Dear Editor,

Hereby we submit the revised version of our manuscript entitled "Climate change, re-/afforestation, and urbanisation impacts on evapotranspiration and streamflow in Europe". We feel we have been served well in the review process with constructive and detailed comments provided by 4 anonymous reviewers, and an additional review from one of our own students from Wageningen. We take the large number of reviewers as a sign that the work has sparked interest. We are also pleased to see that no fundamental flaws were identified during the review process. All reviewers seem to agree that the manuscript is interesting, well written, and is in principle publishable. In response to the main criticism about the potential evapotranspiration used, we have gone through considerable efforts to redo all simulations with CRU PET. As a result, all figures have been updated and the numbers have changed slightly, but it should be noted the conclusions remain unchanged.

We hope that the rebuttal will be positively received by the reviewers.

Best regards, on behalf of the authors,

Ryan Teuling

Bennekom, 7-6-2019

Anonymous Referee #1

Since the forest cover effect is hardcoded in the Budyko model, it will simulate changes in ET. However, this remains an extrapolation, which needs a better validation than what is now presented. The authors mention that the modelled ET average agrees well with the patters of GLEAM. Here I would like to ask the the authors to report statistics of this comparison. Then they also report a correlation with streamflow changes of r = 0.34, which corresponds to an explained variance of 12%, leaving 88% unexplained! Please show the scatterplot. Since there is a need for a validation of the model, I think that the model should be able to predict the observed streamflow changes better than a reference, for example using the changes in precipitation and maybe PET. Only if the Budyko model shows a higher skill I see justification to use that model and its change of the landuse parameterisation for the whole of Europe with confidence.

We agree with the importance of validation as stressed by the referee. In response, we have added a 4*4 contingency table (which was suggested by Anonymous Referee #2 and has the function of a scatterplot) and a comparison of the trends in simulated Q, P and PET over the basins of Stahl et al. (2012). From this analysis (shown in a second new figure) we find that our simulation is closest to the observed change. Both P and PET have larger (wet) bias than the model. In order to compare trends over different units we considered the median change (over all basins) normalized by the IQR (Figure 7). We did not include a quantitative comparison with GLEAM because it is not our goal to match this product. In particular, for forest and urban areas (the focus of our study), we believe our model to be more accurate than GLEAM because it has been constrained by long-term water balance observations rather than being validated on eddy covariance data (as is the case with GLEAM). The comparison is for reference only. We have now made this more clear in the text: "Model simulations are validated and compared …" and "It should be noted that this comparison is added for reference only and should not be seen as validation: GLEAM is not a strictly observational dataset, and it does not necessarily provide better long-term estimates of ET for forest and urban areas."

Land-use change is modelled by changing the parameter in the Budyko model using data from lysimeters. This is quite a central methodological step and ignores differences in scale of a lysimeter with that of a heterogeneous landscape. It also ignores that the parameter in the Budyko model can be different due to climatic variation, in particular the seasonality of rainfall to that of evaporative demand and the rainfall frequency. Jaramillo et al., 2018 HESS showed that there are increases in evaporative fraction, not explained by climate for many catchments in Sweden. Yet the link to changes in forest properties was rather weak. In contrast this study prescribes a distinct effect of forest age, hence there is a strong tendency that this study assesses the upper range of changes in water balance (if the HILDA database actually reflects the changes).

We disagree with the reference to Jaramillo et al. as used here. They find that: "… the positive residual effect occurred along with increasing standing forest biomass in the temperate and boreal basin groups, increasing forest cover in the temperate basin group and no apparent changes in forest species composition in any group". This is fully consistent with our approach of accounting for a) changes in forest cover at the 1 km2 resolution (much smaller than the basins considered in the Jaramillo study) and b) the use of different Budyko parameters for increasing stand age (biomass) as calibrated from unique long-term lysimeter experiments. Hence we believe our study is not at odds with Jamamillo et al. In the revised version, we now refer to this study and also mention seasonality and the synchronicity between precipitation and potential evapotranspiration explicitly as factors that might potentially affect w* in the Discussion (where other factors were already mentioned). Indeed, there is a scale mismatch between the size of the lysimeters and the resolution of the HILDA pixels, but we don't have any reason to assume that the larger forested lysimeters of 625 m2 do not represent the surrounding landscape (why would they have been constructed, operated and maintained for 50+ years if they would not be representative of the larger region?). We now mention the scale of the lysimeters more explicitly in the text ("area varies from 1 to 625~m^2 for the larger lysimeters at Castricum")

The choice of Thornthwaite method for PET is not acceptable for various reasons: a) It underestimates the evaporative demand (PET or Rn/L, see also van der Schrier 2011 or Maes et al., 2018). An annual average of PET of 700mm/yr for Southern Europe is far to low. That is why the authors need to scale it by an arbitrary factor aPET in the Budyko curve. b) Since it is a function of temperature only, it will be overly sensitive to warming trends which is arguably pretty strong for the considered period. It also misses changes in shortwave solar radiation, see e.g. Wild et al., 2007. c) The authors argue for Thornthwaite because of data availability. However, there is data on sunshine duration / cloud cover. Furthermore, the diurnal temperature range correlates with solar radiation and has been used as a proxy for this, e.g. Wild et al., 2007, Makowski et al. 2008.

We agree, and have redone the whole analysis using CRU PET as was suggested in the follow-up comment by this reviewer. This has however not significantly changed the results, nor the fact that scaling is needed to account for annual actual ET rates (including interception evaporation) that exceed PET.

Apart from these major issues I enjoyed reading the paper. It is very well written, is well structured and has appealing figures. The topic is of high relevance for HESS. However, I believe that the validity of the Budyko approach needs to be demonstrated and therefore I recommend major revisions.

Thank you for the kind words. With the simulations now based on CRU PET and additional validation we hope that we can convince the reviewer of the validity of our study.

Minor Remarks:

Introduction, L20ff: it is argued that there are no sufficient studies which treat both landuse change and climate change on streamflow / ET. However, there are studies which indeed try to accomplish this, which I want to bring to the attention of the authors. For example Jaramillo et al., 2018 assessed changes in multiple catchments in Sweden. Renner et al., 2014 assessed observed changes of streamflow in East Germany. Lopez-Moreno et al., 2011 for catchments in Spain.

We now cite the studies by Jaramillo and Renner. It was not clear which study by Lopez-Moreno et al. the reviewer was referring to as no reference was provided. However it should be noted that neither Jaramillo nor Renner discuss or even mention impacts on streamflow (only ET).

Figure 3: color of missing values (NA) should not be white, as indicated in the legend

Fixed

Figures 6,7: there should be a color legend, a 3D color scheme on a map is a beautiful drawing but really difficult to grasp. What is the meaning of grey here? Similar magnitude of all drivers or a missing value? To what reference are the data scaled 2-98%, all of Europe? The choice of rectangular sub-regions seems arbitrary to me. Why not use relevant river basins, where data is available to see if your prediction is indeed pointing in the right direction. For example on P9L10 it is mentioned that Scotland shows dramatic increases in streamflow, is this finding supported by observed changes?

A legend for a 3D colormap would also be 3D (a cube) so not easy to display. We believe this would only add to possible confusion rather than make the interpretation easier. Instead, we have opted for zooming in on (rectangular) regions which show strong changes of different composition, and show the individual contributions and corresponding colors so that the interpolation can be based on those. Since none of the other 4 reviewers had the same comment, we decided to keep the figure as it is (note that this figure was iterated and discussed extensively among the authors). Concerning the Scotland example mentioned by the referee, it should be noted that many of the patterns in the figure result from E-OBS rainfall trends which are known to be robust since they incorporate most available raingauge data. In the case of Scotland, the increasing precipitation is well known (see e.g. https://www.adaptationscotland.org.uk/why-adapt/climate-trends-and-projections). Given that this region is generally energy-limited, it should not come as a surprise that the increasing P acts to increase Q rather than ET. Similar arguments hold for the other examples. To better explain the scaling we changed the sentence into “Each contribution is inversely scaled between the 2nd and 98th percentiles over Europe to reflect its relative importance”.

Table 3: The units in the caption should be km^3/yr and not km/yr. In any case I would prefer fluxes per unit area to allow comparison. Further I think that the total changes in Q / ET should be reported, not just the contributions.

Caption is corrected. We added the total area of the region so that number can easily be converted to per unit area. We did not report the total changes in this table since our aim was to compare contributions.

Fleur Verwaal

The first major comment is about the Thornthwaite method mentioned on page 5 line 1. It is used for calculating PET and requires only the temperature as input (Thornthwaite, 1948). Since there is a warming trend since the 1950s, this choice of method is questionable. Multiple studies, such as Trajkovic and Kolakovic (2009), have found that the radiation-based methods more accurate reproduce reference PET than temperature based methods. Fisher et al (2011) mention that temperature based models estimated 20–30% less than the radiation based models averaged across all their researched sites. It was even stated: "The choice of evapotranspiration model and input data is likely to have a bearing on model fits and predictions when used in analyses of species richness and related phenomena at geographical scales of analysis" (Fisher et al., 2011). Shaw and Riha (2011) state that the Priestley–Taylor equation (a primarily radiation-based model) consistently explained more of the variation in PET than temperature-based methods. The paper of Teuling et al (2019) acknowledges that Thornthwaite method does not always gives the strongest increase in PET values in a warming climate. Considering that the paper aims to understand the effect of climate change on green and blue water fluxes, the effect of a warming trend on the calculated PET values should not be overlooked. The temperature-based PET values will affect the main part of the paper, since it in used in the Budyko model to determine how the average precipitation is portioned between evapotranspiration and streamflow. To improve the quality of the paper, please switch to a radiation based model or add substantial argumentation, on why they picked the Thornthwaite method to calculate the PET over radiation-based methods.

We agree with this comment, which was also brought forward by other reviewers. We have now redone our analysis using CRU PET.

The second major comment is about the observations in this study, which come from lysimeter stations according to P5 Line 23. These lysimeters are assumed to behave similar to landscape elements of 10e6 m2. The locations of these stations, are not evenly spread throughout Europe but mainly constrained central-west, as can be seen in Fig A1. The model forcing is based on interpolated observation from weather stations. The paper states that local land cover impacts on climate, such as enhanced temperature or cloud formation, should not be represented in the forcing dataset. The stations should indeed be carefully selected. WMO (2003) states "Observations of evapotranspiration should be representative of the plant cover and moisture conditions of the general surroundings of the station". Still, the interpolation of lysimeter stations should be representative for the whole of Europe, can this be achieved if the stations are only concentrated in the central-west? It can result in incorrect values near the edges of the maps of Fig 2-7. Please expand the amount and the spread of lysimeter stations or otherwise show the statistics to support the used method.

Apparently the reviewer got the wrong impression that we interpolate lysimeter data. This is incorrect. We optimize a model with lysimeter data that is subsequently for with gridded (interpolated) meteorological observations and used for the spatial prediction of water balance partitioning. We of course agree with the fact that the distribution of the stations is far from ideal, however this is clearly beyond our control. Long-term lysimeter data are rare (most stations are used for experiments rather than climate monitoring). We are among the first to synthesize data from most of the stations into a single modelling framework constrained as much as possible by high-quality observations of water balance partitioning. We believe this approach, which is different to most other studies, adds value to the existing literature on trends that are mostly model-based. It should also be noted that there is no evidence that our model would not work well in other regions, since we are in agreement with many other local studies (like Jaramillo et al. in Sweden, of course in more water limited regions our model would simply be constrained by the water limit which should not be a point of debate when irrigation is not considered). It is assumed that the Budyko framework (just like any other model calibrated on network data such as FLUXNET that does not have a uniform distribution over all climate zones) can be used to transfer local findings to other climate conditions (it should be noted that the Budyko framework in particular was designed to do just that). This is a central assumption in all Budyko studies, and it has long been known that this is a reasonable assumption.

The third major comment is about the temporal scale. In the method section on page 7 (line 7) it was stated that changes over the intermediate 10-year periods (1955–1965 and 2005–2015) were analysed. It was stated "the trends were generally found to be monotonic". Therefore they calculated 10-year climate averages. These were used to force the Budyko model and calculate changes in evapotranspiration and water yield, so it influences the main part of the paper. The simulated continental scale patterns depend on these 10-year climate averages. The choice of words on line 9: "the trends were generally found to be monotonic" raises questions. What were the exceptions? Did this choice of temporal scope have significant effect on the calculated changes in evapotranspiration and water yield? As Zang et al (2004) states, the climatic variables precipitation, temperature, solar radiation and humidity have a large spatial en temporal variability. They interact with the catchment characteristics such as vegetation cover, which is of interest for Teuling et al. Therefore, please choose a smaller temporal scale in which the trends are all found to be monotonic or show the statistics of the trends over the 10-year periods to verify the choice to average them.

The 10 year period was chosen because land use changes in HILDA are reported at this resolution, and not finer. We did not find any major deviations in the general trend in the intermediate periods, so, therefore, we decided to focus only on the oldest and most recent periods covered by all datasets. However, in response to the reviewer (and to reviewer #2 who had a similar remark) we show the maps with 10-year advancing changes in P and PET in this rebuttal. We are happy to include a figure in the main manuscript if the reviewers consider it necessary.

The fourth major comment is about correlation mentioned on page 8 line 9. The paper mentions that their approach is able to reproduce the overall pattern of observed changes in streamflow. It was stated "In spite of the difference in units and the fact that individual basins might have shorter record lengths, the correlation in trends between the basins is 0.34." However, a correlation of 0.34 leaves room for questions, is this correlation sufficient? It means that a large part of the data remains unexplained. The paper states that, the validation shows that their simplified approach is able to capture continental-scale patterns in mean and changes of blue and green water fluxes. Can the correlation of the pattern of observed changes in streamflow be improved by adjusting the input, such as the PET values calculated with a radiation-based model (Considering my first comment)? Please change in input to optimize the correlation or show more elaborate statistics and argumentation on why this correlation is sufficient.

In response to this comment and similar comments by other reviewers, we have extended the validation and included 2 new figures.

Minor Arguments

P1 Line 5: Please replace the term 'green and blue water fluxes' with evapotranspiration and streamflow, to make it understandable without having to read the introduction.

Done

P8 Line 24: simulated ET is shown in figure 5b while it is referenced to 5a

Corrected

P8 Line 25: Observed ET is shown in figure 5a while it is referenced to 5b

Corrected

P9 Line 25: Table 3 list the Europe-wide changes not table 2

This should be Table 2, but since Table 1 needed to be split over 2 pages it received both nr 1 and 2.

P12 Line 1: Change 'Therefor' into Therefore

Done

P11 Line 3: 'WMO recommendations' please include a reference

We included a reference to WMO (1993)

P21 Fig 1: needs revising and clarification:

a. The caption does not fully describe what is displayed in the figure. Please elaborate on the w* values.

Caption has been extended.

b. Yellow line is hardly noticeable, consider changing it to another colour to improve readability

We added the line for reference since the Castricum lysimeter station provides data for bare soil. However this is not used in the analysis since HILDA doesn't have a bare soil class.

c. The legend on the left indicates the colours orange, light and dark green. However, it does not include red and yellow, what do those colours indicate?

The figure mixes land use and stand age effects on w*, hence we use different symbols and colors. To avoid confusion we have added colors for the urban (red) and bare soil (yellow) classes.

d. In the end of the results, it was mentioned that the colours indicate the forest stand age, this should also be mentioned in the caption

Corrected

e. The caption should include describing the grey dashed line as energy limit, to improve the understandability.

Added.

P23 Fig 3: needs revising and clarification: a. The missing values (NA) are indicated by the colour white, however white is already used to indicate another fraction. This brings confusion what the colour is indeed indicating. Please indicate the missing values with another colour.

Figure has been revised.

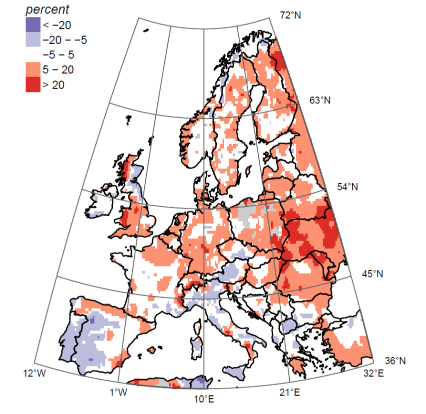
P24 Fig 4: needs revising and clarification: a. Fig 4b and 4d indicate the change for evapotranspiration and streamflow. The change is indicated with green and blue colours to match the evapotranspiration (green) and streamflow (blue). They mention in the caption that they chose to reverse the colour scheme on purpose. However, it works confusing and counterintuitive. My recommendation is to choose a different colour scheme's to match the change in both the figures, without green and blue, to avoid confusion.

We removed reference to blue and green water, yet since we believe it can help the interpretation if colours indicate change in the same direction we have kept the colorbars the same.

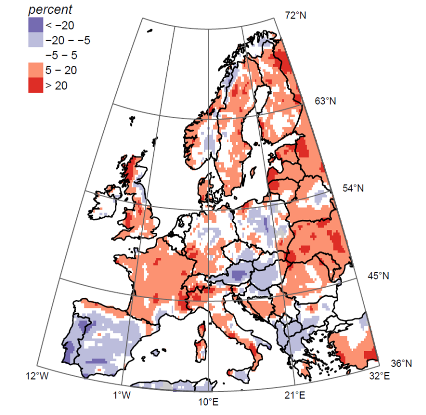
P25 Fig 5: needs revising and clarification: a. In P8L24 and P8L25 there are references to figure 5, I mention below that they reference to the wrong part of the figure. However, one can also consider keeping the reference in that way, and change the order in the figure. In 5a and 5c the observation based ET are shown and in 5b and 5d the simulated ET is shown. When the simulated ET figures are switched to the left, it will fit more clearly in the story line.

Good suggestion, we changed the panel from left to right and vice versa.

