesting than the differences in individual parameter values are the implications of the combined parameter-value changes within the different parameter sets. One approach to evaluate the latter is to run the model for some scenario and to compare the simulated runoff. This could be done with purely synthetic input data. In this study, however, we used the observed meteorological data of all runoff events as "scenarios". Using these input data as climatic driving variables, runoff was simulated with the different parameter sets. Similar to the comparison between observed and simulated peak flows, relative deviations were calculated between the peak flows simulated by the different parameter sets using equation (3), where $Q_{\mathrm{sim}A}$ and $Q_{\mathrm{sim}B}$ are the peak flows simulated with parameter sets selected based on model performance for two different periods (Period A and Period B, respectively):

$$D_i^* = \frac{Q_{\mathrm{sim}A,i} - Q_{\mathrm{sim}B,i}}{\left(Q_{\mathrm{sim}A,i} + Q_{\mathrm{sim}B,i}\right)/2} \quad (3)$$

To summarize, the following procedure was used:

(1) The n best parameter sets for each period i (e.g. before and after a LULC change) were selected (parameter sets $P_{1..n,\ i}$).
(2) All these parameter sets were used to simulate runoff for the events during the entire period.
(3) The simulations using parameter sets from the different periods i were compared by computing the median relative difference for all events in the three groups (large, medium and small events).
(4) Simulations using 30 parameter sets calibrated on Period A were compared with 30 parameter sets calibrated on Period B, which means that there were 900 possible combinations. Based on all these possible combinations distributions of relative differences were derived. These distributions were characterized by their median and percentiles (10 and 90%).

RESULTS

Proof-of-concept for headwater catchments

The best parameter sets resulted in model efficiencies for the different simulation periods of 0.71–0.86 for WS1 and 0.77–0.83 for WS2. This result indicated that the HBV model generally was able to reproduce the observed runoff. However, as expected, the peak flows were not always simulated perfectly. When selecting the best parameter sets according to their performance (evaluated by the model efficiency) during the 1954–1962 period, the medians of the relative residuals of the peak flows during this period were typically around 0.1 (i.e. 10%) (Table 2, left column). For two events in WS1 during 1953 and 1955, the performance was especially poor (most probably due to poor input data) and these two events were excluded from the further analyses.

The residuals increased for the simulation periods following the clear-cut. For WS1 the relative residuals were about 0.4 for the following two periods and decreased again to about 0.2 for the final two simulation periods (Table 2, Fig. 2(a)). For the first two periods following the clear-cut the residuals were significantly larger than for the pre-clear-cut period for the large events; this was also the case for 1978–1986 and 1986–1999 (Table 2). For the control catchment WS2 the relative residuals were smaller than for WS1 and did not differ considerably from those of the calibration period (Fig. 2(b)). However, for large events during 1962–1970 and 1970–1978, there was a statistically significant increase of about 20%, whereas there was a significant decrease for small events for 1986–1999 by about 10%. It is important to note that the p values derived from the analyses of the residuals obtained using the different parameter sets varied considerable (Table 2). This demonstrates the importance of considering parameter uncertainty by using different acceptable parameter sets.

The change in parameter values showed a clear pattern for WS1 (Table 3). These changes should not be interpreted directly as certain physical changes, but allow discussion of the changed system behaviour of a catchment. There was a clear change of values for the groundwater routine parameters. All changes were towards a faster response (increase of recession coefficients (P_{K0}, P_{K1}, and P_{K2}) and decrease of threshold for the fastest outflow (P_{UZL}). Higher recession coefficient values will cause higher peak flows, but also a quicker recession. The decrease of P_{UZL} means that the additional outflow from the upper groundwater box starts contributing to runoff at a smaller storage in this box, again causing higher peak flows and a quicker recession. The only exception was the increase of the routing parameter P_{MAXBAS}. P_{MAXBAS} is a parameter that represents the routing of the simulated flow from the groundwater along the stream network. Higher values mean both an increased delay in and reduction of peak flows, which probably is a compensation for the quicker outflow from the groundwater boxes. The parameter P_{FC}, which is the maximum storage in the soil routine (including vegetation) decreased for the first periods after the clear-cut but increased for the 1986–1999 period. The parameter P_{LP} controls the reduction of potential evaporation as a function of soil moisture storage; P_{LP} is the fraction of the maximum soil storage below which evaporation is reduced linearly, and it increased for the three periods following the clear-cut,

Table 2 Relative differences (-) between model simulations and observations for daily peak flows using parameter sets calibrated for the 1954–1962 period (shaded), the median of the relative differences is given in bold and the 10 and 90% percentiles are given in parentheses, p values are given in italic (median and 10 and 90% percentiles), only cases significantly (median $p < 0.05$) different from the calibration period are shown in the table.

		Period 1954–1962	1962–70	1970–78	1978–86	1986–99
WS1	Large	0.09 (–0.03, 0.21)	0.40 (0.26, 0.50) *0.002 (0.001, 0.004)*	0.33 (0.24, 0.40) *0.003 (0.002, 0.007)*		
	Medium	0.19 (0.09, 0.30)				
	Small	0.10 (–0.03, 0.32)	0.32 (0.19, 0.43) *0.015 (0.003, 0.063)*	0.30 (0.15, 0.45) *0.024 (0.007, 0.079)*		
WS2	Large	0.09 (–0.01, 0.17)	0.15 (0.04, 0.23) *0.030 (0.023, 0.398)*	0.20 (0.11, 0.27) *0.024 (0.007, 0.091)*		
	Medium	–0.02 (–0.13, 0.11)				
	Small	0.10 (–0.03, 0.18)				–0.08 (–0.19, 0.01) *0.030 (0.008, 0.063)*
LOOK	Large	–0.06 (–0.15, 0.05)		0.31 (0.23, 0.38) *0.016 (0.008, 0.032)*		
	Medium	–0.02 (–0.11, 0.13)				
	Small	0.06 (–0.03, 0.13)	0.33 (0.26, 0.40) *0.005 (0.002, 0.016)*			