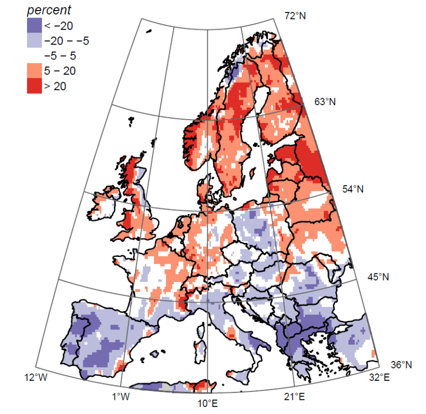
P and PET changes in 10-year periods



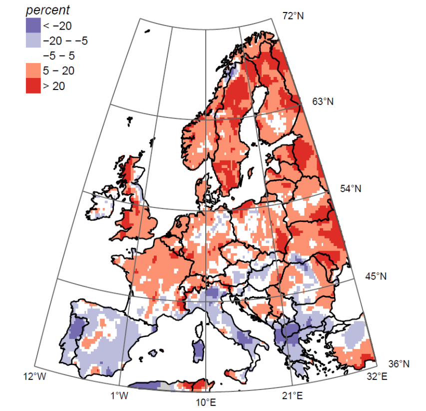
P change 1960-1970



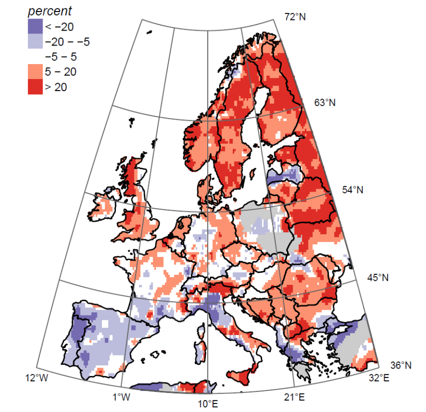
P change 1960-1980



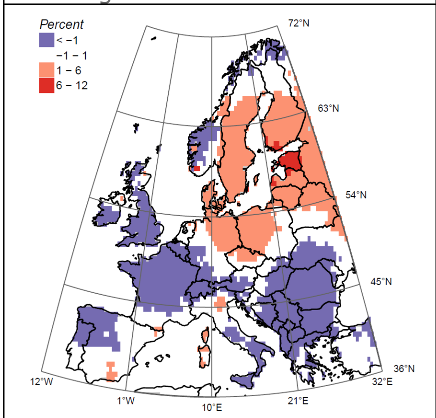
P change 1960-1990



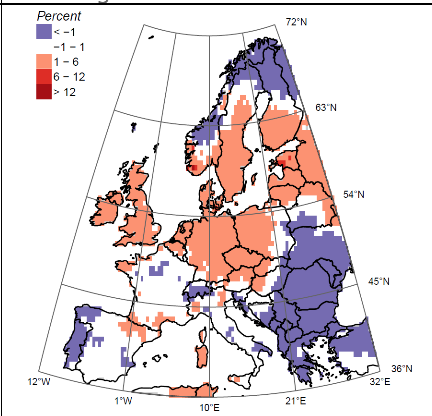
P change 1960-2000



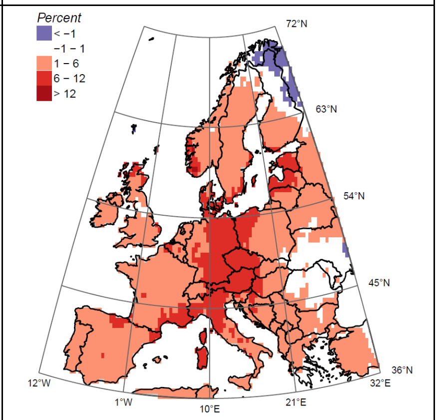
P change 1960-2010



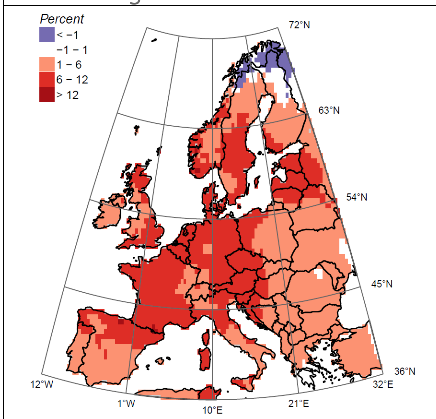
PET change 1960-1970



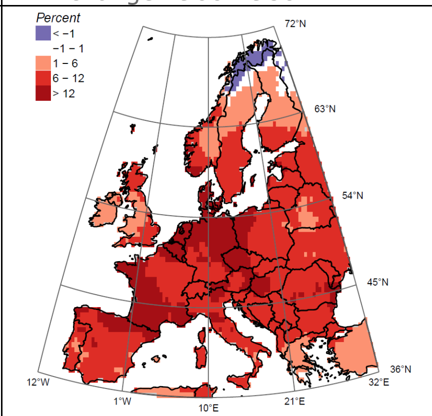
PET change 1960-1980



PET change 1960-1990



PET change 1960-2000



PET change 1960-2010

Anonymous Referee #2

1. The revisited land-use dependent Budyko curves

The land-use attribution relies heavily on the Budyko curves depicted on Figure 1. First I did not understand why this Figure is not discussed in the results section. The authors did a great job in collecting these lysimeter data but the amount of data remains too limited to design the whole modeling framework. Some curves are adjusted on the basis of very few points, e.g. w* is calibrated on the basis of two points for urban areas and these two points are extracted from a unique site of Arnhem. How can we state that this parameter will be representative of all urban areas in Europe? Some land use classes present more experimental points but the w* fitting is far from being satisfying, with large uncertainties, no clear distinct w* values for some classes and again many points are related to the same environmental data (the 26 points originated from only four sites, Table 1). Given the multiple sources of uncertainties, the authors should consider to quantify the parametric uncertainty (the sensitivity of the results to w* fitted values) and should try to validate the fitted w* on independent data (e.g. the streamflow data, see next comment).

As with any study, we are limited by data availability. We did our best to collect as many high-quality observations as possible that reflect ET from a single land use class (note that we excluded catchment-scale estimates because nearly all catchments have a mixed land cover); however, very few long-term datasets are available for urban environments. While we can’t exclude the possibility that the optimum Budyko parameter for urban environments will change slightly when more data become available, we believe our current estimate based on 3 long-term campaigns over Rotterdam, Arnhem and Basel (and not just Arnhem as mentioned by the referee) to be fairly robust, and we also believe that the Budyko framework is a good first order approximation for extrapolation to other climate regions. Note that we managed to find an additional data point for urban areas with respect to the initial version (Basel), which is consistent with the data from Arnhem and Rotterdam.

2. Validation of the attribution results

The authors propose a validation exercise using GLEAM product and streamflow from near-natural catchments. It should be stated that the comparison to GLEAM cannot be viewed as a strict validation since GLEAM relies on hydrological modelling (different to Budyko but still a model using P and PET inputs). The validation using streamflow time series is to my opinion the unique way to perform a real validation with independent data. To perform a rigorous validation studies, the authors could compute for each catchment the observed “Net” change and compare it with the Net change computed by the Budyko-curves. The authors have the material to perform such validation that will provide a clear diagnosis on the method used for attribution. At this stage, the attribution exercise is more a sensitivity analysis, which is not at the level of the ambitious objectives of the study.

We agree such validation should be preferred. However, the streamflow data in Stahl et al. (2011) is not freely available for analysis, so we had to work with the numbers shown in their figures that were kindly provided to us by the main author. So, unfortunately, we do not have the data to perform the analysis suggested by the reviewer. In response to other comments, we have added 2 additional figures on the streamflow validation. We disagree that our exercise is just a sensitivity analysis, since the strength of our approach is that our model is directly constrained by hundreds of years of unique lysimeter and other ET observations that have never been combined before. With respect to the comparison with GLEAM, we added the following sentence to make more clear that this is not a validation: “It should be noted that this comparison is added for reference only and should not be seen as validation: GLEAM is not a strictly observational dataset, and it does not necessarily provide better long-term estimates of ET for forest and urban areas.”

3. Other comments

p.2 l.13-17: this is a repetition with previous sentences.

Sentence removed in line with other comments on the use of green and blue water. This terminology has been removed.

p.3 l.28-29: I disagree with this statement. The impact of urbanization is probably the most sensitive land use change impact on hydrological processes and it is discussed in the early hydrological literature (Leopold, 1968). See also the large sample studies by DeWalle et al. (2000) and the recent reviews on this topic (Oudin et al., 2018; Salvadore et al., 2015).

We don't argue against the fact the urbanization has a strong impact. This is in fact the reason we explicitly accounted for urban effects, and choose to name it in the title. We respectfully disagree with the referee that urbanization has a *stronger* effect than afforestation (which has been studied for a much longer time starting with experiments at Wagon Wheel Gap around 1910). The mentioned studies don't compare the two effects. In fact, we believe our figure 1 to be one of the first figures that allow a direct comparison of urbanisation and deforestation effects from observations. From the differences in w values, it seems that deforestation of a mature needleleaf forest has a slightly bigger impact than urbanization with respect to a grassland reference.

p.4 A discussion on the attribution problem is missing. There is a large existing literature on attribution studies in hydrology and I suggest that the proposed methodology be described upon the several existing attribution studies and associated methodology (see the reviews by Dey and Mishra, 2017 and Wang, 2014).

This is a useful suggestion. We included a some more information on the attribution approaches in the Introduction with reference to the papers mentioned by Dey and Mishra, 2017 and Wang, 2014.

p.6 l.1-5 Is it verified by local measurements of PET? I do not understand how the "c" linear factor might accounts for "all processes affecting yearly ET for tall vegetation".

Sentence changed into "so that ET generally will not exceed aPET even for forested areas" for clarity.

p.7 6-10: Using 10- yr periods to assess hydrological changes is too small with regard to natural climate variability.

We agree that changes over 10-year periods are too short to attribute observed changes to climate change rather than climate variability. However, we aimed to look at both land use and climate, and had to make a trade-off in the temporal and spatial scales used in the analysis. We decided to aim for the finest temporal and spatial scale possible given the resolution of datasets currently available. The rationale behind this strategy is that key land cover changes that impact the terrestrial water budget take place at (much) shorter timescales than 10 years. Therefore, we have used 10 year periods over which land use changes are typically reported and available for analysis. We have made clear in the text that whenever we refer to changes in climate that these can be due to long-term climate change, but also due to decadal variability. After confirming that the climate trends are rather uniform and constant over the analysis period (see figures above in Rebuttal), we decided to analyse the changes over the first and last 10-year period rather than averaging over two 30-year periods which would have smoothed much of both the land use and climate change signal. To clarify this, we added the following sentence: "While the 10-year periods are often considered short for climate change detection, they resulted from a need to balance robust estimation of the mean climate, without averaging out much of the underlying changes in both climate and land use."

p.8 l.28-29: Please clarify the differences in units and how the correlation trends is calculated. Besides, I am not sure that correlation is the more adequate tool, maybe a contingency table would be more appropriate to compare the observed and simulated trends.

We agree with the reviewer and have added a contingency table as a new figure to the validation.

p.10 l.24-25. Please modify the sentence and replace the term believe.

Done

Anonymous Referee #3

I see in the article some interesting aspects that contribute to the literature such as the constraining of the ET from the Budyko model by specific land use-dependent lysimetric data and a detailed analysis of land use changes across Europe, to calculate changes in ET and R. However, I found several weaknesses of the current version that need to be addressed.

1. You say you constrain the w parameter of Eq. 3 with observations of different land types. If I understand correctly, you constrain the w parameter in the locations where you have lysimetric measurements of ET and data on PET and P, and then apply that same w across all the spatial area of that specific land use/cover in combination with local PET and P data to obtain local ET rates across that land use/cover extension.

But the lysimetric observations as you mention, are located in a very small area of Europe. I think that extrapolating those w parameter values to regions like Northern Scandinavia and Iberian Peninsula and other Mediterranean areas is unrealistic. Can't you rely on the work by (Sterling et al., 2013) to improve that constraining exercise or other databases of ET rates? I also think that the land use categories used are to course and omit others such as open-water areas or reservoirs, etc.

It is a central assumption behind all Budyko analyses that catchments can be moved along the Budyko curve by changing climate conditions (see for instance reviews by Dey and Mishra, 2017 and Wang, 2014 mentioned by reviewer #2). The shape of the function is usually assumed to depend on land use and other local conditions. This is exactly the approach that we follow, and we think this approach is well established and fully consistent with other studies that for instance considered paired catchment data from a wider range of climate conditions (Zhang et al., 2001). We believe the strength of our approach is to only make use of observations that reflect hydrological dynamics at the scale at which land use changes occur and are monitored (1 km2). Databases such as the one by Sterling et al. contain larger basins with added uncertainty due to mixed land cover, unknown subsurface flow across catchment boundaries, and uncertain precipitation due to unsampled spatial variability. While we acknowledge the potential impact of limited data, it should be noted that no argument was provided by the reviewer as to why the Budyko approach would not be a valid first order approximation. Concerning the resolution of the HILDA dataset: we agree that improvements in resolution would be interesting. However, such datasets currently do not exist for the length of the study period considered here (i.e., starting long before the satellite era). We believe that in this respect, HILDA reflects the state-of-the-art. We assumed no changes in open water fraction in the analysis period based on the analysis by Fuchs et al. 2014 (Fig 8).

2. I know that the authors are aware of that (Page 10, line 30), but I see that there is no differentiation between irrigated and non-irrigated agriculture. Studies have found continental (Wang and Hejazi, 2011) and worldwide (Jaramillo and Destouni, 2015) driving effects from irrigation on long term ET and ET/P, from a Budyko perspective, and that are evident even at the large-basin scale. I think that a differentiation between irrigated and non-irrigated crops is compulsory for the constraining of the w parameter and the estimation of ET for land use/covers. In the same way, I would say that some of the attribution to reforestation in Southern Europe can be actually irrigation or rain fed agricultural intensification. Please check.

We fully agree with the reviewer that irrigation will have an effect on long-term ET and ET/P. However, it was not our goal to attribute observed changes to all possible causes, but rather to focus on the question how land cover changes compare to climate change impacts. It should be noted, however, that the impacts of reforestation in southern Europe are not related to irrigation but rather to the well-documented effects of land abandonment. This is not an artefact in the land use dataset. At the moment, the HILDA land use data does not distinguish between irrigated and non-irrigated agriculture. We know acknowledge the potential impact on the results in the Discussion, where we also refer to the studies mentioned by the reviewer.

3. Why are the authors using the blue/green water framework, if they are also combining the terminology with fluxes, etc. For instance, they use across the text the terms blue water, runoff, water yield (Page 2 line 6), which appear to be referring to the same. The manuscript needs to be consistent in this way and I would say that green and blue terminology is relevant only when water consumption is being assessed. If not in agreement, please justify the use of such terminology and also cite the main source for such (Falkenmark, 1997).

Other referees had similar comments, as a result we have removed this terminology from the manuscript.

4. It appears that an impact on long-term water partitioning from less now cannot be neglected that easily as stated in Page 6. See (Berghuijs et al., 2014) that also uses a Budyko approach. So at least acknowledge that uncertainty.

We now refer to Berghuijs et al. (2014) in the discussion.

5. The authors justify their work "In spite of the direct link between average green and blue water fluxes, few studies have addressed changes in both fluxes simultaneously. However, they omitted many works precisely doing that: (Rost et al., 2008; Siebert and Döll, 2010). I also find missing important works on the effects of forest change across Europe and from a Budyko framework perspective that have been omitted here (Jaramillo et al., 2018; Renner et al., 2014). These four studies would for sure enrich the discussion in relation to the attribution of the observed R and ET changes to forest change in Europe. Their findings should also support several of the statements expressed by the authors and interpreted from their results.

References to Jaramillo et al. (2018) and Renner et al. (2014) have been added. Since we have removed mention of green and blue water, we did not add references to Rost et al. (2008) and Siebert and Döll, (2010).

Anonymous Referee #4

This paper assesses the contributions of land use and climate changes to historical changes in streamflow and evapotranspiration in Europe. This is done using a stationary Budyko approach for water partitioning constrained by lysimeter observations and adopting historical land use reconstructions and gridded climate data at a high resolution of 1 km x 1 km. The resulting simulated changes in streamflow and evapotranspiration are in line with observed counterparts, although the comparison of streamflow changes is less straightforward. The contributions of land use change and climate change (through precipitation and evapotranspiration) are assessed for Europe and analysed in detail for eight selected regions.

Overall, the paper is well written and presents interesting results for the European continent. The authors use informative and well-prepared figures to illustrate their results. Several issues need attention such as the use of the terms green and blue water, the use of one (high) value to adapt the potential evapotranspiration and the aggregation of positive and negative contributions of land use or climate change at continental or large river basin scale. These and other specific comments can be found below. The paper includes many typos, several examples and other technical corrections can be found below as well.

We thank the reviewer for the kind remarks and detailed comments aimed at improving our work.

Specific comments

1. P2, L3-17: The terms green and blue water are not appropriately used here. The total evaporative flux also includes blue water from irrigation with surface water or groundwater (see e.g. Falkenmark, 2000; Oki and Kanae (2006); Falkenmark and Rockström, 2010) and hence the total evapotranspiration cannot be equated with the green water flux. Furthermore, the blue water flux does not only include streamflow (lines 15-16) but also groundwater flow (as briefly mentioned in line 5). The authors are suggested to remove the terms green water and blue water from the manuscript to avoid any confusion and to be consistent in terminology throughout the manuscript. There is no need to use the terms green and blue water (and water yield), since the focus is on streamflow and evapotranspiration.

We agree with the suggestion and have removed all reference to green and blue water.

2. P2, L11-12: Why is this in particular true for Europe? I would expect that uncertainties and limitations in observations and models in other parts of the world are at least comparable to those in Europe, but probably often larger.

This statement reflects on the impacts of climate change, not on uncertainties in modelling. These are indeed likely larger in other regions. Many impacts of climate change are relatively strong in Europe.

3. P4, L30: The readability of section 2 and also section 3 can be improved by distributing the contents of these sections over a few sub-sections.

Since this was not mentioned in any of the other 4 reviews, we decided not to use subsections. While on the lengthy side, the sections 2 and 3 are in fact shorter than the Introduction.

4. P5, L20-23: This sentence seems to be contradictory. The FLUXNET observations are not used in this study, because they are assumed to be more reliable for longterm water balance assessments. However, this study also considers long-term water balances. This should be better explained.

We have split the sentence for clarity.

5. P6, L10: A c-value of 1.8 is high and apparently seems to be used for all grid cells in Europe. This value implies that about 45% (0.8/1.8) of the energy used for evaporation is not included in the calculation of the potential evapotranspiration. Is this related to the simplified method (Thornthwaite) used to estimate the potential evapotranspiration? Which mechanisms (besides advection) are responsible for this? Is it reasonable to use the same c-value for all land use types? For instance, due to evaporation of intercepted water, you might expect higher c-values for forests compared to cropland and grassland. In summary, the use of a constant and

high value for correcting the potential evapotranspiration seems to be doubtful and partly limits the conclusions which can be drawn based on this study.

First, it should be noted that we no longer use Thornthwaite for the calculation of PET. The use of CRU PET and a slightly different scaling (we have allowed the Plynlimon observation to remain to the left of the energy limit even with aPET instead of PET) has led to a lower value of *c* (namely 1.6). It should be noted that we don't assume this value to be reasonable for all land use types, rather it indicates how much more the ET *could be if the land use was different* (optimal for ET). The factor is not a crop factor. The main reason with taking this approach rather than a more conventional crop factor is that the land use differences remain similar in the Budyko plot as to how they are in the observations (different crop factor would bring all the curves together), and that we keep most observations within the physical energy and water limits. The inconvenient truth is that annual ET for forested plots is often much higher than one would expect based on PET, but this is not a new finding in itself. We want to stress that the use of aPET does not affect the results in any way, because it is a scaling in model parameters that is balanced by an opposite scaling in forcing (because we use aPET rather than PET to drive the model). We have made this more clear in the manuscript by adding: "aPET should thus be interpreted as the maximum total yearly evapotranspiration that would occur under given climate conditions (PET/P) and under optimal vegetation conditions (i.e.\ vegetation that is most efficient in returning precipitation to the atmosphere, in this case needleleaf forest), rather than a land use-specific crop factor. It should be stressed that this scaling is only done to move most observations within the energy and water limits in the Budyko space (so that we can obtain a fit with Eq.~1), and that it has no impact on the results since the model is subsequently forced with aPET rather than PET."

6. P8, L5 and L14: The authors seem to mean something different with central-western Europe in these two lines, where firstly they seem to refer to Belgium and the Netherlands and secondly to Switzerland and Austria and parts of Germany and France. Try to be more specific here.

Done.

7. P8, L9-10: Is it logical that the mean evapotranspiration is highest due to pronounced topography? Although the term 'pronounced topography' is not completely clear, in general evaporation rates will decrease with altitude.

This is due to the impact of increased P dominating the slight decrease of PET with altitude. This sentence has been rephrased: "in regions with topographically-enhanced precipitation and/or forest cover".

8. P9, 25-27: Is it sensible to determine the net effect of for instance precipitation by balancing positive contributions from the north with negative contributions from the south? The net effect obscures the real contributions and potentially associated problems; however, these net effects are an important element of the main conclusions and the abstract. I would recommend the authors to reformulate relevant parts of the manuscript and highlight positive and negative contributions rather than net contributions.

Table 2 summarizes the results for Europe. A more regional-scale perspective is provided in (current) Figures 8 and 9, which list the regional contributions and net changes for several areas. However, we agree with the reviewer that more emphasis should be put on the positive and negative contributions at the regional scale. This is done by adding more numbers to the Results and Abstract, which has partly been rewritten and extended.

9. P10, L1-3: Is it useful to compare the sensitivity of streamflow to past land use changes with the effects of future changes? In order to interpret the differences between these two studies, the reader should at least have information on the magnitude of the future land use changes and the approach employed in the 2009 study. For instance, the way streamflow is determined in this study probably will be very different from the way streamflow has been determined in the 2009 study.

We agree with the reviewer that a comparison between the studies should be done with care. Our goal here was simply to provide a reference to other work in the Rhine Basin, hence we added "The strong sensitivity … seemingly contradicts"

10. P11, L13-25: What is the role of other variables than temperature and radiation in the determination of PET (i.e. humidity and wind) and what is the effect of excluding these variables on the results?

The analysis has been redone with CRU PET in response to the comments by reviewer #1

11. P11, L26-28: The statement that socio-economic impacts relate more directly to blue water fluxes compared to green water fluxes is not correct. Green water is the main source of water to produce food, feed, bioenergy, etc. (e.g. Oki and Kanae, 2006) and as such changes in green water availability and fluxes will have a large socio-economic impact.

This sentence was rephrased.

12. P25, Figure 5: Can streamflow be validated on the rate of change or only based on the patterns of change, since the units of c. and d. are different?

See also the response to other reviewers. If we would have had access to the raw data, we would have validated using the same units and the same time period (these slightly differ from basin to basin). However due to data restrictions we could only use the data as was shown in the Stahl et al. paper (which was kindly provided to us by the main author).

Technical corrections

1. P1, L12: 'Mediterranean' instead of 'Mediterrenean', see also e.g. page 2, line 22 and page 7, line 23.

Changed.

2. P1, L15: The meaning of ET is not clear here.

Changed.

3. P5, L12: ". . . the magnitude of this effect . . ."; which effect is meant here?

We changed "magnitude" into "this dependency"

4. P5, L17: 'coarse' instead of 'course'.

Changed

5. P6, L19: ". . . and lateral transport between . . ."; between what?

Changed into "… and net lateral groundwater in-/outflow can be neglected"

6. P7, L10: 'purposes' instead of 'porpuses'.

Changed

7. P7, L23: 'Iberian' instead of 'Iberina'.

Changed

8. P7, L27: 'Romania' instead of 'Romenia'.

Changed

9. P8, L5: 'increase' instead of 'increased'.

Changed

10. P9, L1: 'separate' instead of 'separate', see also line 26 on this page.

Changed

11. P9, L1: The rescaling of the contributions is not clear to me.

Changed into “is rescaled inversely from the 2nd to the 98th percentile of its spatial distribution over Europe”

12. P9, L25: ‘Table 3’ instead of ‘Table 2’.

Tables 1 and 2 are supposed to become a single table in the final layout, so Table 3 is in fact table 2.

13. P9, L26: ‘precipitation’ instead of ‘prepitation’.

Changed

14. P9, L34: ‘4 km^3/year’ instead of 4 km^3’? And ‘-2 km^3/year’ instead of -2 km^3’?

Changed

15. P10, L16: ‘physical’ instead of ‘phyical’.

Changed

16. P26-27, Figure 6-7: ‘Table 3’ instead of ‘Table 2’.

Tables 1 and 2 are supposed to become a single table in the final layout, so Table 3 is in fact table 2.

17. P28, Table 1: How is it possible to use a reference from 1975 to obtain data from periods until 1996? This needs to be adapted.

Indeed this is not possible. The 1975 reference is the best reference to the site, but it does not describe the recent data. Therefore we changed the column title into “Site reference/Source”

18. P28-29, Table 1-2: Which minimum and maximum values are meant for **? And what is the unit? Where can we find the *** in the tables?

We forgot the *** in the table for Plynlimon, which apparently caused confusion. Fixed now.

19. P30, Table 3: ‘km^3 y^-1’ instead of ‘km y^-1’?

Corrected

Climate change, re-/afforestation, and urbanisation impacts on evapotranspiration and streamflow in Europe

Adriaan J. Teuling[1], Emile De Badts[1], Femke A. Jansen[1], Richard Fuchs[2], Joost Buitink[1], Anne J. Hoek van Dijke[1,3,4], and Shannon Sterling[5]

[1]Hydrology and Quantitative Water Management Group, Wageningen University & Research, Wageningen, The Netherlands
[2]Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research, Atmospheric Environmental Research (IMK-IFU), Group of Land Use Change and Climate, Garmisch-Partenkirchen, Germany
[3]Luxembourg Institute of Science and Technology, Belvaux, Luxembourg
[3]Laboratory of Geo-information Science and Remote Sensing, Wageningen University & Research, Wageningen, The Netherlands
[5]Department of Earth Sciences, Dalhousie University, Halifax, Canada

Correspondence: Adriaan J. Teuling (ryan.teuling@wur.nl)

Abstract. Since the 1950s, Europe has ~~seen~~ undergone large shifts in climate and land cover. Previous assessments of past and future changes in evapotranspiration or streamflow have either focussed on land use/cover or climate contributions, or ~~have focussed~~ on individual catchments under specific climate conditions but not on all aspects at larger scales. Here, we aim to understand how decadal changes in climate (e.g., precipitation, temperature) and land use (e.g., de-/afforestation, urbanization) have impacted the amount and distribution of water resources availability (both evapotranspiration and streamflow) across Europe since the 1950s. To this end, we simulate the distribution of ~~green and blue water fluxes~~ average evapotranspiration and streamflow at high-resolution (1 km^2) by combining a) a steady-state Budyko model for water balance partitioning constrained by long-term (lysimeter) observations across different land-use types, b) a novel decadal high-resolution historical land use reconstruction, and c) gridded observations of key meteorological variables. The continental-scale patterns in the simulations agree well with coarser-scale observation-based estimates of evapotranspiration, and also with observed changes in streamflow from small basins across Europe. We find that strong shifts in the continental-scale patterns of evapotranspiration and streamflow have occured ~~from 1950 to~~ between the period around 1960 and 2010. ~~In~~

In much of central-western Europe, our results show an increase in evapotranspiration in the order of 5–15% between 1955–1965 and 2005–2015, whereas much of the Scandinavian peninsula shows increases exceeding 15%. The Iberian peninsula and other parts of the Meditteranean show a decrease in the order of 5–15%. A similar north-south gradient was found for changes in streamflow, although changes in central-western Europe were generally small. Strong decreases and increases exceeding 45% were found in parts of the Iberian and Scandinavian peninsulas, respectively. In Sweden, for example, increased precipitation ~~dominates effects of~~ is a larger driver than large scale re- and afforestation, leading to increases in both streamflow and evapotranspiration. In most of the ~~Mediterrenean~~Mediterranean, decreased precipitation combines with increased forest cover and potential evapotranspiration to reduce streamflow. In spite of considerable local and regional scale complexity, the ~~Europe-wide net contribution of~~ response of net actual evapotranspiration to changes in land use, precipitation~~and potential evapotranspiration changes to changes in ET is similar with~~ , and potential evaporation is remarkably uniform

across Europe, increasing ∼~~40~~35–60 km^3/y, equivalent to the discharge of a large river. For streamflow, effects of changes in precipitation (∼95 km^3/y) dominate land use and potential evapotranspiration contributions ~~with~~ (∼~~90~~45–60 km^3/y~~compared to ∼45 km^3/y~~). Locally, increased forest cover, forest stand age and urbanisation have lead to significant decreases and increases of available streamflow, even in catchments that are considered to be near-natural.

1 Introduction

Streamflow provides an integrated signal both in space and time of all upstream changes in the terrestrial hydrological cycle. At smaller timescales of days to weeks, streamflow reflects the weather conditions and precipitation in the recent past. At longer (multi-year) timescales, over which internal catchment storage changes become much smaller than the amount of water passing through the catchment system, streamflow reflects the amount of water that passes through aquifers and dams (the "water yield"), which is the portion of precipitation that is not returned to the atmosphere via evapotranspiration~~(the so-called "green water" flux, here used to indicate the total evaporative flux)~~. The water yield represents the average water flux that can potentially be exploited for human benefit in a sustainable way~~, and is nowadays often referred to as "blue water" availability or "blue water" flux~~. Quantifying and understanding past and future changes in ~~blue water availability ,~~ water availability in rivers and groundwater systems, reflecting the integrated signal of all net changes in the water cycle upstream, is not only of key importance to water resources management and planning, but it is also a major scientific challenge given the uncertainties and limitations in both observations and models (Wang, 2014). This is in particular true for Europe, where strong changes in land use (in particular urbanisation, re- and afforestation, see Fuchs et al., 2013), and climate (van der Schrier et al., 2013; Caloiero et al., 2018; Bach et al., 2018) have occurred since the 1960s. ~~In the following, we will use several terms interchangeably: green water flux or evapotranspiration (ET) to describe to the total average evaporative return flux to the atmopshere (including interception evaporation), and blue water flux, water yield, or streamflow to describe the average flux of water from a land area (although it should be noted that not every pixel considered might be directly connected to a stream).~~

Several studies have focussed on large-scale changes in ~~green~~ evapotranspiration and/or ~~blue water fluxes~~water availability. In one of the first global studies, Milly et al. (2005) analysed climate-driven changes in ~~blue~~ water availability from an ensemble of climate models and found a general drying of transitional regions and a wetting of current humid and colder regions. Over Europe, the study reported a strong latitudinal gradient in average ~~blue~~ water fluxes increasing in strength from the 20$^{\text{th}}$ to the 21$^{\text{st}}$ century, with decreasing availability trends in the Meditarrenean, and increasing trends in ~~Northern~~ northern Europe. Gerten et al. (2008) showed that globally, precipitation changes were the biggest drivers of changes in runoff, but also land use change had a considerable effect. Changes in ~~Northern Hemisphere~~ northern hemisphere streamflow over the past decades have likely also been impacted by decadal changes in solar radiation (the so-called global "dimming" and "brightening", see Teuling et al., 2009; Gedney et al., 2014). Other studies have focussed on the impact of anthropogenic land-cover change on

~~green water fluxes~~evapotranspiration. Sterling et al. (2012) found a 5% reduction in global ~~ET~~ evapotranspiration due to land cover conversion, resulting in a 7.6% increase in global average streamflow. Other studies have highlighted strong decadal-scale variability in global average ~~ET~~ evapotranspiration over the recent decades related to El Niño–Southern Oscillation (Jung et al., 2010; Miralles et al., 2014). In spite of the direct link between average ~~green and blue water fluxes~~evapotranspiration and streamflow, few studies have addressed changes in both fluxes simultaneously.

Because streamflow is impacted by many factors, which often have opposing effects, changes in streamflow should be considered at small scales at which indivual factors can be understood and quantified rather than at larger river basin scales. Although several long discharge records exist for large river basins, changes that occur at the sub-basin level are often obscured by opposing effects in other parts of the basin. In a landmark study, Stahl et al. (2010) addressed this limitation by analysing streamflow changes in Europe from a dataset of relatively small river basins with limited human influence. They reported a diverging pattern of streamflow trends over the past decades, with negative trends in annual mean streamflow in many parts of the ~~Mediterrenean and Central~~ Mediterranean and central Europe, and predominantly positive trends in ~~Western~~ western Europe and parts of Scandinavia. While the longer-term and long-range variability of streamflow in these basins and its relation to circulation indices is generally well understood at the interannual and decadal timescales (Gudmundsson et al., 2011; Hannaford et al., 2013), significant uncertainty exists in understanding the regional-scale variability in trends since these are not well reproduced by the current generation of hydrological models (Stahl et al., 2012). Previous ~~case studies for catchments~~ regional case studies across Europe have reported a sensitivity of long-term water balance partitioning to both climate and land use change ~~(Parkin et al., 1996; van Roosmalen et al., 2009; van der Velde et al., 2013; Pijl et al., 2018)~~ (Parkin et al., 1996; van Roosmalen et al., 2009; van der Velde et al., 2013; Renner et al., 2014; Pijl et al., 2018). Thus, quantifying changes in streamflow requires accounting for changes in climate (precipitation and potential evapotranspiration) as well as changes in land use and/or land cover (Stonestrom et al., 2009). But whereas assessing the impact of climate on average streamflow is relatively straightforward, the role of land cover requires a more careful consideration.

At the smaller scale, land use, in particular forest cover, has longsince been known to have a strong impact on average streamflow or water yield, with forested catchments having a much lower water yield compared to non-forested catchments (Bosch and Hewlett, 1982; Zhang et al., 2001; Brown et al., 2005; Farley et al., 2005; van Dijk and Keenan, 2007; Filoso et al., 2017). Based on analysis of paired catchment observations, a large majority of studies have found that removal of forest leads to an increase in water yield. While this is likely linked to higher average evapotranspiration over forest, the reverse has been reported for dry and warm summer conditions based on eddy-covariance observations from FLUXNET (Teuling et al., 2010). Somewhat surprisingly, average evapotranspiration rates for forested FLUXNET sites are on generally slightly lower than for non-forested sites (Williams et al., 2012), which is seemingly inconsistent with other studies (e.g., Zhang et al., 2001) where annual evapotranspiration was inferred from the water balance (the so-called “forest evapotranspiration paradox”, see Teuling, 2018). A possible explanation for this discrepancy is the role of interception (van Dijk et al., 2015). Several studies (e.g., Gash et al., 1980; Zimmermann et al., 1999) have shown that interception can constitute a major term in the water balance of forested ecosystems, in particular in humid conditions (Calder, 1976; Ramírez et al., 2018). Controlled experiments on large non-weighable lysimeters covered with forest have shown that growing forest strongly reduces the water yield (Tollenaar and

Ryckborst, 1975; Harsch et al., 2009; Müller, 2009; Teuling, 2018), and that this effect is somewhat larger for coniferous than for deciduous species. This is in line with results from a large number of basins in Sweden, where increases in forest cover and biomass (age) were the main factor explaining observed trends in inferred evapotranspiration (Jaramillo et al., 2018). This shows that both forest cover area but also stand age need to be taken into account when evaluating land use change effects on water balance partitioning.

In contrast to forest cover, few studies have quantified the effects of urban area and urbanization on the long-term water balance partitioning~~are largely unknown~~. Runoff from urban areas is typically measured with focus on short-term dynamics and event runoff ratios (Berthier et al., 1999) or runoff produced by impervious areas only (Boyd et al., 1993; Shuster et al., 2005). Evapotranspiration from urban areas, or the other hand, is typically measured or analysed over individual elements that make up the urban landscape, such as (un)paved areas (Ramamurthy and Bou-Zeid, 2014), green roofs, or trees (Pataki et al., 2011). Few studies have measured ~~ET~~ evapotranspiration at the urban landscape scale. In a study comparing measurements made over the Dutch cities of Rotterdam and Arnhem, Jacobs et al. (2015) found ~~ET~~ evapotranspiration rates to be generally low, and to quickly drop in the days following rainfall reflecting a strongly water-limited system. Similar results were found for the Swiss city of Basel (Christen and Vogt, 2004). This suggests that urban areas, because of their limited capacity to store water, might have much lower ~~green water fluxes~~ evapotranspiration and as a result ~~much higher blue water fluxes~~ might generate much higher streamflow than other land use types. This was also reported by DeWalle et al. (2000) based on statistical analysis of long-term streamflow record in the United States. They found strong increases in streamflow in areas with heavy urbanisation, which was attributed to a decrease in evapotranspiration.

In order to isolate and/or attribute the hydrological impact of climate change from that of changes in land use, ~~models with varying levels of complexity are used.~~ different methods exist (see reviews by Wang, 2014; Dey and Mishra, 2017). The methods can be categorized into experimental approaches, hydrological modeling, conceptual approaches, and analytical approaches (Dey and Mishra, 2017). Typically, hydrological or land surface models run at hourly or daily resolution are used (Bosmans et al., 2017; Breuer et al., 2009; Viney et al., 2009; Dwarakish and Ganasri, 2015; Pijl et al., 2018). Such models often contain a high number of poorly-constrained parameters and parameterizations, leading to large uncertainty in trend estimates (Arnell, 2011), or even disagreement in the direction of simulated trends (Melsen et al., 2018). When the research focus is on robust simulation of long-term rather than short-term changes, low-dimensional models with well-constrained parameters often perform well (Choudhury, 1999; Zhang et al., 2008). The Budyko model (Budyko, 1974) is an example of such a ~~model~~ conceptual approach which allows for evaluating combined land use and climate impacts on water availability (see, for example Jiang et al., 2015). In spite of its extreme simplicity (parameterizations typically have only one parameter reflecting land surface characteristics), it has been applied successfully in numerous studies focussing on different controls on long-term water balance partitioning (Zhang et al., 2004; Roderick and Farquhar, 2011; Xu et al., 2013; Greve et al., 2014; Xu et al., 2014; Creed et al., 2014; Jiang et al., 2015; Zhang et al., 2016; Wei et al., 2018). Although it is generally applied at coarse global grid resolution or to large river basins, other studies (e.g. Zhang et al., 2004; Redhead et al., 2016) have found the model to also work well for smaller basins or grid cells (< 10 km^2). This opens up the possibility for robust and parsimonious modeling of hydrological impacts at high spatial resolution.