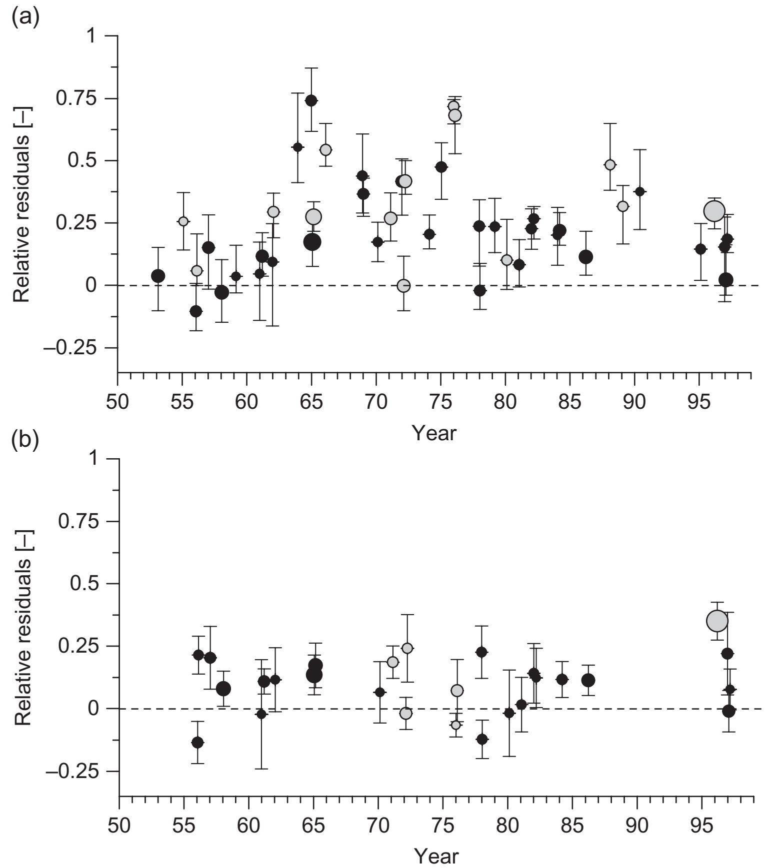

Fig. 2 Model residuals for peak flows simulated with the best parameter sets for the period 1954–1962 for: (a) WS1 and (b) WS2. Median values (circle) and range of 90% of the simulations using different parameter sets. The size of the circle indicates the relative magnitude of the peak flow. Black circles indicate rain events; grey circles indicate rain-on-snow events (i.e. events when there were at least 10 mm of snow storage prior to the event).

indicating a reduction in simulated evaporation rates. For the control catchment WS2 parameter changes were minor for most parameters with the exception of the parameters of the groundwater routine (Table 3).

The combined effect of the changed parameter values can be evaluated by simulation of the same runoff event using the same meteorological input but different collections of parameter sets. This can be illustrated by simulations using the meteorological input from the highest observed events at WS1. For all these meteorological input series higher peak flow values were simulated using the collections of parameter sets that performed best during the post-treatment periods compared with the collection of parameter sets from the pre-treatment period (Fig. 3). Results were similar for WS1 when using the meteorological input series from all observed runoff events. For WS1 the largest peaks were simulated when using the best parameter sets for the period 1954–1962 and peak flow simulations clearly decreased when using parameter sets from the other periods (Table 4 and Fig. 4). For the large flow events the difference was more than 20% compared to the best parameter sets for the period directly following the clear-cut and decreased to 13% for parameter sets from the period 1986–1994 (Table 4). Other detected changes indicated a decrease of peak flows for parameter sets for longer periods following the clear-cut. For the control catchment, WS2, the differences were much smaller (usually below 5%; Table 4). Again the different parameter sets resulted in rather wide distributions of relative differences and it was only for the cases where the median of this distribution was larger than ~10%, or lower than −10%, that both the 10 and 90% percentiles had the same sign (i.e. both positive or both negative).

Application to the 62 km² Lookout Creek catchment

The best parameter sets for LOOK resulted in model efficiencies for the different simulation periods of 0.80 to 0.88. The largest increase was observed for the

Table 3 Change in parameter values for the catchments WS1 and WS2 relative to the parameter values for the period 1954–1962 and for Lookout Creek (LOOK) relative to the parameter values for the period 1986–1994.

Parameter	Change in WS1 parameter values relative to the 1954–1962 values				Change in WS2 parameter values relative to the 1954–1962 values				Change in LOOK parameter values relative to the 1986–1994 values			
	1962–1970	1970–1978	1978–1986	1986–1994	1962–1970	1970–1978	1978–1986	1986–1994	1954–1962	1962–1970	1970–1978	1978–1986
P_{TT}							−−		+++		−−−	+
P_{CFMAX}				−								+++
P_{SFCF}						−				+++		
P_{CFR}												
P_{CWH}												
P_{FC}	−−−	−		++					−−−		−−−	−−−
P_{LP}	+++	+	+++							+++		
P_{BETA}										+++		+++
P_{CET}												
P_{PERC}												
P_{UZL}	−−	−−−	−−−	−−−					+	+++		
P_{K0}	+++		+++	+++	++		++	+++		+++		−
P_{K1}	+++	+++	+++	+++	++	+++	++					
P_{K2}	+++	+++	++	+++			−−−	−−−			−−−	−−−
P_{MAXBAS}	++	+++	+++	+++			+++	+++	−−−	−−−	−−	−−−

+ and - indicate an increase and decrease, respectively, of median of parameter values. +++/−−−: $p < 0.005$, ++/−−: $p < 0.01$; +/−: $p < 0.02$.