The strong impact of land use on water balance partitioning at smaller scales, combined with the large-scale land use changes that have occurred over Europe over the past decades, leads to the question how they have impacted changing patterns of ~~green and blue water fluxes~~evapotranspiration and streamflow. Previous assessments of past and future changes in ~~streamflow~~ water balance partitioning have either focussed on land use/cover (Sterling et al., 2012) or climate contributions (Wilby, 2006; Gardner, 2009; Hannaford et al., 2013), or have focussed on smaller catchments under particular climate conditions ~~(van Roosmalen et al., 2009; Pijl et al., 2018)~~(van Roosmalen et al., 2009; Renner et al., 2014; Pijl et al., 2018). Therefore, we aim to understand how recent decadal changes in climate (e.g., precipitation, temperature) and land use (de-/afforestation, urbanization) have impacted the amount and distribution of water resources availability across Europe since the 1950s. We address the hypothesis that land cover changes play a much more important role at the European scale than previously reported, even in basins which are assumed to have a limited human influence on the water cycle. To this end, we simulate the distribution of ~~green and blue water fluxes~~ evapotranspiration and streamflow at high-resolution (1 km^2) by combining a) a steady-state Budyko model for water balance partitioning constrained by long-term observations across different land-use types, b) a novel decadal high-resolution historical land use reconstruction, and c) gridded observations of key meteorological variables. Simulations will be evaluated against state-of-the-art observation-based assessments of evapotranspiration and observed changes in streamflow.

2 Methods and Data

Central to our approach is the formulation of the Budyko model as used by Zhang et al. (2004). As with any Budyko approach, it follows the central ~~assumption~~ assumptions that the fraction of precipitation that returns to the atmosphere as evapotranspiration ET depends on the ratio between the average potential evapotranspiration PET and average precipitation P, rather than on their absolute values, and that a catchments' ET, when a catchment is subjected to a range of climate conditions, follows a single path in the ET/P,PET/P-space. Good fit with observations at several spatial scales show that this assumption is ~~justified. Here, we will calculate PET by the Thornthwaite method (see, e.g., van der Schrier et al., 2011). This method only requires temperature as input, and as a result it is not sensitive to changes in other variables affecting evaporation, such as vapor pressure deficit, wind speed, or net radiation. While the benefit of this approach is that the analysis can be carried out beyond the record for routine incoming shortwave and/or net radiation observations, it has as main drawback that temperature is not always a reliable proxy for radiation in particular under global warming. The potential implications of this assumption are discussed in Section 4.~~

generally justified. In the work by Zhang et al. (2004), the following equation was proposed for the dependency of ET/P on PET/P:

$$\frac{\mathrm{ET}}{P} = 1 + \frac{\mathrm{PET}}{P} - \left[1 + \left(\frac{\mathrm{PET}}{P}\right)^{w}\right]^{1/w} \quad (1)$$

in which w is a model parameter which is typically linked to catchment and/or vegetation properties (Li et al., 2013). Zhang et al. (2004) found $w = 2.63$ to best fit observations for Australian catchments, with slightly lower values for grassed ($w =$

2.55) and higher for forested catchments ($w = 2.84$). While these different values confirm that w depends on land surface characteristics, the magnitude of this ~~effect~~ dependency at the scale of individual land use elements, rather than catchments with a land use mixtures of varying degrees, is probably larger. Based on analysis of remotely sensed Normalized Difference Vegetation Index (NDVI) and gridded global fields of ET, PET, and P at the $0.5°$ resolution, Greve et al. (2014) reported values of 3.05 for grid cells with an NDVI of around 0.8, whereas grid cells with an NDVI of around 0.2 where found to follow $w = 1.63$. In a similar study but using observed streamflow rather than estimated ~~E~~ET, Li et al. (2013) found w to depend on the basin-average fractional vegetation cover M according to $w = 2.36 \times M + 1.16$. These studies show that w can show considerable variation even at relatively ~~course~~ coarse scales.

In order to get the most realistic values for w for application at smaller scales (~ 1 km^2) at which land use is often fairly homogeneous and the effects on water balance partitioning are most pronounced, we constrain w by the best available observations for different land use types and made under European climate conditions. It should be noted that widely available FLUXNET observations are not used in this study, because they might show the opposite land use ET signal from water balance-based ~~studies~~ observations (the so-called forest evapotranspiration paradox, see Teuling, 2018)~~which are assumed~~. The latter are assumed here to be more reliable for ~~long-term water balance analysis~~our application. The observations used in this study primarily come from the long-term lysimeter stations, such as the ones at Rietholzbach (Seneviratne et al., 2012), St. Arnold (Harsch et al., 2009), Brandis (Haferkorn and Knappe, 2002), Eberswalde-Britz (Müller, 2009), Castricum (Tollenaar and Ryckborst, 1975), and Rheindahlen (Xu and Chen, 2005), several of which were also analysed in a previous study by Teuling (2018). This data is complemented by observations from a natural lysimeter at Plynlimon (Calder, 1976) under more humid climate conditions and flux observations made over the cities of ~~Arnhem~~ Basel (Christen and Vogt, 2004), Arnhem, and Rotterdam (Jacobs et al., 2015). Long-term data is preferred to minimise impacts of interannual storage variations (Istanbulluoglu et al., 2012). By relying on lysimeter observations to constrain our Budyko parameters, we implicitly assume lysimeters (area ~~1–625~~varies from 1 to 625 m^2 for the larger lysimeters at Castricum) to behave similar to landscape elements of 10^6 m^2 (our grid cell size). The data is shown in Fig. 1 and listed in Table 1. It should be noted that the stations are not distributed evenly across Europe, but are mainly constrained to ~~Central-Western~~ central-western Europe (Fig. A1).

Due to the smaller scale than applied in previous Budyko analyses, we initially find many points, in particular observations from forested lysimeters, to be located above the energy-limit (grey dashed line in Fig. 1). This indicates that the long-term average yearly evapotranspiration (ET) exceeds the average potential evapotranspiration (PET). This is possible, for instance, due to evaporation of interception water by energy not captured in the formulation of PET (van Dijk et al., 2015). Therefore, we correct for the underestimation by introducing a so-called adjusted potential evapotranspiration (aPET) which is assumed to be proportional to the potential evapotranspiration and accounts for all processes affecting yearly ET for tall vegetation (including evaporation of intercepted water through advection) so that ET generally will not exceed aPET even for forested areas:

$$\mathrm{aPET} = c \times \mathrm{PET}, \tag{2}$$

resulting in the following expression for the Budyko-curve:

$$\frac{\mathrm{ET}}{P} = 1 + \frac{\mathrm{aPET}}{P} - \left[1 + \left(\frac{\mathrm{aPET}}{P}\right)^{w^*}\right]^{1/w^*} \tag{3}$$

in which w^* is the value for w when aPET rather than PET is used. ~~The~~ aPET should thus be interpreted as the maximum total yearly evapotranspiration that would occur under given climate conditions (PET/P) and under optimal vegetation conditions (i.e. vegetation that is most efficient in returning precipitation to the atmosphere, in this case needleleaf forest), rather than a land use-specific crop factor. It should be stressed that this scaling is only done to move most observations within the energy and water limits in the Budyko space (so that we can obtain a fit with Eq. 1), and that it has no other impact on the results since the model is subsequently forced with aPET rather than PET. The resulting values for w^* that match the (lysimeter) observations are shown in Fig. 1. It was found that ~~$c = 1.8$~~ $c = 1.6$ was required to ensure ~~all~~ most observations would be located on the right-side of the energy-limit ($\mathrm{ET}/P = \mathrm{PET}/P$). Subsequently, in all analysis we replaced PET with aPET, including Eq. 1 but also in the atmospheric forcing fields. It should be noted that while this procedure results in lower values for w^* that cannot be directly compared to values for w reported in previous studies, most of the simulated ET values are identical to the ones that would be simulated with the original model. We find the highest w^* for full-grown forest, indicating that any change towards this state due to re- or afforestation will increase ET given the same climate (P and PET). We distinguish between young stands (age $<$ 10 years, assumed to behave similar to crop-/grasslands based on the data in Fig. 1 with w^*=1.7), intermediate (age 10–20 years, w^*=2.3) and older stands (age $>$ 20 years, w^*=3.1), see also Fig. 1. In this way, we implicitly account for effects of increasing biomass, tree height and stand age on ET and water yield reported in previous studies (Harsch et al., 2009; Jaramillo et al., 2018; Teuling, 2018). Conversly, urban areas have a low w^* of 1.3, indicating that urbanisation will generally decrease ET. Finally, the long-term average streamflow or ~~blue water flux~~ water yield at the pixel level is calculated from the catchment water balance:

$$Q \approx P - \mathrm{ET} \tag{4}$$

under the assumption that storage changes (such as snow, soil moisture, groundwater) and ~~lateral transport between~~ net lateral groundwater in-/outflow can be neglected at the decadal (10-year) timescale. This timescale is chosen to align with the temporal resolution of the land use dataset, and to minimize possible impacts of storage changes.

As input to our model as described above (Eqs. 2–4), we use gridded datasets of land cover and meteorological observations. All calculations ~~where~~ were performed at a 1×1 km spatial resolution, which were later rescaled to a coarser resolution for visualization purpose. Historic land-change information is based on the HIstoric Land Dynamics Assessment (HILDA, v2.0) model reconstruction of historic land cover/use change (Fuchs et al., 2013, 2015a, b). This data-driven reconstruction approach used multiple harmonized and consistent data streams such as remote sensing, national inventories, aerial photographs, statistics, old encyclopedias and historic maps to reconstruct historic land cover at a 1×1 km spatial resolution for the period 1900 to 2010 in decadal time steps. The reconstruction provides information for six different land cover/use categories: forest, grassland (incl. pastures, natural grasslands and shrublands), cropland, settlements/urban, water bodies and other (i.e. bare rock, glaciers etc.). Here we only use the forest, grass-/cropland, and settlement classes. The reconstruction considers gross land changes, the sum of all area gains and losses that occur within an area and time period, unlike other reconstructions that focus on net changes only, calculated by area gain minus the area losses. Details on the net versus gross changes can be found in Fuchs et al. (2015a). The gross changes are used to derive forest stand age. ~~We distinguish between young stands (age $<$ 10 years, assumed to behave similar to crop-/grasslands), intermediate (age 10–20 years) and older stands (age $>$ 20 years), see also Fig. 1.~~ Previous research has shown that not accounting for gross land use changes in reconstruction led to serious underestimations in the amount of total land use changes that have occurred (Fuchs et al., 2015a). The E-OBS ~~v17 gridded datasets~~ v18 gridded dataset (Haylock et al., 2008) of observed precipitation ~~and temperature~~ at 0.25° resolution and the CRU TS v4.02 gridded dataset (Harris et al., 2014) of observed potential evapotranpiration at 0.5° resolution were used to force the model (~~Eq.~~ Eqs. 2 and 3). ~~Temperature was used to calculate PET according to the Thorntwaithe method.~~ Based on the joint availability of ~~both~~ the HILDA, CRU and E-OBS datasets, we selected two 10-year periods which were considered for analysis: 1955–1965 and 2005–2015. In the following, we will refer to these periods as 1960 and 2010 for simplicity. While the 10-year periods are often considered short for climate change detection, they resulted from a need to balance robust estimation of the mean climate, without averaging out much of the underlying changes in both climate and land use. Changes over the intermediate 10-year periods were analysed, but since the trends were ~~generally~~ found to be mostly monotonic the results are not shown here (except for validation ~~porpuses~~ purposes in Fig. 5).

Model simulations are validated and compared against observed yearly average streamflow changes in near-natural catchments and observation-based average evapotranspiration. The relative streamflow changes for the period 1962–2004 (normalized by the standard deviation of yearly streamflow) were ~~taken directly from~~ used as presented in Stahl et al. (2010, their Figure 2). Average evapotranspiration was derived from GLEAM v3.2a (Martens et al., 2017). The contribution of P, PET and land use (through w^*) was assessed by performing separate simulations in which only one of the three factors was varied while the others were kept constant at their 1960s reference.

3 Results

Recent changes in climate have lead to substantial changes in the magnitude and distribution of precipitation and potential evapotranspiration, the two main climate drivers in the Budyko model (Eq. 1) that determine how average precipitation is partitioned between evapotranspiration and streamflow. Average precipitation during the reference period shows a general decrease towards the East (Fig. 2a). Superimposed on this large-scale pattern are local areas with higher precipitation along the coastal areas in the ~~West~~ west and/or in mountainous regions. Changes in average precipitation over the study period show a strong ~~North-South~~ north-south gradient (Fig. 2b): Most of the ~~Mediterrenean~~Mediterranean, in particular the ~~Iberina Peninsula~~Iberian peninsula, shows a decline in precipitation, whereas ~~Northern~~ northern Europe, in particular the British Isles, the Scandinavian ~~Peninsula~~ peninsula and Finland, have seen strong increases in average precipitation regionally exceeding 20%. In contrast to precipitation, potential evapotranspiration shows a strong latitudinal gradient (Fig. 2c) with lower values (PET around 400 mm/y) in Scandinavia and higher (~~around 700~~regionally exceeding 1000 mm/y) in the ~~Mediterrenean~~Mediterranean. Changes in potential evapotranspiration (Fig. 2d) are predominantly positive ~~(decreasing values in Romania likely reflect a data-quality issue)~~

and highest in ~~Central~~ central Europe reflecting the higher increase in average temperatures and shortwave radiation. In general, these strong changes in climate forcing (P and PET) are likely to be reflected in continental-scale patterns of changes in water availability.

In addition to climate, also land use and land cover in Europe have seen large scale shifts over the past 60 years, albeit on a more local scale. Figure 3 shows the mean forest and urban fraction for the reference period, as well as the fractional change over the period 1960–2010. While forest cover is widespread over most of Europe (Fig. 3a), most extensive forest regions can be found in central-western Europe, Sweden and Finland. Forest cover has increased considerably over most of Europe (Fig. 3b) following abandonment of less-productive agricultural areas and intensification of forestry and forest management, with Sweden (Ericsson et al., 2000) and the ~~Mediterrenean~~ Mediterranean region showing the strongest changes. It should be noted that areas where forest cover has declined are virtually absent. This is also true for change in urban area. The average urban fraction is highest in central-western Europe (Fig. 3c), ~~and this~~ in particular in Belgium, the German Ruhr area, and The Netherlands. This is also the region that has seen the strongest ~~increased~~ increase (Fig. 3d). Changes in urban area are generally more localized in nature than changes in forest cover.

Patterns of mean and changes in evapotranspiration and water yield were calculated by forcing the Budyko model with subsequent 10-year averages of climate forcing and land use at a 1×1 km resolution. Figure 4 shows the resulting continental-scale patterns. The mean evapotranspiration in the reference period (Fig. 4a) is highest in central Europe, locally exceeding ~~800~~600 mm/y, in regions with ~~pronounced topography~~ topographically-enhanced precipitation and/or forest cover. The Nordic countries and the Iberian ~~Peninsula~~ peninsula generally have lower values (<400 mm) due to more pronounced energy and water limitation, respectively. Changes in evapotranspiration show a strong latitudinal gradient (Fig. 4b). Changes exceeding +15% are found in large parts of Scotland, Sweden, Finland, and Estonia, whereas most of ~~Central-Western~~ central-western Europe shows a smaller increases in the order of 10%. Decreases of similar magnitude occur in parts of the Iberian ~~Peninsula~~ peninsula and Italy. Average streamflow (Fig. 4c) is highest in ~~Central-Western~~ central-western Europe (locally exceeding 600 mm/y), in particular in mountainous areas that receive larger amounts of precipitation. Streamflow of less than 150 mm/y is found in the large parts of Sweden, Finland, Spain, ~~Rumenia~~ Romania and Bulgaria. Changes in water yield (Fig. 4d) show a roughly similar pattern to changes in evapotranspiration, however the changes are much stronger in magnitude. Decreases in the ~~Mediterrenean~~ Mediterranean locally exceed −45%, where increases in Sweden and Finland exceed +45%. Both the changes in evapotranspiration and streamflow show considerable regional variability superimposed on the large-scale patterns.

In order to assess the quality of the simulated evapotranspiration and streamflow and the changes therein, we evaluate our simulations against observation-based estimates of average evapotranspiration (Martens et al., 2017) over the more recent period 1980–2017 (it should be noted that currently no gridded evapotranspiration estimates are available that cover our complete study period) as well as observed changes in streamflow reported by Stahl et al. (2010) that cover most of our study period. The pattern of simulated ET (Fig. 5a) closely resembles the pattern as produced by GLEAM version 3.2a (Martens et al., 2017, data shown in Fig. 5b). It should be noted that this comparison is added for reference only and should not be seen as validation: GLEAM is not a strictly observational dataset, and it does not necessarily provide better long-term estimates of ET for forest and urban areas. The Budyko model produces slightly lower values in Eastern Europe and the Iberian ~~Peninsula~~peninsula, but slightly higher values in Sweden and Finland. At the regional scale, our simulations show more variability due to the higher resolution of the forcing and land use datasets. In addition to matching the pattern of average ET, our approach is also able to reproduce the overall pattern of observed changes in streamflow (Fig. 5c,d). ~~In spite of the difference in units and the fact that individual basins might have shorter record lenghts, the correlation in trends between the basins is 0.34.~~ The simulations agree with the observed declines in average streamflow in much of ~~Southern and Central~~ southern and central Europe, and increases in the more mountainous, coastal and/or ~~Northern regions.~~ northern regions. The two-dimensional frequency distribution (Fig. 6) confirms the capability of our approach in reproducing the observed trends in Figs. 5d, with a much higher frequency in the outer quartiles along the diagonal (12% and 9.7% compared to 6.25% expectation) than across the diagonal (4.8% and 2.8% of catchments). It should be noted that a higher-order validation on trends is subject to more noise than validation on mean fields, and a perfect match should not be expected also due to the difference in normalization. Figure 7 shows that our simulations also add information with respect to trends in forcing (P and PET), where PET and to a lesser extent P show a predominant increase over all basins while observed trends center around zero change. Overall, the validation shows that our simplified approach is able to capture continental-scale patterns in mean and changes of ~~blue and green water fluxes~~evapotranspiration and streamflow.

~~Changes~~ In order to understand how changes in fluxes are driven by local changes in climate and land use~~. Figures 6 and 7~~ , Figures 8 and 9 show how the contribution of the main drivers (precipitation, PET, and land use) to changes in evapotranspiration ~~and streamflow, respectively, vary~~ (Fig. 8) and streamflow (Fig. 9), varies across Europe. This is done by plotting each contribution (as determined from simulations where the other drivers where kept constant) as a ~~seperate~~ separate RGB component, whereby each contribution is rescaled inversely from the 2nd to the 98th percentile of its spatial distribution over Europe. The resulting colormap thus has a 3D color legend. From the distribution of colors, covering most of the possible colors, it can be readily seen that contribution of individual drivers shows a strong variability. Magenta indicates that land use-induced changes in evapotranspiration and streamflow are widespread but generally local in character. Yellow colors occur widely in a latitudinal band between 45°N and 54°N, indicating that changes in PET have a strongest influence on water balance partitioning in transitional regions, but less so in water-limited and humid ~~Northern~~ northern regions. Finally, the relative impact of precipitation is strongest above 54°N and below 45°N. It should be noted that these continental-scale patterns differ from their changes, which are much more uniform (e.g. PET changes in Fig. 2d are fairly homogeneous).

~~The~~ For a more quantitative regional assessment, the subpanels in Figures ~~6 and 7~~ 8 and 9 zoom in on several regions~~, further illustrating~~ . These further illustrate the strong regional divergence in changes in water flux partitioning. In the ~~Southern~~ southern Highlands of Scotland (Figs. ~~6a~~8a/~~7a~~9a), a strong increase in precipitation has lead to a strong net increase in streamflow of +~~339~~362 mm/y, only slightly counteracted by opposing PET and land use (afforestation) effects. Urbanisation in the Paris metropolitan area (Figs. ~~6b~~8b/~~7b~~9b) has act to reduced ET (−18 mm/y), but combines with increased P to a significant increase in streamflow (+~~44~~38 mm/y). In the Landes forest region (Figs. ~~6c~~8c/~~7c~~9c), individual effects are small but combine to a strong (−~~79~~90 mm/y) reduction in water yield. ET changes in the Seville region (Figs. ~~6d~~8d/~~7d) are relatively small~~ 9d) are moderate (−53 mm/y) due to opposing contributions of precipitation decline and afforestation, but these effects combine into a strong reduction on streamflow (−~~100~~80 mm/y). In Sweden, ET changes (Figs. ~~6e~~8e/~~6f~~8f) are stronger in the middle of the country where widespread afforestation and precipitation increase combine (+95 mm/y). As a result, increases in streamflow are stronger in the ~~South~~ south (+120 mm/y) where land use contributions do not reduce the effect of precipitation increase (Figs. ~~7e~~9e/~~7f~~9f). In ~~Central~~ central Austria (Figs. ~~6g~~8g/~~7g~~9g), PET increases dominate the net ET change (+~~31~~45 mm/y), but combine with precipitation reduction into a strong reduction of water yield (−~~94~~108 mm/y). In the Bulgarian Smolyan Province (Figs. ~~6h~~8h/~~7h~~9h), contributions combine into a strong ET increase (+~~69~~80 mm/y) but largely cancel out in the net effect on water yield (−~~8~~19 mm/y). The examples highlight the fact that locally, individual changes are often amplified or counteracted by other changes, but because of the water balance constraint this is only true for impacts on either evapotranspiration or streamflow.

When the results are averaged over the continental scale, land use plays a more important role than suggested by Figure ~~6.~~ 8. Table 2 lists the Europe-wide changes in ~~green or blue water fluxes~~ evapotranspiration and streamflow as induced by the three main drivers. While changes in ET induced by ~~prepitation~~ precipitation are largest when positive and negative contributions are considered ~~seperately~~separately, the net effect is smaller since decreases in P in southern Europe are largely balanced by increases in the northern parts. As a result, net effects of land use and PET on ET are comparable to those of precipitation (around 40 $\mathrm{km^3\,y^{-1}}$ each), with land use having the largest contribution. These contributions correspond to nearly 1300 $\mathrm{m^3\,s^{-1}}$, the equivalent of the discharge of a large river. The effects on streamflow differ slightly, with P dominating both the positive and net contributions. When zooming in on the near-natural catchments used by Stahl et al. (2010), a different picture is obtained. The contribution of P is less strong, likely because most of the catchments are located in ~~Central-Western~~ central-western Europe where precipitation changes have been modest (Fig. 2b) compared to, for instance, Sweden. The net change in ET is mainly driven by land use and PET. For streamflow changes, P is the largest net contributor at ~~nearly~~ around 4 $\mathrm{km^3\,y^{-1}}$, but land use contributes significantly with nearly -2 $\mathrm{km^3\,s^{-1}}$. For individual large river basins, such as the Rhine basin shown here, the impacts can differ significantly. Rather than precipitation, land use and PET are found to be the main drivers of changes in ~~blue water fluxes~~ streamflow over the past decades. The strong sensitivity of streamflow to past land use changes seemingly contradicts the small land use effects under future land use scenarios for this catchment found in previous studies (e.g. Hurkmans et al., 2009).

4 Discussion

Our results on changes in water balance partitioning over Europe are in line with many more local or regional-scale studies. In some regions, studies have found little to no trends due to dominance of natural variability on change indicators (Hannaford, 2015). For the 6.5 $\mathrm{km^2}$ Hupsel Brook catchment in the east of the Netherlands, Brauer et al. (2018) reported no significant trend in annual runoff since the mid 1970s. In one of the few studies on long-term in situ observations of ET, Seneviratne et al. (2012) reported no significant trends of annual ET at the Rietholzbach lysimeter in north-eastern Switzerland. These findings are consistent with the results on changes in ET and Q presented in Figs. 4b and 4d. Other regions have seen negative trends. The decline in water yield in the Ebro river has been attributed to land abandonment (López-Moreno et al., 2011), whereas precipitation decline has been identified as an additional factor in most of the Iberian peninsula (Lorenzo-Lacruz et al., 2012).

In Austria, increased P and PET has been identified as factors driving ET increase (Duethmann and Blöschl, 2018). In Sweden, Jaramillo et al. (2018) found little change in the ratio ET/P in spite of strong increases in P and PET. Also these findings are consistent with our results. This shows that even using gridded observations contain consistent information for local-scale change analysis.

The modelling approach followed here is simplified both in terms of number of model parameters, land use classes, and the parameterization of climate. While the single model parameter w^* correlates with ~~phyical~~ physical land surface properties, it does not have a direct physical meaning (although expressions can be derived linking Budyko parameters to vegetation and climate characteristics, see Gerrits et al., 2009). Therefore w^* might also change with mean climate conditions~~and~~, the synchronicity between precipitation and potential evapotranspiration, changing snow conditions (Berghuijs et al., 2014) and/or vegetation phenology (Donohue et al., 2007)~~, an effect that cannot~~ . This could not be investigated due to a lack of observations in ~~Southern and Northern~~ southern and northern Europe. It has also been argued that the success of Budyko approaches can be partly explained by the possible adaptation of vegetation to difference in climate seasonality and soil type (Gentine et al., 2012), which would be a strong argument in favor of using such simplified models. We also use a limited number of land use classes. This number is constrained both by the limited availability of accurate estimates of long-term water balance partitioning for different land use types, as well as by the limited number of land use classes in the HILDA land use reconstruction. Nonetheless, ~~we believe our simulations to capture the first-order~~ our simulations capture the most important land use and climate-induced impacts.

The lysimeter observations include land use with some of the highest and lowest reported ET rates, making it unlikely that we underestimate the land use-induced variability in ET. And whereas there can be considerable variability in average ET within land use classes, for instance due to vegetation and/or soil type (Haferkorn and Knappe, 2002), this variability is typically small compared to the possible range of ET over all land use classes. The range in w^* values is also consistent, at least qualitatively, with estimates in previous studies (Li et al., 2013; Greve et al., 2014). Our modeling approach did not explicitly consider effects other than atmospheric temperature as climate drivers of ET. For instance, the impact of rising CO_2 levels on transpiration (Piao et al., 2007) were not considered, although the effects of CO_2 where found to be small compared to effects of forest stand age (Jaramillo et al., 2018). Also the impact of agricultural intensification (Liu et al., 2015) and irrigation on ET were not considered, although in some regions the effect of irrigation can be considerable (Siebert and Döll, 2010; Jaramillo and Destouni, 2015). Both can be expected to lead to higher ET and lower ~~Q~~ Q. While such processes can have strong impacts locally and regionally, other studies have shown small effect under European conditions (e.g., van Roosmalen et al., 2009). It should be mentioned that other, more rigurous, methods have been applied at smaller scales based on multiple working hypotheses (Harrigan et al., 2014), that allow for identification of additional factors driving hydrologic change.

The model forcing is based on interpolated observations from weather stations. The location of these stations generally follows WMO recommendations (see e.g. Ehinger, 1993), and as a result there is a lack of meteorological observations in, near, or above forests (Frenne and Verheyen, 2016), or in urban areas. Large forest or urban areas, however, are known to impact their own weather for instance due to enhanced temperature ~~(the well-known urban-heat island effect)~~(the well-known urban-heat island effect), cloud formation ~~(as has been observed over the larger French forest regions of Landes and Sologne, see Teuling et al., 2017)~~ (as has been observed over the larger French forest regions of Landes and Sologne and cities of Paris and London, see Teuling et al., 2017 , or rainfall (as has been shown by modeling experiments for the Dutch Veluwe forest region, see ter Maat et al., 2013). Such local land cover impacts on climate are unlikely to be represented correctly in the forcing dataset used in this study which is based on interpolation of weather station data. Also the quality of the data underlying the E-OBS and HILDA datasets used in this study might differ between countries. As a result, the datasets might induce "jumps" near to borders as can be seen in some of the maps. These inconsistencies will likely be fixed in future releases of the datasets, and do not impact the overall conclusions of this study.

The model forcing of potential evapotranspiration is determined ~~by a simple temperature-based parameterization , namely the Thornthwaith method~~using the Penman-Moneith parameterization (Harris et al., 2014), which accounts for temperature, radiation, humidity, and wind speeds effects on evapotranspiration. The benefit of this approach is that ~~simulations can be done in a consistent and robust manner for a longer historical period~~it is the most physical model for potential evapotranspiration, but the larger number of variables involved also increases the risk of spurious trends. Routine observations of ~~global~~ net radiation, needed to force more complex parameterizations such as the Penman-Monteith equation, are only available for the most recent decades from either stations or satellite. Often, they are calculated from other (uncertain) input data. This raises the question whether decadal trends in radiation (i.e., global dimming and brightening, see Wild, 2016), are correctly represented in long-term PET datasets based on Penman-Monteith. Potentially, trends in PET might be underestimated. A major disadvantage of ~~the~~ simpler temperature-based methods is that, while they correctly follow the intra-annual variations in energy, they might be too sensitive to interannual and decadal variations in temperature that are independent of radiation trends (Sheffield et al., 2012). ~~In spite of the possible overestimation of the temperature effect on PET trends, we believe the impact on our main results to~~ The different between temperature-based and more physical representations will be minimal in particular in drier (semi-arid) regions with seasonal water limitation due to the reduced sensitivity of ET to PET (van der Schrier et al., 2011). ~~This is partly since the temperature trends in Europe have been impacted by decadal trends in radiation (i.e., global dimming and brightening, see Wild, 2016), and also because even with the possible overestimation the trends in ET and Q are still dominated by changes in land use and precipitation.~~ It has also been reported that ~~Thornthwaite does not always gives~~ temperature-based methods such as Thornthwaite do not always give the strongest increase in PET ~~values~~ in a warming climate when compared to other more physically-based methods ~~such as Penman-Monteith (Prudhomme and Williamson, 2013)~~(Prudhomme and Williamson, 2013), suggesting that PET-induced changes in water balance partitioning should be interpreted with care.

Changes in climate and land use generally affect both the average ~~green and blue water flux~~evapotranspiration and streamflow. But whereas changes in ~~green water fluxes~~ evapotranspiration are needed to explain changes in ~~blue water~~streamflow, the socio-economic impact ~~relates~~ relate more directly to ~~blue water fluxes since they reflect~~ streamflow since this reflects average fresh water availability. This is of particular relevance in the Mediterranean region, where a decline in ~~blue water flux~~ water yield or streamflow reflects a decrease in water available for irrigation and agricultural production downstream. Our results indicate that land use changes in the more mountainous areas in the ~~Mediterrenean~~ Mediterranean have contributed significantly to reductions in streamflow. Conversely, increasing ~~blue water fluxes in Northern~~ streamflow in northern Europe might be beneficial to other sectors as such the hydropower industry. The finding that land use change effects are of similar magnitude as climate change effects on water availability also has important implications beyond the yearly average values. Extremes will likely also be impacted by land use, yet current drought projections for Europe (Forzieri et al., 2014; Samaniego et al., 2018) or assessments of changes in floods (e.g. Hall et al., 2014) do not take into account past and/or future land cover changes. Not accounting for land use change will likely lead to regional over- or underestimation of changes in water availability. ~~Therefor~~Therefore, land use change impacts on ~~green and blue water fluxes~~ evapotranspiration and streamflow need to be considered in conjunction with climate change impacts.

5 Conclusions

In this study, we investigated the role of changes in land use and climate in Europe from 1960 to 2010 on average evapotranspiration and streamflow. In our modeling approach, we combined a state-of-the-art land use reconstruction with gridded observational datasets of climate forcing and a Budyko-model constrained with ET observations from several long-term lysimeter stations. Based on the model results, it was shown that land use changes have had net impacts on evapotranspiration that are generally comparable in size to those caused by changes in precipitation and potential evapotranspiration. Evapotranspiration increased in response to land use (mainly large-scale re- and afforestation) and climate change in most of Europe, with the Iberian pensinsula and other small ~~part of the Mediterrenean~~ parts of the Mediterranean being exceptions with negative trends. Streamflow changes were dominated by a strong positive contribution of precipitation increases in ~~Northern~~ northern Europe. Land use and potential evapotranspiration had smaller effects of opposite sign, resulting in small net streamflow changes over Europe. The analysis revealed considerable complexity at smaller scales, with most of the possible combinations between positive and negative contributions of precipitation, land use, and potential evapotranspiration occuring at some locations. This was true for effects on evapotranspiration and discharge. Most pressing, we find that in much of the ~~Mediterrenean~~Mediterranean, land use and climate change combine to further reduce ~~blue water fluxes~~streamflow and water availability.

6 ~~Acknowledgements~~

~~We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). We thank Diego Miralles for providing the GLEAM data for validation.~~

Data availability. The HILDA land change dataset is available at https://www.wur.nl/en/Research-Results/Chair-groups/Environmental-Sciences/Laboratory-of-Geo-information-Science-and-Remote-Sensing/Models/Hilda.htm. E-OBS v18 precipitation can be downloaded from https://www.ecad.eu//download/ensembles/download.php. CRU TS v4.02 potential evapotranspiration is available from https://crudata.uea.ac.uk/cru/data/hrg/. All hydroclimatic observations used to constrain the Budyko model and their references are listed in Table 1.

Figure A1. Location of the stations and sites listed in Table 1.

Appendix A: Appendix A

Author contributions. AJT designed the study and wrote the manuscript. EDB carried out the study under supervision of AJT and FJ. RF provided the HILDA data. All authors contributed to the writing and the interpretation of the results.

Competing interests. The authors declare no competing interests.

Acknowledgements. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). We thank Diego Miralles for providing the GLEAM data, and Kerstin Stahl for the streamflow change data used for validation. We thank four anonymous referees and Fleur Verwaal (MSc student Earth & Environment at Wageningen University) for their constructive feedback during the interactive discussion that helped to improve the manuscript.

References

Arnell, N. W.: Uncertainty in the Relationship between Climate Forcing and Hydrological Response in UK Catchments, Hydrology and Earth System Sciences, 15, 897–912, https://doi.org/https://doi.org/10.5194/hess-15-897-2011, 2011.

Bach, A. F., van der Schrier, G., Melsen, L. A., Tank, A. M. G. K., and Teuling, A. J.: Widespread and Accelerated Decrease of Observed Mean and Extreme Snow Depth Over Europe, Geophysical Research Letters, 45, 12,312–12,319, https://doi.org/10.1029/2018GL079799, 2018.

Berghuijs, W. R., Woods, R. A., and Hrachowitz, M.: A Precipitation Shift from Snow towards Rain Leads to a Decrease in Streamflow, Nature Climate Change, 4, 583–586, https://doi.org/10.1038/nclimate2246, 2014.

Berthier, E., Andrieu, H., and Rodriguez, F.: The Rezé Urban Catchments Database, Water Resources Research, 35, 1915–1919, https://doi.org/10.1029/1999WR900053, 1999.

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, https://doi.org/10.1016/0022-1694(82)90117-2, 1982.

Bosmans, J. H. C., van Beek, L. P. H., Sutanudjaja, E. H., and Bierkens, M. F. P.: Hydrological Impacts of Global Land Cover Change and Human Water Use, Hydrology and Earth System Sciences, 21, 5603–5626, https://doi.org/https://doi.org/10.5194/hess-21-5603-2017, 2017.

Boyd, M. J., Bufill, M. C., and Knee, R. M.: Pervious and Impervious Runoff in Urban Catchments, Hydrological Sciences Journal, 38, 463–478, https://doi.org/10.1080/02626669309492699, 1993.

Brauer, C. C., van der Velde, Y., Teuling, A. J., and Uijlenhoet, R.: The Hupsel Brook Catchment : Insights from Five Decades of Lowland Observations, Vadose Zone Journal, 17, https://doi.org/10.2136/vzj2018.03.0056, 2018.

Breuer, L., Huisman, J. A., Willems, P., Bormann, H., Bronstert, A., Croke, B. F. W., Frede, H. G., Gräff, T., Hubrechts, L., Jakeman, A. J., Kite, G., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindström, G., Seibert, J., Sivapalan, M., and Viney, N. R.: Assessing the Impact of Land Use Change on Hydrology by Ensemble Modeling (LUCHEM). I: Model Intercomparison with Current Land Use, Advances in Water Resources, 32, 129–146, https://doi.org/10.1016/j.advwatres.2008.10.003, 2009.

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, https://doi.org/10.1016/j.jhydrol.2004.12.010, 2005.

Budyko, M. I.: Climate and Life, Academic Press, New York, 1974.

Calder, I. R.: The Measurement of Water Losses from a Forested Area Using a “Natural” Lysimeter, Journal of Hydrology, 30, 311–325, https://doi.org/10.1016/0022-1694(76)90115-3, 1976.