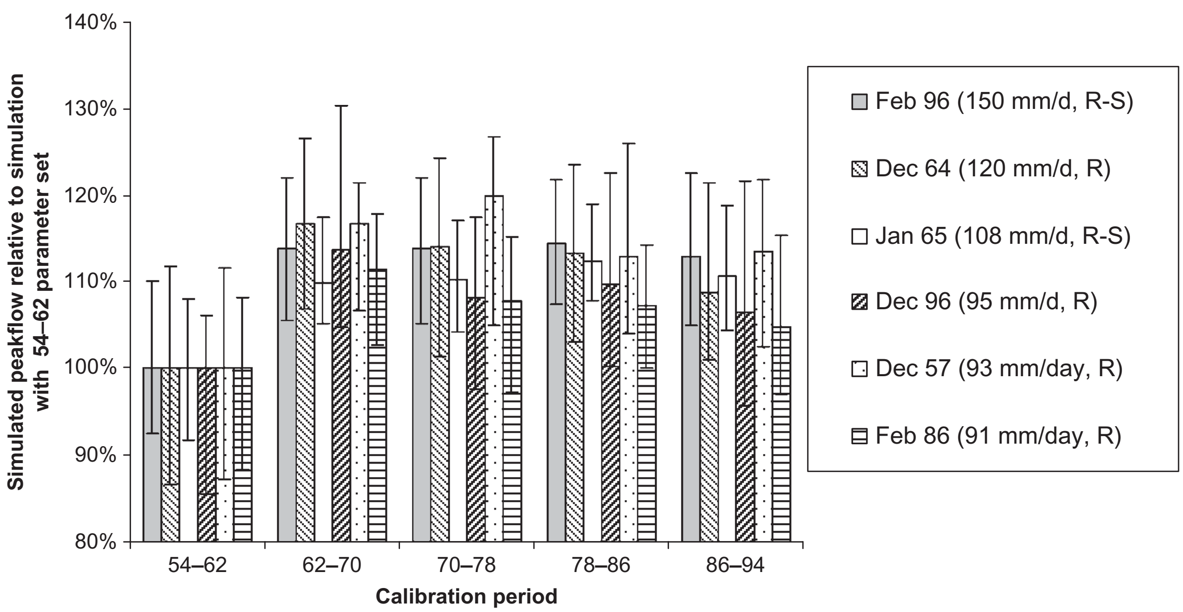

Fig. 3 Comparison of simulations of the six largest events on record for WS1 using the best parameter sets for different time periods. The bars indicate the median simulation of peak flow and the error bars indicate the range of 90% of the simulations. Note that all events were rescaled relative to the median simulation for the 1954–1962 period. (R: rain events, R-S: rain-on-snow events).

1962–1970 period, while differences were not significant for the later periods (Table 2, Fig.5). The changes were less pronounced at the LOOK scale than for WS1. For the analysis of parameter values we used the period 1986–1994 as a reference period for LOOK, because this was the period with the least clear-cutting activity in the preceding 10-year period (Fig.1). We found some significant changes in parameter values for LOOK (Table 3). The routing parameter MAXBAS was lower for all other periods, indicating

Table 4 Simulations of runoff from the three catchments WS1, WS2 and Lookout Creek (LOOK) using different parameter sets. Median relative differences when using parameter sets determined based on two calibration periods to simulate all peak flows (for large peak flows).

Period used to select first group of parameter sets	Period used to select second group of parameter sets				
	1954–1962	1962–1970	1970–1978	1978–1986	1986–1994
WS1					
1954–1962	**0**	**0.205**	**0.188**	**0.153**	**0.126**
1962–1970		0	–0.018	–0.057	–0.081
1970–1978			0	–0.038	–0.062
1978–1986				0	–0.030
1986–1994					0
WS2					
1954–1962	0	0.050	0.018	0.025	–0.027
1962–1970		0	–0.033	–0.026	–0.085
1970–1978			0	–0.007	–0.046
1978–1986				0	–0.054
1986–1994					0
LOOK					
1954–1962	0	**0.083**	–0.029	–0.054	–0.026
1962–1970		0	**–0.126**	**–0.148**	**–0.114**
1970–1978			0	–0.015	0.010
1978–1986				0	0.028
1986–1994					0

Bold: both 10 and 90% percentiles of the distribution of relative differences had the same sign, i.e. both positive, or both negative.

a faster streamflow response for the larger basin. The most significant parameter changes in LOOK were for the period 1962–1970, i.e. the period corresponding to the largest clear-cutting activities (Table 3). For this period the snow correction factor (P_{SFCF}) increased, indicating an increased snow accumulation. The increase of the evaporation parameter P_{LP} might indicate reduced evaporation, but the increase of shape factor P_{BETA} partly compensates for this. P_{BETA} controls the division of precipitation and snowmelt between water contributing to the soil storage and groundwater recharge; increasing values mean that for a certain (simulated) soil storage a larger portion of the incoming precipitation and snowmelt is added to the soil storage and can eventually evaporate. The increase of the recession coefficient for the upper outflow (P_{K0}) indicates a faster response for the largest events, but on the other hand the threshold (P_{UZL}) for this outflow contribution increased. The parameter set collection for LOOK for the period 1962–1970 predicted significantly higher peak flows than those of all other periods with a difference of about 10% (Fig. 6). Differences among the simulations using parameter set collections for other periods were smaller and not significant (Table 4).

DISCUSSION

Parameter change detection: a new way forward

Model parameters are highly interdependent. This causes the well-known problem of parameter identification (Beven, 2001, 2006), but also makes it difficult to relate parameter value changes to LULC changes. In our analyses, the parameter value changes for WS1 after the clear-cut showed a rather clear pattern of faster streamflow response to rainfall, whereas the pattern was less obvious for the entire LOOK where changes were less distinct. While the importance of parameter uncertainty has been emphasized in many recent studies (Beven, 2001, 2006), most LULC change detection model studies (Kuczera, 1987; Brandt *et al.*, 1988; Andréassian *et al.*, 2003; Kundzewicz & Robinson, 2004) have not incorporated parameter uncertainty. Our work shows the value of such an approach. The ranges for both model residuals and model predictions varied considerably implying the risk for significant over- or underestimation of change if only a single parameter set is used. If only one parameter set is used, then a difference between two periods may be the result of parameter uncertainty and not a change of the physical processes. The rather wide distributions of