Caloiero, T., Caloiero, P., and Frustaci, F.: Long-Term Precipitation Trend Analysis in Europe and in the Mediterranean Basin, Water and Environment Journal, 32, 433–445, https://doi.org/10.1111/wej.12346, 2018.

Choudhury, B.: Evaluation of an Empirical Equation for Annual Evaporation Using Field Observations and Results from a Biophysical Model, Journal of Hydrology, 216, 99–110, https://doi.org/10.1016/S0022-1694(98)00293-5, 1999.

Christen, A. and Vogt, R.: Energy and Radiation Balance of a Central European City, International Journal of Climatology, 24, 1395–1421, https://doi.org/10.1002/joc.1074, 2004.

Creed, I. F., Spargo, A. T., Jones, J. A., Buttle, J. M., Adams, M. B., Beall, F. D., Booth, E. G., Campbell, J. L., Clow, D., Elder, K., Green, M. B., Grimm, N. B., Miniat, C., Ramlal, P., Saha, A., Sebestyen, S., Spittlehouse, D., Sterling, S., Williams, M. W., Winkler, R., and Yao, H.: Changing Forest Water Yields in Response to Climate Warming: Results from Long-Term Experimental Watershed Sites across North America, Global Change Biology, 20, 3191–3208, https://doi.org/10.1111/gcb.12615, 2014.

DeWalle, D. R., Swistock, B. R., Johnson, T. E., and McGuire, K. J.: Potential Effects of Climate Change and Urbanization on Mean Annual Streamflow in the United States, Water Resources Research, 36, 2655–2664, https://doi.org/10.1029/2000WR900134, 2000.

Dey, P. and Mishra, A.: Separating the Impacts of Climate Change and Human Activities on Streamflow: A Review of Methodologies and Critical Assumptions, Journal of Hydrology, 548, 278–290, https://doi.org/10.1016/j.jhydrol.2017.03.014, 2017.

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, 11, 983–995, https://doi.org/https://doi.org/10.5194/hess-11-983-2007, 2007.

Duethmann, D. and Blöschl, G.: Why Has Catchment Evaporation Increased in the Past 40 Years? A Data-Based Study in Austria, Hydrology and Earth System Sciences, 22, 5143–5158, https://doi.org/https://doi.org/10.5194/hess-22-5143-2018, 2018.

Dwarakish, G. S. and Ganasri, B. P.: Impact of Land Use Change on Hydrological Systems: A Review of Current Modeling Approaches, Cogent Geoscience, 1, 1115 691, https://doi.org/10.1080/23312041.2015.1115691, 2015.

Ehinger, J.: Siting and Exposure of Meteorological Instruments, WMO/TD-No. 589 Instruments and Observing Methods Report 55, World Meteorological Organization, Geneva, 1993.

Ericsson, S., Östlund, L., and Axelsson, A.-L.: A Forest of Grazing and Logging: Deforestation and Reforestation History of a Boreal Landscape in Central Sweden, New Forests, 19, 227–240, https://doi.org/10.1023/A:1006673312465, 2000.

Farley, K. A., Jobbágy, E. G., and Jackson, R. B.: Effects of Afforestation on Water Yield: A Global Synthesis with Implications for Policy, Global Change Biology, 11, 1565–1576, https://doi.org/10.1111/j.1365-2486.2005.01011.x, 2005.

Filoso, S., Bezerra, M. O., Weiss, K. C. B., and Palmer, M. A.: Impacts of Forest Restoration on Water Yield: A Systematic Review, PLOS ONE, 12, e0183 210, https://doi.org/10.1371/journal.pone.0183210, 2017.

Forzieri, G., Feyen, L., Rojas, R., Flörke, M., Wimmer, F., and Bianchi, A.: Ensemble Projections of Future Streamflow Droughts in Europe, Hydrology and Earth System Sciences, 18, 85–108, https://doi.org/https://doi.org/10.5194/hess-18-85-2014, 2014.

Frenne, P. D. and Verheyen, K.: Weather Stations Lack Forest Data, Science, 351, 234–234, https://doi.org/10.1126/science.351.6270.234-a, 2016.

Fuchs, R., Herold, M., Verburg, P. H., and Clevers, J. G. P. W.: A High-Resolution and Harmonized Model Approach for Reconstructing and Analysing Historic Land Changes in Europe, Biogeosciences, 10, 1543–1559, https://doi.org/10.5194/bg-10-1543-2013, 2013.

Fuchs, R., Herold, M., Verburg, P. H., Clevers, J. G. P. W., and Eberle, J.: Gross Changes in Reconstructions of Historic Land Cover/Use for Europe between 1900 and 2010, Global Change Biology, 21, 299–313, https://doi.org/10.1111/gcb.12714, 2015a.

Fuchs, R., Verburg, P. H., Clevers, J. G. P. W., and Herold, M.: The Potential of Old Maps and Encyclopaedias for Reconstructing Historic European Land Cover/Use Change, Applied Geography, Complete, 43–55, https://doi.org/10.1016/j.apgeog.2015.02.013, 2015b.

Gardner, L. R.: Assessing the Effect of Climate Change on Mean Annual Runoff, Journal of Hydrology, 379, 351–359, https://doi.org/10.1016/j.jhydrol.2009.10.021, 2009.

Gash, J. H. C., Wright, I. R., and Lloyd, C. R.: Comparative Estimates of Interception Loss from Three Coniferous Forests in Great Britain, Journal of Hydrology, 48, 89–105, https://doi.org/10.1016/0022-1694(80)90068-2, 1980.

Gedney, N., Huntingford, C., Weedon, G. P., Bellouin, N., Boucher, O., and Cox, P. M.: Detection of Solar Dimming and Brightening Effects on Northern Hemisphere River Flow, Nature Geoscience, 7, 796–800, https://doi.org/10.1038/ngeo2263, 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, https://doi.org/10.1029/2012GL053492, 2012.

Gerrits, A. M. J., Savenije, H. H. G., Veling, E. J. M., and Pfister, L.: Analytical Derivation of the Budyko Curve Based on Rainfall Characteristics and a Simple Evaporation Model, Water Resources Research, 45, https://doi.org/10.1029/2008WR007308, 2009.

Gerten, D., Rost, S., von Bloh, W., and Lucht, W.: Causes of Change in 20th Century Global River Discharge, Geophysical Research Letters, 35, https://doi.org/10.1029/2008GL035258, 2008.

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., and Seneviratne, S. I.: Global Assessment of Trends in Wetting and Drying over Land, Nature Geoscience, 7, 716–721, https://doi.org/10.1038/ngeo2247, 2014.

Gudmundsson, L., Tallaksen, L. M., Stahl, K., and Fleig, A. K.: Low-Frequency Variability of European Runoff, Hydrology and Earth System Sciences, 15, 2853–2869, https://doi.org/https://doi.org/10.5194/hess-15-2853-2011, 2011.

Haferkorn, U. and Knappe, S.: Long-Term Lysimeter Experiments - Findings from Brandis Lysimeter Station 1st Communication, Archives of Agronomy and Soil Science, 48, 485–491, https://doi.org/10.1080/03650340215651, 2002.

Hall, J., Arheimer, B., Borga, M., Brázdil, R., Claps, P., Kiss, A., Kjeldsen, T. R., Kriaučiūnienė, J., Kundzewicz, Z. W., Lang, M., Llasat, M. C., Macdonald, N., McIntyre, N., Mediero, L., Merz, B., Merz, R., Molnar, P., Montanari, A., Neuhold, C., Parajka, J., Perdigão, R. a. P., Plavcová, L., Rogger, M., Salinas, J. L., Sauquet, E., Schär, C., Szolgay, J., Viglione, A., and Blöschl, G.: Understanding Flood Regime Changes in Europe: A State-of-the-Art Assessment, Hydrology and Earth System Sciences, 18, 2735–2772, https://doi.org/https://doi.org/10.5194/hess-18-2735-2014, 2014.

Hannaford, J.: Climate-Driven Changes in UK River Flows: A Review of the Evidence, Progress in Physical Geography: Earth and Environment, 39, 29–48, https://doi.org/10.1177/0309133314536755, 2015.

Hannaford, J., Buys, G., Stahl, K., and Tallaksen, L. M.: The Influence of Decadal-Scale Variability on Trends in Long European Streamflow Records, Hydrology and Earth System Sciences, 17, 2717–2733, https://doi.org/https://doi.org/10.5194/hess-17-2717-2013, 2013.

Harrigan, S., Murphy, C., Hall, J., Wilby, R. L., and Sweeney, J.: Attribution of Detected Changes in Streamflow Using Multiple Working Hypotheses, Hydrology and Earth System Sciences, 18, 1935–1952, https://doi.org/https://doi.org/10.5194/hess-18-1935-2014, 2014.

Harris, I., Jones, P. D., Osborn, T. J., and Lister, D. H.: Updated High-Resolution Grids of Monthly Climatic Observations – the CRU TS3.10 Dataset, International Journal of Climatology, 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Harsch, N., Brandenburg, M., and Klemm, O.: Large-Scale Lysimeter Site St. Arnold, Germany: Analysis of 40 Years of Precipitation, Leachate and Evapotranspiration, Hydrology and Earth System Sciences, 13, 305–317, https://doi.org/https://doi.org/10.5194/hess-13-305-2009, 2009.

Haylock, M. R., Hofstra, N., Tank, A. M. G. K., Klok, E. J., Jones, P. D., and New, M.: A European Daily High-Resolution Gridded Data Set of Surface Temperature and Precipitation for 1950–2006, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2008JD010201, 2008.

Hurkmans, R. T. W. L., Terink, W., Uijlenhoet, R., Moors, E. J., Troch, P. A., and Verburg, P. H.: Effects of Land Use Changes on Streamflow Generation in the Rhine Basin, Water Resources Research, 45, https://doi.org/10.1029/2008WR007574, 2009.

Istanbulluoglu, E., Wang, T., Wright, O. M., and Lenters, J. D.: Interpretation of Hydrologic Trends from a Water Balance Perspective: The Role of Groundwater Storage in the Budyko Hypothesis, Water Resources Research, 48, https://doi.org/10.1029/2010WR010100, 2012.

Jacobs, C., Elbers, J., Brolsma, R., Hartogensis, O., Moors, E., Rodríguez-Carretero Márquez, M. T., and van Hove, B.: Assessment of Evaporative Water Loss from Dutch Cities, Building and Environment, 83, 27–38, https://doi.org/10.1016/j.buildenv.2014.07.005, 2015.

Jaramillo, F. and Destouni, G.: Local Flow Regulation and Irrigation Raise Global Human Water Consumption and Footprint, Science, 350, 1248–1251, https://doi.org/10.1126/science.aad1010, 2015.

Jaramillo, F., Cory, N., Arheimer, B., Laudon, H., van der Velde, Y., Hasper, T. B., Teutschbein, C., and Uddling, J.: Dominant Effect of Increasing Forest Biomass on Evapotranspiration: Interpretations of Movement in Budyko Space, Hydrology and Earth System Sciences, 22, 567–580, https://doi.org/https://doi.org/10.5194/hess-22-567-2018, 2018.

Jiang, C., Xiong, L., Wang, D., Liu, P., Guo, S., and Xu, C.-Y.: Separating the Impacts of Climate Change and Human Activities on Runoff Using the Budyko-Type Equations with Time-Varying Parameters, Journal of Hydrology, 522, 326–338, https://doi.org/10.1016/j.jhydrol.2014.12.060, 2015.

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., de Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., Mu, Q., Mueller, B., Oleson, K., Papale, D., Richardson, A. D., Roupsard, O., Running, S., Tomelleri, E., Viovy, N., Weber, U., Williams, C., Wood, E., Zaehle, S., and Zhang, K.: Recent Decline in the Global Land Evapotranspiration Trend Due to Limited Moisture Supply, Nature, 467, 951–954, https://doi.org/10.1038/nature09396, 2010.

Li, D., Pan, M., Cong, Z., Zhang, L., and Wood, E.: Vegetation Control on Water and Energy Balance within the Budyko Framework, Water Resources Research, 49, 969–976, https://doi.org/10.1002/wrcr.20107, 2013.

Liu, Y., Pan, Z., Zhuang, Q., Miralles, D. G., Teuling, A. J., Zhang, T., An, P., Dong, Z., Zhang, J., He, D., Wang, L., Pan, X., Bai, W., and Niyogi, D.: Agriculture Intensifies Soil Moisture Decline in Northern China, Scientific Reports, 5, 11 261, https://doi.org/10.1038/srep11261, 2015.

López-Moreno, J. I., Vicente-Serrano, S. M., Moran-Tejeda, E., Zabalza, J., Lorenzo-Lacruz, J., and García-Ruiz, J. M.: Impact of Climate Evolution and Land Use Changes on Water Yield in the Ebro Basin, Hydrology and Earth System Sciences, 15, 311–322, https://doi.org/https://doi.org/10.5194/hess-15-311-2011, 2011.

Lorenzo-Lacruz, J., Vicente-Serrano, S. M., López-Moreno, J. I., Morán-Tejeda, E., and Zabalza, J.: Recent Trends in Iberian Streamflows (1945–2005), Journal of Hydrology, 414-415, 463–475, https://doi.org/10.1016/j.jhydrol.2011.11.023, 2012.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: Satellite-Based Land Evaporation and Root-Zone Soil Moisture, Geoscientific Model Development, 10, 1903–1925, https://doi.org/https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Melsen, L. A., Addor, N., Mizukami, N., Newman, A. J., Torfs, P. J. J. F., Clark, M. P., Uijlenhoet, R., and Teuling, A. J.: Mapping (Dis)Agreement in Hydrologic Projections, Hydrology and Earth System Sciences, 22, 1775–1791, https://doi.org/https://doi.org/10.5194/hess-22-1775-2018, 2018.

Milly, P. C. D., Dunne, K. A., and Vecchia, A. V.: Global Pattern of Trends in Streamflow and Water Availability in a Changing Climate, Nature, 438, 347–350, https://doi.org/10.1038/nature04312, 2005.

Miralles, D. G., van den Berg, M., Gash, J. H., Parinussa, R. M., de Jeu, R. A., Beck, H. E., Holmes, T. R., Jiménez, C., Verhoest, N., Dorigo, W. A., Teuling, A. J., and Dolman, A. J.: El Niño–La Niña Cycle and Recent Trends in Continental Evaporation, Nature Climate Change, 4, 122–126, https://doi.org/http://dx.doi.org/10.1038/nclimate2068, 2014.

Müller, J.: Forestry and Water Budget of the Lowlands in Northeast Germany — Consequences for the Choice of Tree Species and for Forest Management, Journal of Water and Land Development, 13a, 133–148, https://doi.org/10.2478/v10025-010-0024-7, 2009.

Parkin, G., O'Donnell, G., Ewen, J., Bathurst, J. C., O'Connell, P. E., and Lavabre, J.: Validation of Catchment Models for Predicting Land-Use and Climate Change Impacts. 2. Case Study for a Mediterranean Catchment, Journal of Hydrology, 175, 595–613, https://doi.org/10.1016/S0022-1694(96)80027-8, 1996.

Pataki, D. E., McCarthy, H. R., Litvak, E., and Pincetl, S.: Transpiration of Urban Forests in the Los Angeles Metropolitan Area, Ecological Applications: A Publication of the Ecological Society of America, 21, 661–677, 2011.

Piao, S., Friedlingstein, P., Ciais, P., de Noblet-Ducoudré, N., Labat, D., and Zaehle, S.: Changes in Climate and Land Use Have a Larger Direct Impact than Rising CO2 on Global River Runoff Trends, Proceedings of the National Academy of Sciences, 104, 15 242–15 247, https://doi.org/10.1073/pnas.0707213104, 2007.

Pijl, A., Brauer, C. C., Sofia, G., Teuling, A. J., and Tarolli, P.: Hydrologic Impacts of Changing Land Use and Climate in the Veneto Lowlands of Italy, Anthropocene, 22, 20–30, https://doi.org/10.1016/j.ancene.2018.04.001, 2018.

Prudhomme, C. and Williamson, J.: Derivation of RCM-Driven Potential Evapotranspiration for Hydrological Climate Change Impact Analysis in Great Britain: A Comparison of Methods and Associated Uncertainty in Future Projections, Hydrology and Earth System Sciences, 17, 1365–1377, https://doi.org/https://doi.org/10.5194/hess-17-1365-2013, 2013.

Ramamurthy, P. and Bou-Zeid, E.: Contribution of Impervious Surfaces to Urban Evaporation, Water Resources Research, 50, 2889–2902, https://doi.org/10.1002/2013WR013909, 2014.

Ramírez, B. H., Melsen, L. A., Ganzeveld, L., Leemans, R., and Teuling, A. J.: Tropical Montane Cloud Forests in the Orinoco River Basin: Inferring Fog Interception from through-Fall Dynamics, Agricultural and Forest Meteorology, 260-261, 17–30, https://doi.org/10.1016/j.agrformet.2018.05.016, 2018.

Redhead, J. W., Stratford, C., Sharps, K., Jones, L., Ziv, G., Clarke, D., Oliver, T. H., and Bullock, J. M.: Empirical Validation of the InVEST Water Yield Ecosystem Service Model at a National Scale, Science of The Total Environment, 569-570, 1418–1426, https://doi.org/10.1016/j.scitotenv.2016.06.227, 2016.

Renner, M., Brust, K., Schwärzel, K., Volk, M., and Bernhofer, C.: Separating the Effects of Changes in Land Cover and Climate: A Hydro-Meteorological Analysis of the Past 60 Yr in Saxony, Germany, Hydrology and Earth System Sciences, 18, 389–405, https://doi.org/https://doi.org/10.5194/hess-18-389-2014, 2014.

Roderick, M. L. and Farquhar, G. D.: A Simple Framework for Relating Variations in Runoff to Variations in Climatic Conditions and Catchment Properties, Water Resources Research, 47, https://doi.org/10.1029/2010WR009826, 2011.

Samaniego, L., Thober, S., Kumar, R., Wanders, N., Rakovec, O., Pan, M., Zink, M., Sheffield, J., Wood, E. F., and Marx, A.: Anthropogenic Warming Exacerbates European Soil Moisture Droughts, Nature Climate Change, 8, 421, https://doi.org/10.1038/s41558-018-0138-5, 2018.

Seneviratne, S. I., Lehner, I., Gurtz, J., Teuling, A. J., Lang, H., Moser, U., Grebner, D., Menzel, L., Schroff, K., Vitvar, T., and Zappa, M.: Swiss Prealpine Rietholzbach Research Catchment and Lysimeter: 32 Year Time Series and 2003 Drought Event, Water Resources Research, 48, https://doi.org/10.1029/2011WR011749, 2012.