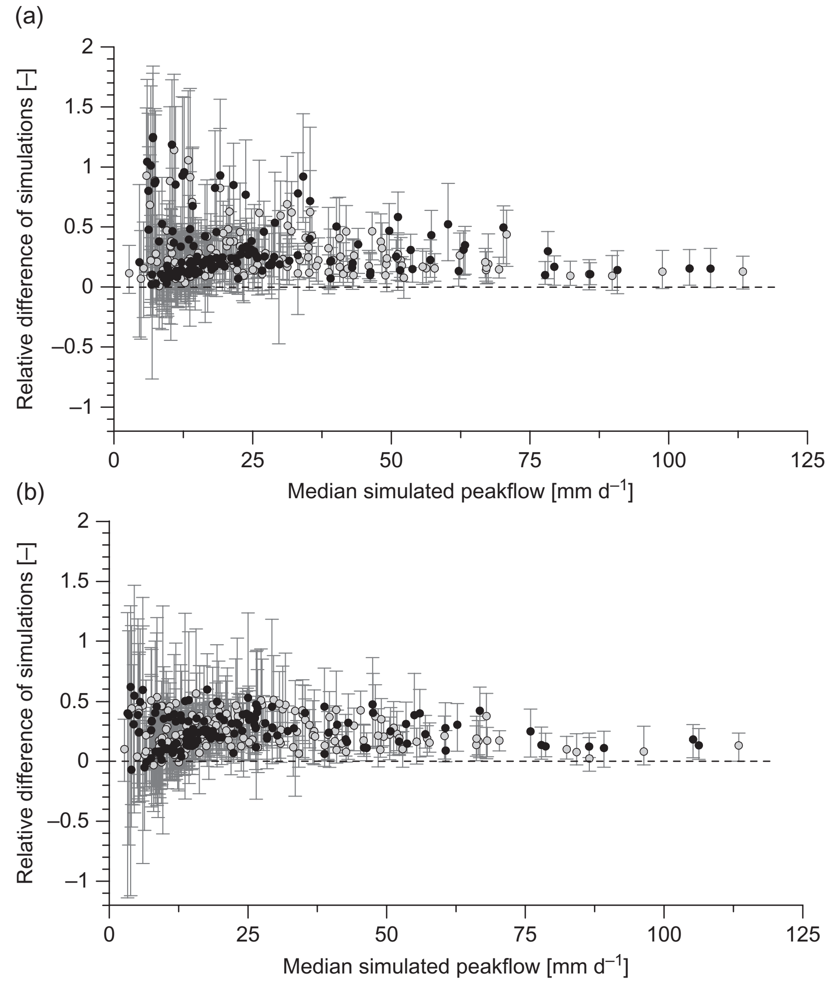

Fig. 4 Relative differences for WS1 between peak flow simulated (a) with the best parameter sets for the periods 1954–1962 and 1962–1970 respectively, and with those for the periods 1954–1962 and 1970–1978 (b), respectively. Median values (circle) and range of 90% of the simulations. Black circles indicate rain events and grey circles indicate rain-on-snow events (i.e. events when there were at least 10 mm of snow storage prior to the event). A specific discharge of 50 mm d^{-1} corresponds approximately to a 0.5 year event.

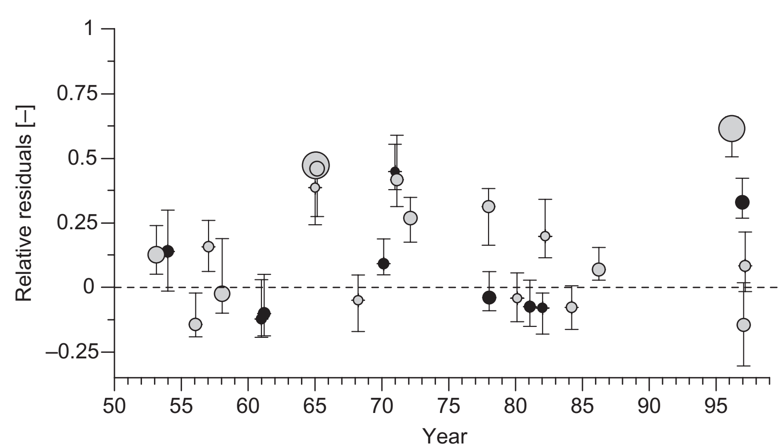

Fig. 5 Model residuals for peak flows simulated with the best parameter sets for the period 1954–1962 for LOOK. Median values (circle) and range of 90% of the simulations using different parameter sets. The size of the circle indicates the relative magnitude of the peak flow. Black circles indicate rain events; grey circles indicate rain-on-snow events (i.e. events when there were at least 10 mm of snow storage prior to the event).

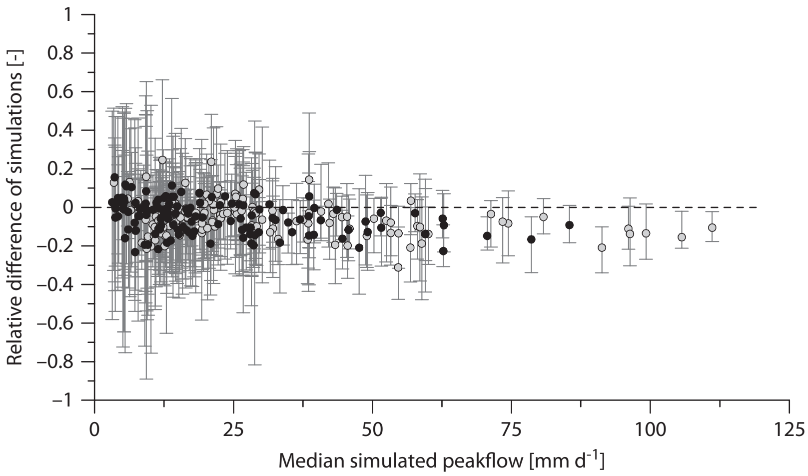

Fig. 6 Relative differences for LOOK between peak flow simulated with the best parameter sets for the periods 1962–1970 and 1986–1994, respectively. Median values (circle) and range of 90% of the simulations. Black circles indicate rain events and grey circles indicate rain-on-snow events (i.e. events when there was snow storage of at least 10 mm prior to the event). A specific discharge of 50 mm d^{-1} corresponds approximately to a 0.5 year event.

p values in our statistical analysis of the model residuals (Table 2) illustrate this risk.

The issue of parameter uncertainty is especially important when examining the change of parameter values. Even without a LULC change in the catchment, one might see changes in single parameters due to parameter uncertainty. In our analysis we addressed this problem by comparing a more robust measure of parameter values, namely parameter-value distributions, rather than single calibrated values. However, some values for the groundwater routine were found to change significantly for the control catchment for the different calibration periods, which clearly demonstrates that changes in individual parameters cannot solely be attributed to land-use changes.

Parameter interactions can complicate the interpretation of changes in parameter values. Therefore we also examined changes in system behaviour with the model-to-model comparison, i.e. the comparison of model simulations using different sets of parameters. The comparison of peak flows simulated using the best parameter sets for the different periods is a way to overcome the problems of parameter interactions. Instead of looking at individual parameter values, the system behaviour is investigated. We are not aware of papers that have used such a model-to-model comparison before, but would recommend this type of analysis for further studies. Unlike the analysis of changes in single parameters, the model-to-model comparison allows assessment of integrated catchment behaviour. In contrast to the analysis of residuals, this approach is less sensitive to errors in precipitation or runoff observations during single events.

In this study we used the HBV model, but the approach can easily be used with other runoff models. Such models might vary from more physically-based models (compared to the highly conceptual HBV model) to black-box models. The important point is that the model must be able to capture the rainfall–runoff relationship and to reproduce the runoff series. In other words, the particular structure of a model is of minor importance to the results, as long as the model can be calibrated to fit the observed hydrograph. The correct simulation of internal variables, which is crucial in many other model applications, is not prerequisite for this type of model application. For change detection, any black box (e.g. neural network) model that enables one to examine the catchment in a system response function manner can be used. The advantage of using a conceptual hydrological model is that we have a sense of what certain parameters might mean. If many catchments were being analysed, one could conceivably use this knowledge of how parameters change to apply the model in a forward (scenario modelling) sense. We used the HBV model in this study because we wanted a model where we had an idea of how the changing parameter values relate to the changing system behaviour, and also a model with few parameters that could be found through calibration. It should also be noted that the model performance of the HBV model in HBV WS1 and WS2 was similar to that of the more complex DHSVM model used by Waichler *et al.* (2005) for these same catchments. Had we used a more complex model with a large number of parameters, we could never have constrained these parameters by calibration. A practical reason for choosing a simple model is the computational demand; the approach presented in this paper requires thousands of multi-year model runs, which would not be feasible with a more complex model such as DHSVM. Having a more physically-based model with parameters that are more physically measurable (e.g. measured hydraulic conductivity values that can be made in the field and related directly to a physical model parameter) does not help us in such change-detection analysis. This is because we want, and need, the calibration itself to actually detect the changes. If the objective is the prediction, rather than the detection, of changes a more physically-based model would be necessary.

In the modelling approach for change detection the model is used for runoff reconstruction. This can be done with simple models as long as they can be fitted acceptably well to the observed runoff data. We might have been able to obtain a slightly better fit with a more complex model, but as parameter uncertainty would have increased, probably the results overall would not have been more certain. It is important to note that if a model is used for change prediction (e.g. Tague & Band, 2001) it becomes important that the relevant processes are modelled correctly. The simplified formulation for evaporation in the HBV model where interception is implicitly included in the soil routine is useful for change detection as it reduces parameter uncertainty. For change prediction, however, it is obviously important to distinguish between interception, transpiration and soil evaporation. A similar argument can be made for the importance of spatial patterns of model inputs and runoff routing, for which more detailed model formulations are needed when a model is used for change prediction rather than change detection (Uhlenbrook *et al.*, 2003; Blöschl *et al.*, 2007).

Peak flow responses in small and large basins

Jones & Grant (1996) examined paired discharges for 150–375 storm events for five basin pairs, using the long time series data record from the HJ Andrews Experimental Forest in Oregon and surrounding catchments in the Oregon Cascades. Jones & Grant (1996) reported that forest harvesting increased peak discharges by as much as 50% in small basins (60–101 ha) and 100% in large basins (60–600 km^2). The main mechanism that they suggested was responsible was the increased drainage efficiency of basins due to the integration of road/patch clear-cut networks with the pre-existing stream channel networks. In a statistical re-examination of the paired catchment findings that followed, Thomas & Megahan (1998) concluded that the Jones & Grant statistical analysis did not allow detection of any effect of cutting on peak flows in one of the large basin pairs and results were inconclusive in the two other large basins. While Thomas & Megahan (1998) conceded that peak flows were increased by up to 90% for the smallest peak events on the headwater clear-cut catchments, percent treatment effects were found to decrease as event size increased and were not detectable for flows with 2-year return intervals or greater on either of the headwater treated catchments. In the Beschta *et al.* (2000) statistical re-analysis of the Jones & Grant paired catchment data, peak flow increases averaged less than 13–16% after treatment for 1-year recurrence interval events and 6–9% for 5-year recurrence interval events. For the large basins, peak flow increases were weakly related to harvesting but were generally small (1–7%).

For the large Lookout Creek (LOOK) catchment, our analyses indicate a 10–30% change in peak flows on average, which is intermediate between that of Jones & Grant (1996), who reported a 100% change during the maximum occurrence of clear-cut activities (1962–1970), and those of Thomas & Megahan (1998). Our increases, while moderate, are larger than those reported by Beschta *et al.* (2000).

One obvious reason for the smaller changes in the Lookout Creek catchment is that here only a portion of the catchment was harvested. The smaller response to land cover changes might also be explained by the observation that the importance of other catchment characteristics than land cover increases for larger scales (Uhlenbrook *et al.*, 2003; Blöschl *et al.*, 2007). Scale issues together with the additional changes due to forest road construction might also be an explanation for the different parameters value changes which were observed for LOOK compared to WS1.