Sheffield, J., Wood, E. F., and Roderick, M. L.: Little Change in Global Drought over the Past 60 Years, Nature, 491, 435–438, https://doi.org/10.1038/nature11575, 2012.

Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., and Smith, D. R.: Impacts of Impervious Surface on Watershed Hydrology: A Review, Urban Water Journal, 2, 263–275, https://doi.org/10.1080/15730620500386529, 2005.

Siebert, S. and Döll, P.: Quantifying Blue and Green Virtual Water Contents in Global Crop Production as Well as Potential Production Losses without Irrigation, Journal of Hydrology, 384, 198–217, https://doi.org/10.1016/j.jhydrol.2009.07.031, 2010.

Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M., and Jódar, J.: Streamflow Trends in Europe: Evidence from a Dataset of near-Natural Catchments, Hydrology and Earth System Sciences, 14, 2367–2382, https://doi.org/https://doi.org/10.5194/hess-14-2367-2010, 2010.

Stahl, K., Tallaksen, L. M., Hannaford, J., and van Lanen, H. A. J.: Filling the White Space on Maps of European Runoff Trends: Estimates from a Multi-Model Ensemble, Hydrology and Earth System Sciences, 16, 2035–2047, https://doi.org/https://doi.org/10.5194/hess-16-2035-2012, 2012.

Sterling, S. M., Ducharne, A., and Polcher, J.: The Impact of Global Land-Cover Change on the Terrestrial Water Cycle, Nature Climate Change, 3, 385–390, 2012.

Stonestrom, D. A., Scanlon, B. R., and Zhang, L.: Introduction to Special Section on Impacts of Land Use Change on Water Resources, Water Resources Research, 45, https://doi.org/10.1029/2009WR007937, 2009.

ter Maat, H. W., Moors, E. J., Hutjes, R. W. A., Holtslag, A. a. M., and Dolman, A. J.: Exploring the Impact of Land Cover and Topography on Rainfall Maxima in the Netherlands, Journal of Hydrometeorology, 14, 524–542, https://doi.org/10.1175/JHM-D-12-036.1, 2013.

Teuling, A. J.: A Forest Evapotranspiration Paradox Investigated Using Lysimeter Data, Vadose Zone Journal, 17, 0, https://doi.org/10.2136/vzj2017.01.0031, 2018.

Teuling, A. J., Hirschi, M., Ohmura, A., Wild, M., Reichstein, M., Ciais, P., Buchmann, N., Ammann, C., Montagnani, L., Richardson, A. D., Wohlfahrt, G., and Seneviratne, S. I.: A Regional Perspective on Trends in Continental Evaporation, Geophysical Research Letters, 36, https://doi.org/10.1029/2008GL036584, 2009.

Teuling, A. J., Seneviratne, S. I., Stöckli, R., Reichstein, M., Moors, E., Ciais, P., Luyssaert, S., van den Hurk, B., Ammann, C., Bernhofer, C., Dellwik, E., Gianelle, D., Gielen, B., Grünwald, T., Klumpp, K., Montagnani, L., Moureaux, C., Sottocornola, M., and Wohlfahrt, G.: Contrasting Response of European Forest and Grassland Energy Exchange to Heatwaves, Nature Geoscience, 3, 722–727, https://doi.org/10.1038/ngeo950, 2010.

Teuling, A. J., Taylor, C. M., Meirink, J. F., Melsen, L. A., Miralles, D. G., van Heerwaarden, C. C., Vautard, R., Stegehuis, A. I., Nabuurs, G.-J., and de Arellano, J. V.-G.: Observational Evidence for Cloud Cover Enhancement over Western European Forests, Nature Communications, 8, 14 065, https://doi.org/10.1038/ncomms14065, 2017.

Theeuwes, N. E., Barlow, J. F., Teuling, A. J., Grimmond, C. S. B., and Kotthaus, S.: Persistent Cloud Cover over Mega-Cities Linked to Surface Heat Release, npj Climate and Atmospheric Science, 2, 15, https://doi.org/10.1038/s41612-019-0072-x, 2019.

Tollenaar, P. and Ryckborst, H.: The Effect of Conifers on the Chemistry and Mass Balance of Two Large Lysimeters in Castricum (The Netherlands), Journal of Hydrology, 24, 77–87, https://doi.org/10.1016/0022-1694(75)90143-2, 1975.

van der Schrier, G., Jones, P. D., and Briffa, K. R.: The Sensitivity of the PDSI to the Thornthwaite and Penman-Monteith Parameterizations for Potential Evapotranspiration, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2010JD015001, 2011.

van der Schrier, G., van den Besselaar, E. J. M., Tank, A. M. G. K., and Verver, G.: Monitoring European Average Temperature Based on the E-OBS Gridded Data Set, Journal of Geophysical Research: Atmospheres, 118, 5120–5135, https://doi.org/10.1002/jgrd.50444, 2013.

van der Velde, Y., Lyon, S. W., and Destouni, G.: Data-Driven Regionalization of River Discharges and Emergent Land Cover–Evapotranspiration Relationships across Sweden, Journal of Geophysical Research: Atmospheres, 118, 2576–2587, https://doi.org/10.1002/jgrd.50224, 2013.

van Dijk, A. I. J. M. and Keenan, R. J.: Planted Forests and Water in Perspective, Forest Ecology and Management, 251, 1–9, https://doi.org/10.1016/j.foreco.2007.06.010, 2007.

van Dijk, A. I. J. M., Gash, J. H., van Gorsel, E., Blanken, P. D., Cescatti, A., Emmel, C., Gielen, B., Harman, I. N., Kiely, G., Merbold, L., Montagnani, L., Moors, E., Sottocornola, M., Varlagin, A., Williams, C. A., and Wohlfahrt, G.: Rainfall Interception and the Coupled Surface Water and Energy Balance, Agricultural and Forest Meteorology, 214-215, 402–415, https://doi.org/10.1016/j.agrformet.2015.09.006, 2015.

van Roosmalen, L., Sonnenborg, T. O., and Jensen, K. H.: Impact of Climate and Land Use Change on the Hydrology of a Large-Scale Agricultural Catchment, Water Resources Research, 45, https://doi.org/10.1029/2007WR006760, 2009.

Viney, N. R., Bormann, H., Breuer, L., Bronstert, A., Croke, B. F. W., Frede, H., Gräff, T., Hubrechts, L., Huisman, J. A., Jakeman, A. J., Kite, G. W., Lanini, J., Leavesley, G., Lettenmaier, D. P., Lindström, G., Seibert, J., Sivapalan, M., and Willems, P.: Assessing the Impact of Land Use Change on Hydrology by Ensemble Modelling (LUCHEM) II: Ensemble Combinations and Predictions, Advances in Water Resources, 32, 147–158, https://doi.org/10.1016/j.advwatres.2008.05.006, 2009.

Wang, X.: Advances in Separating Effects of Climate Variability and Human Activity on Stream Discharge: An Overview, Advances in Water Resources, 71, 209–218, https://doi.org/10.1016/j.advwatres.2014.06.007, 2014.

Wei, X., Li, Q., Zhang, M., Giles-Hansen, K., Liu, W., Fan, H., Wang, Y., Zhou, G., Piao, S., and Liu, S.: Vegetation Cover—Another Dominant Factor in Determining Global Water Resources in Forested Regions, Global Change Biology, 24, 786–795, https://doi.org/10.1111/gcb.13983, 2018.

Wilby, R. L.: When and Where Might Climate Change Be Detectable in UK River Flows?, Geophysical Research Letters, 33, https://doi.org/10.1029/2006GL027552, 2006.

Wild, M.: Decadal Changes in Radiative Fluxes at Land and Ocean Surfaces and Their Relevance for Global Warming, Wiley Interdisciplinary Reviews: Climate Change, 7, 91–107, https://doi.org/10.1002/wcc.372, 2016.

Williams, C. A., Reichstein, M., Buchmann, N., Baldocchi, D., Beer, C., Schwalm, C., Wohlfahrt, G., Hasler, N., Bernhofer, C., Foken, T., Papale, D., Schymanski, S., and Schaefer, K.: Climate and Vegetation Controls on the Surface Water Balance: Synthesis of Evapotranspiration Measured across a Global Network of Flux Towers, Water Resources Research, 48, https://doi.org/10.1029/2011WR011586, 2012.

Xu, C.-Y. and Chen, D.: Comparison of seven models for estimation of evapotranspiration and groundwater recharge using lysimeter measurement data in Germany, Hydrological Processes, 19, 3717–3734, https://doi.org/10.1002/hyp.5853, 2005.

Xu, X., Liu, W., Scanlon, B. R., Zhang, L., and Pan, M.: Local and Global Factors Controlling Water-Energy Balances within the Budyko Framework, Geophysical Research Letters, 40, 6123–6129, https://doi.org/10.1002/2013GL058324, 2013.

Xu, X., Yang, D., Yang, H., and Lei, H.: Attribution Analysis Based on the Budyko Hypothesis for Detecting the Dominant Cause of Runoff Decline in Haihe Basin, Journal of Hydrology, 510, 530–540, https://doi.org/10.1016/j.jhydrol.2013.12.052, 2014.

Zhang, L., Dawes, W. R., and Walker, G. R.: Response of Mean Annual Evapotranspiration to Vegetation Changes at Catchment Scale, Water Resources Research, 37, 701–708, https://doi.org/10.1029/2000WR900325, 2001.

Zhang, L., Hickel, K., Dawes, W. R., Chiew, F. H. S., Western, A. W., and Briggs, P. R.: A Rational Function Approach for Estimating Mean Annual Evapotranspiration, Water Resources Research, 40, https://doi.org/10.1029/2003WR002710, 2004.

Zhang, L., Potter, N., Hickel, K., Zhang, Y., and Shao, Q.: Water Balance Modeling over Variable Time Scales Based on the Budyko Framework – Model Development and Testing, Journal of Hydrology, 360, 117–131, https://doi.org/10.1016/j.jhydrol.2008.07.021, 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, https://doi.org/10.1002/2015GL066952, 2016.

Zimmermann, L., Frühauf, C., and Bernhofer, C.: The Role of Interception in the Water Budget of Spruce Stands in the Eastern Ore Mountains/Germany, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 24, 809–812, https://doi.org/10.1016/S1464-1909(99)00085-4, 1999.

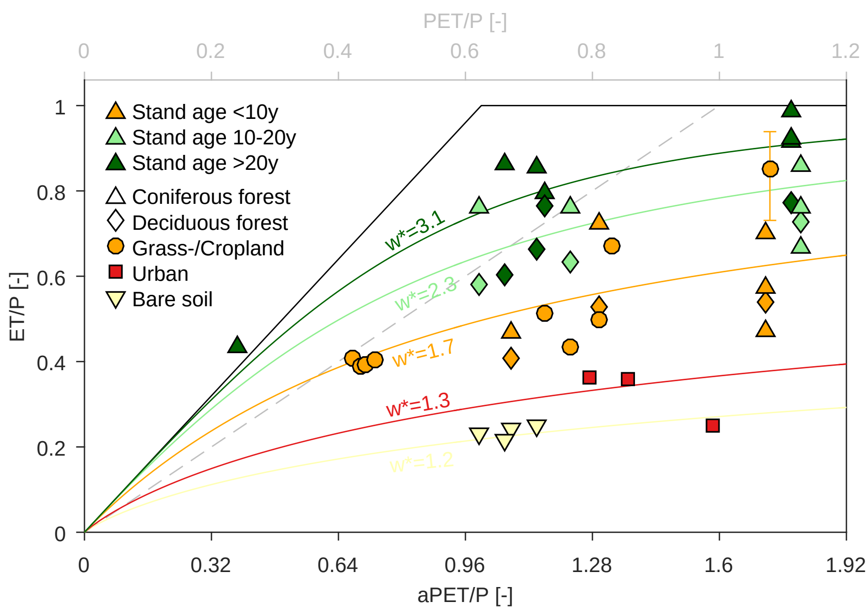


Figure 2. Climate and land use controls on water balance partitioning from long-term flux observations. See Table 1 for origin of data points. The errorbar indicates the total spread over multiple lysimeters at the Brandis site with different soil types (Haferkorn and Knappe, 2002). Curves are based on Eq. 3. Note that symbol shape indicates land cover, but colors indicate stand age in case of forest. The dashed grey line indicates the energy limit for non-adjusted PET.

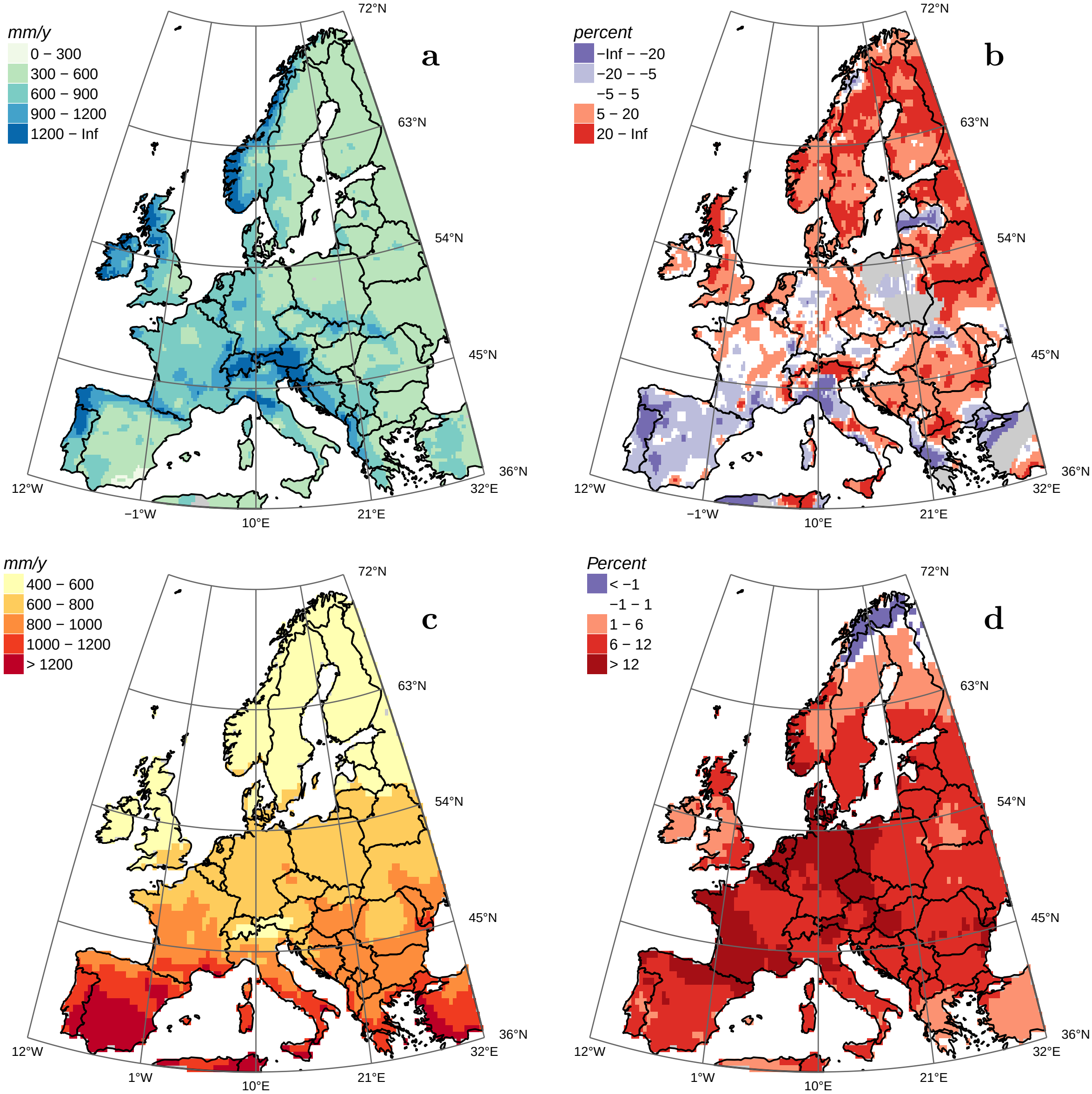


Figure 3. Climate characteristics over the period 1960–2010. Left panels show the mean precipitation and potential evapotranspiration (**a** and **c**, respectively), while the right panels indicate the change over the period 1960–2010 for precipitation (**b**) and potential evapotranspiration (**d**).

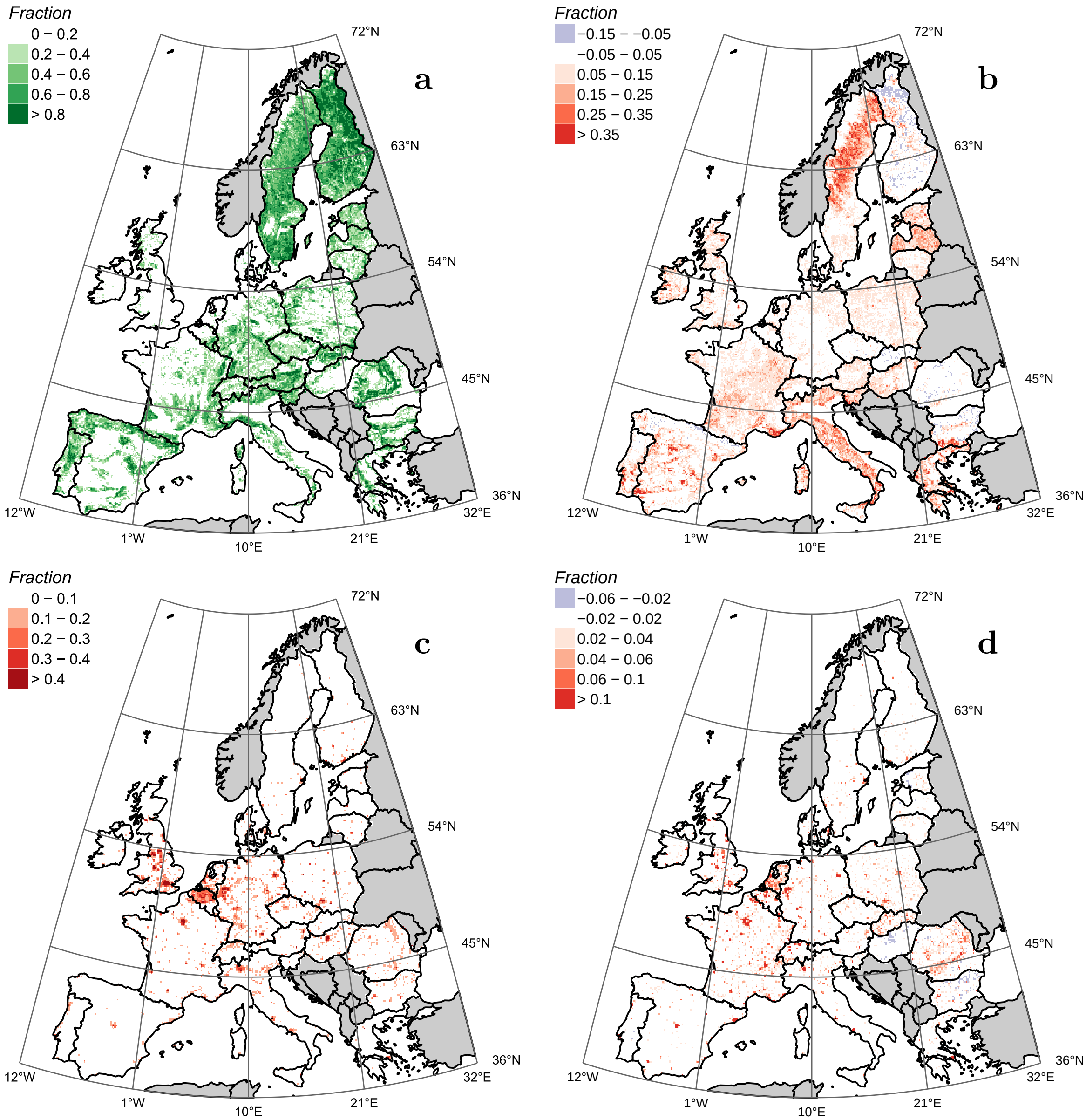


Figure 4. Land cover characteristics over the period 1960–2010. Left panels show the mean forest cover and urban fraction in 1960 (**a** and **c**, respectively), while the right panels indicate the change between the periods 1960 and 2010 for forest cover (**b**) and urban area (**d**).