Much of the so-called Jones & Grant debate is based on statistical tests used for comparing pre- to post-logging streamflows and not the physical process controlling rainfall–runoff relations and how they are affected by forest removal. In the last paper in this series, Jones & Grant (2001, p. 177) note that: "… *the issue cannot be resolved with statistics based on a mere handful of extreme flood events. Future physical process based modelling and field studies will improve our understanding of forest harvest effects on these rare big floods* …". The modelling approach also is limited by the fact that the data set is dominated by a population of sub-annual peak flows (Beschta *et al*., 2000), that is peak flows with a recurrence interval of <1 year (that we term small and medium peak flows in our analyses), whereas extreme events, by definition, are rare. Nevertheless, the change detection approach described in this paper has the advantage that it combines a statistical analysis of runoff values and the analysis of changes in the system behaviour, as quantified by hydrological model parameters.

Ultimately, the best way forward will be to illuminate the black box of forested catchments and understand flow pathways, residence times, and stream sources (and we are actively engaged in these studies at the HJ Andrews – see McGuire *et al*., 2005, 2007). Nevertheless, the change detection modelling approach may be a way to deal with the many data sets where controversy lingers (and new sites where controversy will undoubtedly arise).

A self-critique of our change detection analyses and the approach in general

The change detection approach is not without its faults and our analyses not without their ambiguity. We observed some changes in residuals for our control catchment WS2, and we acknowledge this to be a problem. Thus false positive change (type I error) is possible even when there is no LULC change (as in the case of our control). No change was detected for WS1 after 1978 compared to pre-harvest conditions. This might indicate that the hydrological recovery was complete by then. However, there is also the risk for type II errors, i.e. no detection of an actual change. This is especially possible when changes are relatively small. In general though, we found big changes in residuals for WS1, intermediate changes for LOOK, and relatively minor changes for WS2 (the control catchment).

One extremely important point for any evaluation of runoff changes is the need for consistent data. In the case of paired catchment studies this applies to runoff measurements in both catchments. Both driving variables and observed runoff have to be consistent over time when using the model approaches, as discussed in this paper. The results might be biased if there were changes in measurement techniques or the location (or surrounding) of the measurement site, and there is no correction for these changes. It is acceptable if, for instance, precipitation measurements are somewhat incorrect all the time, because such time-invariant errors are implicitly taken into account by the model. On the other hand, misleading results might be obtained if the measurements change over time. For example, if the type of raingauge is changed and the systematic underestimation of measured precipitation is reduced, simulated runoff will apparently increase even if there is no change in the catchment.

For our test case we tested the homogeneity of the precipitation data using data from the station at Cascadia State Park and a double-mass curve analysis. We also computed annual ratios of the precipitation sums at the two stations. Both analyses provided no evidence for any non-homogeneity. Testing the modelling approach for the control catchment can also be seen as a test for homogeneity. If there had been inhomogeneities in the time series for precipitation and temperature, this would have resulted in false detections of runoff changes for the control catchment.

CONCLUDING REMARKS

The effect of forest harvesting and road construction on peak flow responses in streams continues to pose important questions despite decades of paired catchment studies and hydrological research by various groups around the world. The change-detection modelling approach described in this paper provides a useful alternative to the headwater-scale paired catchment approach to evaluate the hydrological effects of a LULC change, particularly where a suitable control catchment does not exist. The lack of a suitable control is often the case for larger catchments. Whereas paired catchment studies rely on consistent runoff measurements in two catchments (a control and a treatment), the modelling approach described in this paper requires runoff data from only one catchment. However, data consistency is crucial, not only for runoff, but also for the driving variables (in this study precipitation and temperature). In the approach presented herein, the model is not used to make predictions of changes but rather for runoff reconstruction (analysis of model residuals) or to characterize and compare the runoff dynamics (analyses of model parameters calibrated for different periods and model simulations with those parameters). Although this modelling effort can be undertaken only if the necessary meteorological and streamflow data are available, it has fewer requirements than when using a model to predict runoff associated with different treatment scenarios. In the latter case, it must be determined whether or not the chosen model is able to make reliable scenario predictions, although this aspect often is not tested thoroughly. If properly tested, a model that provides scenario predictions is a powerful tool for catchment management. Using a change detection approach as described in this paper might provide results that can contribute to models capable of forward predictions of LULC change scenarios.

Acknowledgements We thank Julia Jones for early discussions and Gordon Grant, Fred Swanson, Stefan Uhlenbrook and one anonymous reviewer for valuable comments on an earlier version of the manuscript. We also thank Jan-Olov Persson, Department of Mathematics, Stockholm University, for fruitful discussions on the statistical analyses. This work was supported by USFS grant PNW 03-CA-11261921-529 and NSF EAR DEB0218088. The Swedish STINT Foundation and TU Delft are also thanked for their support of JJM during the final writing stages.

REFERENCES

Andréassian, V. (2004) Waters and forests: from historical controversy to scientific debate. *J. Hydrol*. **291**(1-2), 1–27.

Andréassian, V., Parent, E. & Michel, C. (2003) A distribution-free test to detect gradual changes in watershed behavior. *Water Resour. Res*. **39**(9), 1252, doi:1210.1029/2003WR002081.

Bergström, S. (1976) Development and application of a conceptual runoff model for Scandinavian catchments. SMHI Reports RHO no. 7, Norrköping, Sweden.

Bergström, S. (1992). The HBV model - its structure and applications. SMHI Reports RH no. 4, Norrköping, Sweden.

Berris, S. N. & Harr, R. D. (1987) Comparative snow accumulation and melt during rainfall in forested and clear-cut plots in the western cascades of Oregon. *Water Resour. Res.* **23**, 135–142.

Beschta, R. L., Pyles, M. R., Skaugset, A. E. & Surfleet, C. G. (2000) Peakflow responses to forest practices in the western cascades of Oregon, USA. *J. Hydrol*. **233**(1-4), 102–120.

Beven, K. (2001) How far can we go in distributed hydrological modelling? *Hydrol. Earth Syst. Sci*. **5**(1), 1–12.

Beven, K. (2006) A manifesto for the equifinality thesis. *J. Hydrol*. **320**(1-2), 18–36.

Blöschl, G., Ardoin-Bardin, S., Bonell, M., Dorninger, M., Goodrich, D., Gutknecht, D., Matamoros, D., Merz, B., Shand, P. & Szolgay, J. (2007) At what scales do climate variability and land cover change impact on flooding and low flows? *Hydrol. Processes* **21**(9).

Bonell, M. & Bruijnzeel, L. A. (2005) *Forests Water and People in the Humid Tropics: Past, Present and Future Hydrological Research for Integrated Land and Water Management*. Cambridge University Press, Cambridge.

Bowling, L. C., Storck, P. & Lettenmaier, D. P. (2000) Hydrologic effects of logging in western Washington, United States. *Water Resour. Res*. **36**(11), 3223–3240.

Brandt, M., Bergstrom, S. & Gardelin, M. (1988) Modeling the effects of clearcutting on runoff – examples from Central Sweden. *Ambio* **17**(5), 307–313.

Bruijnzeel, L. A. (2004) Hydrological functions of tropical forests: not seeing the soil for the trees? *Agri. Ecosyst. Environ*. **104**(1), 185–228.

Daly, C., Neilson, R. P. & Phillips, D. L. (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Met.* **33**(2), 140–158.

Eisenbies, M. H., Aust, W. M., Burger, J. A. & Adams, M. B. (2007) Forest operations, extreme flooding events, and considerations for hydrologic modeling in the Appalachians—A review. *For. Ecol. Manag.* **242**(2-3), 77–98.

Grant, G. E. & Wolff, A. L. (1991) Long-term patterns of sediment transport after timber harvest, western Cascade Mountains, Oregon, USA. In: *Sediment and Stream Water Quality in a Changing Environment: Trends and Explanation*. IAHS, 203.

Hamilton, L. S. & King, P. N. (1983) *Tropical Forested Watersheds: Hydrologic and Soils Response to Major Uses or Conversions*, Westview Press, Boulder, Colorado, USA.

Harr, R. D. (1977) Water flux in soil and subsoil on a steep forested slope. *J. Hydrol*. **33**, 37–58.

Harr, R. D. (1986) Effects of clearcutting on rain-on-snow runoff in western Oregon: a new look at old studies. *Water Resour. Res*. **22**(7), 1095–1100.

Harr, R. D. & McCorison, F. M. (1979) Initial effects of clearcut logging on size and timing of peak flows in a small watershed in western Oregon. *Water Resour. Res*. **15**(1), 90–94.

Hewlett, J. D. (1982) Forests and floods in the light of recent investigation. In: *Hydrological Processes of Forested Areas*, 543–559, National Research Council of Canada Publ. 20548, Ottawa.

Jones, J. A. (2000) Hydrologic processes and peak discharge response to forest removal, regrowth, and roads in 10 small experimental basins, western Cascades, Oregon. *Water Resour. Res*. **36**(9), 2621–2642.

Jones, J. A. & Grant, G. E. (1996) Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon. *Water Resour. Res*. **32**(4), 959–974.

Jones, J. A. & Grant, G. E. (2001) Comment on "Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon" by J. A. Jones and G. E. Grant. *Water Resour. Res*. **37**(1), 179–180.

Kuczera, G. (1987), Prediction of water yield reductions following a bushfire in ash-mixed species eucalypt forest. *J. Hydrol*. **94**(3-4), 215–236.

Kundzewicz, Z. W. & Robson, A. J. (2004) Change detection in hydrological records – a review of the methodology. *Hydrol. Sci. J.* **49**(1), 7–19.

Lindström, G. & Bergström, S. (1992) Improving the HBV and PULSE-models by use of temperature anomalies. *Vannet i Norden* **1**, 16–23.

Lindström, G., Johansson, B., Persson, M., Gardelin, M. & Bergström, S. (1997) Development and test of the distributed HBV-96 hydrological model. *J. Hydrol*. **201**(1-4), 272–288.

Lookingbill, T. R. & Urban, D. L. (2003) Spatial estimation of air temperature differences for landscape-scale studies in montane environments. *Agric. For. Met*. **114**(3-4), 141–151.

McGuire, K., Weiler, M. & McDonnell, J. (2007), Integrating tracer experiments with modeling to assess runoff processes and water transit times. *Adv. Water Resour.* **30**(4), 824–837.

McGuire, K. J., McDonnell, J. J., Weiler, M., Kendall, C., McGlynn, B. L., Welker, J. M. & Seibert, J. (2005) The role of topography on catchment-scale water residence time. *Water Resour. Res*. **41**, W05002, doi:05010.01029/02004WR003657.

Moore, R. D. & Scott, D. F. (2005) Camp Creek revisited: streamflow changes following salvage harvesting in a medium-sized, snowmelt-dominated catchment. *Can. Water Resour. J.* **30**(4), 331–344.

Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models, Part I – A discussion of principles. *J. Hydrol*. **10**(3), 282–290.

Rothacher, J. (1965) Streamflow from small watersheds on the western slope of the Cascade Range of Oregon. *Water Resour. Res*. **1**(1), 125–134.

Rothacher, J., Dyrness, C. T. & Fredriksen, R. L. (1967) Hydrologic and related characteristics of three small watersheds in the Oregon Cascades. Rep., US Dept. Agric. Forest Service, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, USA.

Schnorbus, M. & Alila, Y. (2004) Forest harvesting impacts on the peak flow regime in the Columbia Mountains of southeastern British Columbia: An investigation using long-term numerical modeling. *Water Resour. Res*. **40**(5), W05205, doi:05210.01029/02003WR002918.

Seibert, J. (1997) Estimation of parameter uncertainty in the HBV model. *Nordic Hydrol.* **28**(4-5), 247–262.

Siriwardena, L., Finlayson, B. L. & McMahon, T. A. (2006) The impact of land use change on catchment hydrology in large catchments: the Comet River, Central Queensland, Australia. *J. Hydrol*. **326**(1-4), 199–214.

Sollins, P., Grier, C. C., McCorison, F. M., Cromack, K. J., Fogel, R. & Fredriksen, R. L. (1980) The internal element cycles of an old-growth Douglas-fir ecosystem in western Oregon. *Ecol. Monogr.* **50**, 261–285.

Tague, C. & Band, L. (2001) Simulating the impact of road construction and forest harvesting on hydrologic response. *Earth Surf. Processes Landf.* **26**(2), 135–151.

Thomas, R. B. & Megahan, W. F. (1998) Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon: a second opinion. *Water Resour. Res*. **34**(12), 3393–3403.

Thomas, R. B. & Megahan, W. F. (2001) Comment on "Peak flow responses to clear-cutting and roads in small and large basins, western Cascades, Oregon" by J. A. Jones & G. E. Grant – Reply. *Water Resour. Res*. **37**(1), 181–183.

Troendle, C. A., Wilcox, M. S., Bevenger, G. S. & Porth, L. S. (2001) The Coon Creek Water Yield Augmentation Project: implementation of timber harvesting technology to increase streamflow. *For. Ecol. Manag.* **143**(1-3), 179–187.

Uhlenbrook, S., McDonnell, J. & Leibundgut, C. (2003) Runoff generation and implications for river basin modelling special issue. *Hydrol. Processes* **17**(2).

Uhlenbrook, S., Seibert, J., Leibundgut, C. & Rodhe, A. (1999) Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrol. Sci. J.* **44**(5), 779–797.

Waichler, S. R., Wemple, B. C. & Wigmosta, M. S. (2005) Simulation of water balance and forest treatment effects at the H.J. Andrews Experimental Forest. *Hydrol. Processes* **19**(16), 3177–3199.

Wei, X., Sun, G., Liu, S., Jiang, H., Zhou, G. & Dai, L. (2008) The forest–streamflow relationship in China: a 40-year retrospect. *J. Am. Water Resour. Assoc.* **44**(5), 1076–1085.

APPENDIX

A short description of the HBV model

The HBV (*Hydrologiska byråns vattenavdelning*) model is a conceptual runoff model that can be used to simulate daily catchment runoff based on observed time series of daily rainfall, temperature, and potential evaporation. Precipitation is considered to be either snow or rain depending on whether the temperature is above or below a threshold temperature, P_{TT} (°C). All precipitation simulated to be snow, *i.e.*, falling when the temperature is bellow P_{TT}, is multiplied by a snowfall correction factor, P_{SFCF} (-), which compensates for systematic errors in the snowfall measurements and evaporation from the snow pack in the model (the latter is not simulated explicitly). Snowmelt, M (mm d^{-1}) is calculated with the degree-day method using the degree-day factor P_{CFMAX} (mm d^{-1} °C^{-1}) (equation (A1)). Meltwater and rainfall is retained within the snow pack until it exceeds a certain fraction, P_{CWH} (-), of the water equivalent of the snow. When temperatures drop below P_{TT} the amount of refreezing liquid water within the snow pack, R (mm d^{-1}) is computed using a refreezing coefficient, P_{CFR} (-) (equation (A2)).

$$M = P_{CFMAX} \cdot (T(t) - P_{TT}) \tag{A1}$$

$$R = P_{CFR} \cdot P_{CFMAX} \cdot (P_{TT} - T(t)) \tag{A2}$$

Based on the amount of rainfall and snow melt at a certain day, $P(t)$ (mm d^{-1}), the flux to the groundwater, $F(t)$ (mm d^{-1}), is computed; the remaining part of $P(t)$ is added to the soil box. The partition is a function of the ratio between current water content of the soil box ($S_{SOIL}(t)$, mm) and its maximum value (P_{FC}, mm; equation (A3)). Actual evaporation from the soil box equals the potential evaporation if S/P_{FC} is above P_{LP} (-) time P_{FC}, while a linear reduction is used when S/P_{FC} is below this value (equation (A4)).

$$\frac{F(t)}{P(t)} = \left(\frac{S_{SOIL}(t)}{P_{FC}}\right)^{P_{BETA}} \tag{A3}$$

$$E_{act} = E_{pot} \cdot \min\left(\frac{S_{SOIL}(t)}{P_{FC} \cdot P_{LP}},\ 1\right) \tag{A4}$$

Groundwater recharge is added to the upper groundwater box (S_{UZ}, mm). P_{PERC} (mm d^{-1}) defines the maximum percolation rate from the upper to the lower groundwater box (S_{LZ}, mm). Runoff from the groundwater boxes is computed as the sum of two or three linear outflow equations (P_{K0}, P_{K1} and P_{K2}, d^{-1}) depending on whether S_{UZ} is above a threshold value, P_{UZL} (mm), or not (equation (A5)). This runoff is finally transformed by a triangular weighting function defined by the parameter P_{MAXBAS} (equation (A6)) to give the simulated runoff (mm d^{-1}).

$$Q_{GW}(t) = P_{K2} \cdot S_{LZ} + P_{K1} \cdot S_{UZ} + P_{K0} \cdot \max(S_{UZ} - P_{UZL}, 0) \tag{A5}$$

$$Q_{sim}(t) = \sum_{i=1}^{P_{MAXBAS}} c(i) \cdot Q_{GW}(t - i + 1)$$

$$\text{where } c(i) = \int_{i-1}^{i} \frac{2}{P_{MAXBAS}} - \left|u - \frac{P_{MAXBAS}}{2}\right| \cdot \frac{4}{P_{MAXBAS^2}}\, du \tag{A6}$$

The long-term mean values of the potential evaporation, $E_{pot,M}$, for a certain day of the year are corrected to its value at day t, $E_{pot}(t)$, by using the deviations of the temperature, $T(t)$ at a certain day, from its long-term mean, T_M, and a correction factor, P_{CET} (°C^{-1}) (equation (A7)).

$$E_{POT}(t) = (1 + P_{CET} \cdot (T(t) - T_M)) \cdot E_{POT,M} \tag{A7}$$

$$\text{but } 0 \leq E_{POT}(t) \leq 2 \cdot E_{POT,M}$$