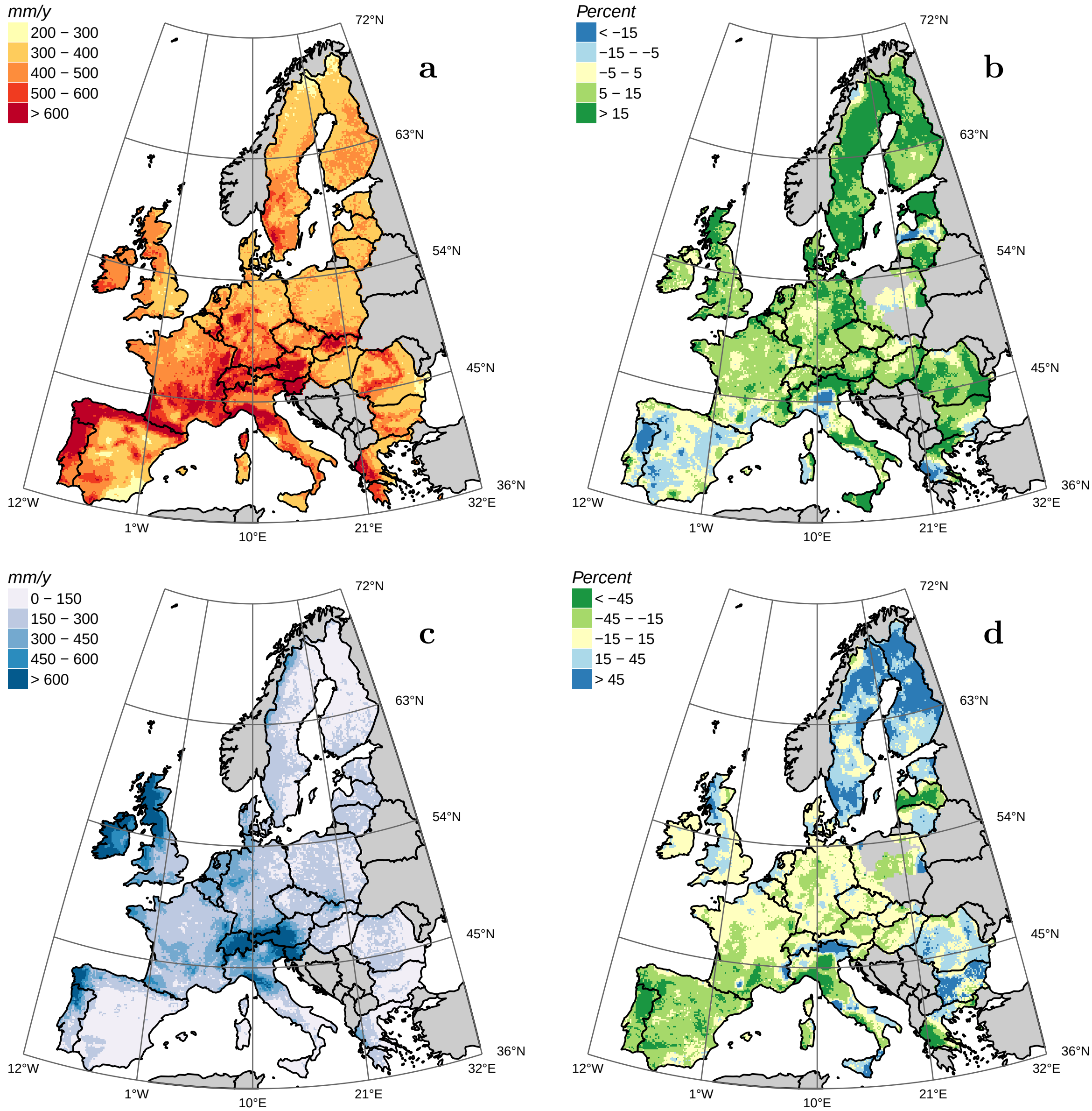


Figure 5. Simulated water balance partitioning over the period 1960–2010. Left panels show the mean evapotranspiration ~~or green water flux~~ and streamflow ~~or blue water flux~~ in 1960 (**a** and **c**, respectively), while the right panels indicate the change between the periods 1960 and 2010 for evapotranspiration (**b**) and streamflow (**d**).~~The colors in the right panels have been chosen such that an increase in blue/green water flux also is shown in blue/green, although it should be noted that a decrease in one flux does not directly translate into an increase in the other.~~

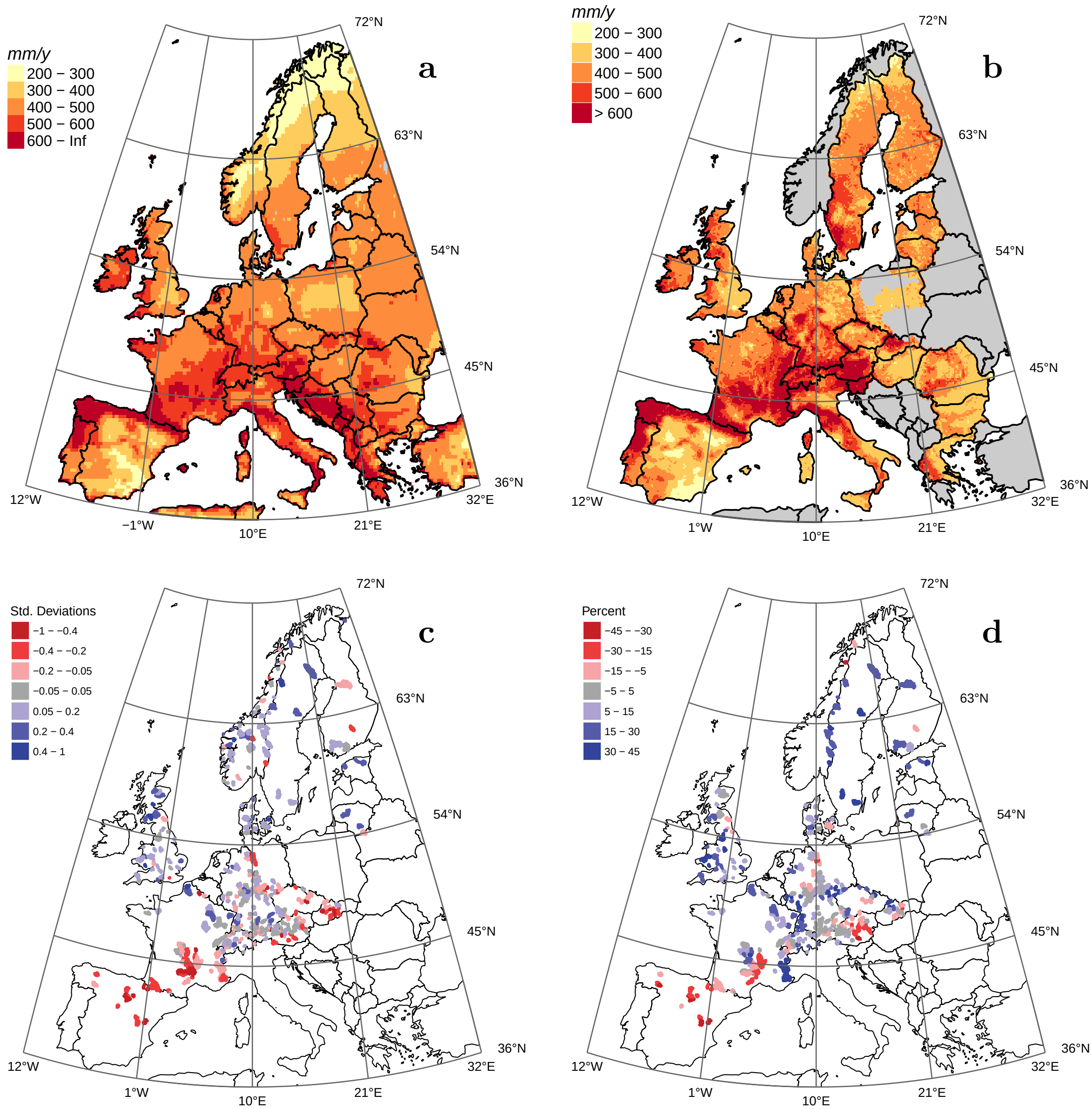


Figure 6. Validation of simulated hydrological fluxes across Europe. **a** Simulated ET average over the 10-year periods 1990, 2000, and 2010. **b** Observation-based ET average over the period 1980–2017 from GLEAM version 3.1 (Martens et al., 2017). ~~**b**~~ **c** Simulated ~~ET average over~~ changes in streamflow between the ~~10-year~~ periods ~~1990, 2000,~~ 1960 and ~~2010. **c**~~ 2000. **d** Observed changes in streamflow over the period 1962–2004 taken from Stahl et al. (2010, their Figure 2). ~~**d** Simulated changes in streamflow between the periods 1960 and 2000.~~ Note the difference in units between ~~observations~~ simulations (**c**) and ~~simulations~~ observations (**d**) because the approach followed in this study does not allow for normalisation by interannual streamflow variability. Observed trends might also be calculated for shorter periods within the period 1962–2004. It should also be noted that ET validation is done for the mean flux, whereas streamflow is validated on the rate of change rather than the mean.

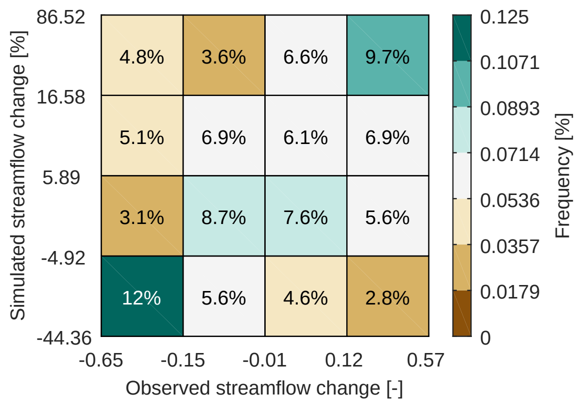


Figure 7. ~~Distribution~~ Two-dimensional quartile distribution of ~~absolute contribution of climate (P and PET) and land use (LU) changes on changes in evapotranspiration over the period 1960–2010~~observed versus simulated streamflow. ~~Colors reflect the relative importance of land use (LU, in magenta or RGB 0,255,255), precipitation (P, in cyan or 255,0,255), and PET (yellow, 255,255,0). Each contribution is inversely scaled between the 2nd and 98th percentiles to reflect its relative importance over Europe. As a result, white (255,255,255) indicates locations where all contributions~~ Note that observed streamflow changes are ~~below their 2nd percentile~~normalized by interannual streamflow variability, ~~and black (0,0,0) indicates locations where all contributions~~ whereas simulated changes are ~~above~~ normalized by their ~~98th percentile~~1960s values.~~Side panels show the absolute contribution of LU, P, and PET and the net change for selected regions. **a** Southern Highlands (Scotland), **b** Paris metropolitan area (France), **c** Landes forest region (France), **d** Seville region (Spain), **e** Central Sweden, **f** Southern Sweden, **g** Styria region (Austria), **h** Smolyan Province (Bulgaria). Domain-averages are listed in Table 2.~~

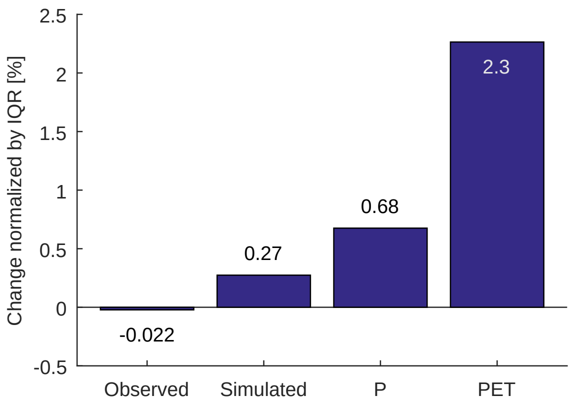


Figure 8. ~~Distribution~~ Comparison of ~~absolute contribution of climate~~ median change normalized by the Inter Quartile Range (~~P and PET~~IQR) for observed and ~~land use (LU) changes on changes in~~ simulated streamflow ~~over the period 1960–2010~~and climate forcing. ~~See caption Fig. 5~~ Normalization was done in order to allow for ~~explanation~~ direct comparison of ~~colors. Side panels show~~ the ~~absolute contribution of LU~~changes reported by Stahl et al. (2010), ~~P~~who reported change normalised by interannual streamflow variability, and ~~PET and the net~~ other fluxes with change ~~for selected regions. **a** Southern Highlands (Scotland), **b** Paris metropolitan area (France), **c** Landes forest region (France), **d** Seville region (Spain), **e** Central Sweden, **f** Southern Sweden, **g** Styria region (Austria), **h** Smolyan Province (Bulgaria). Domain-averages are listed~~ expressed in ~~Table 2.~~percentage.

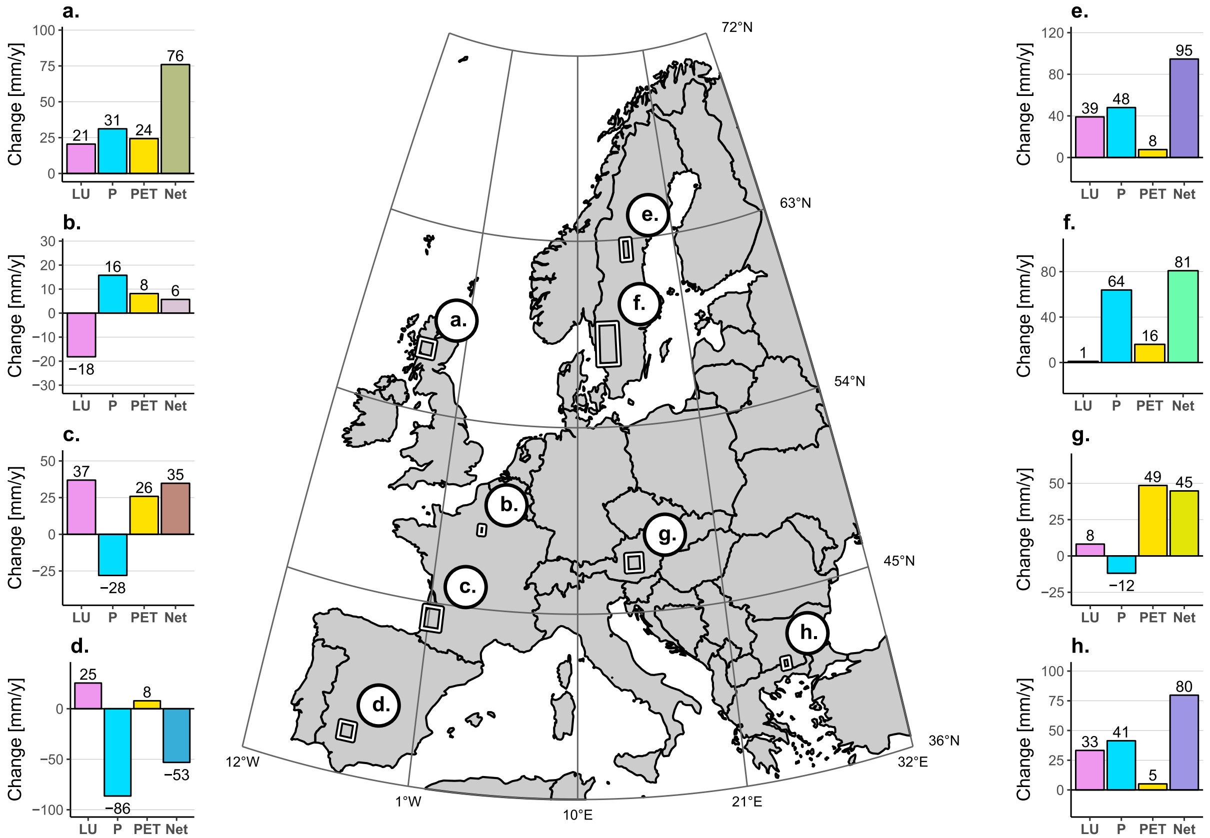


Figure 9. Distribution of absolute contribution of climate (P and PET) and land use (LU) changes on changes in evapotranspiration over the period 1960–2010. Colors reflect the relative importance of land use (LU, in magenta or RGB 0,255,255), precipitation (P, in cyan or 255,0,255), and PET (yellow, 255,255,0). Each contribution is inversely scaled between the 2nd and 98th percentiles over Europe to reflect its relative importance. As a result, white (255,255,255) indicates locations where all contributions are below their 2nd percentile, and black (0,0,0) indicates locations where all contributions are above their 98th percentile. Side panels show the absolute contribution of LU, P, and PET and the net change for selected regions. **a** Southern Highlands (Scotland), **b** Paris metropolitan area (France), **c** Landes forest region (France), **d** Seville region (Spain), **e** Central Sweden, **f** Southern Sweden, **g** Styria region (Austria), **h** Smolyan Province (Bulgaria). Domain-averages are listed in Table 2.

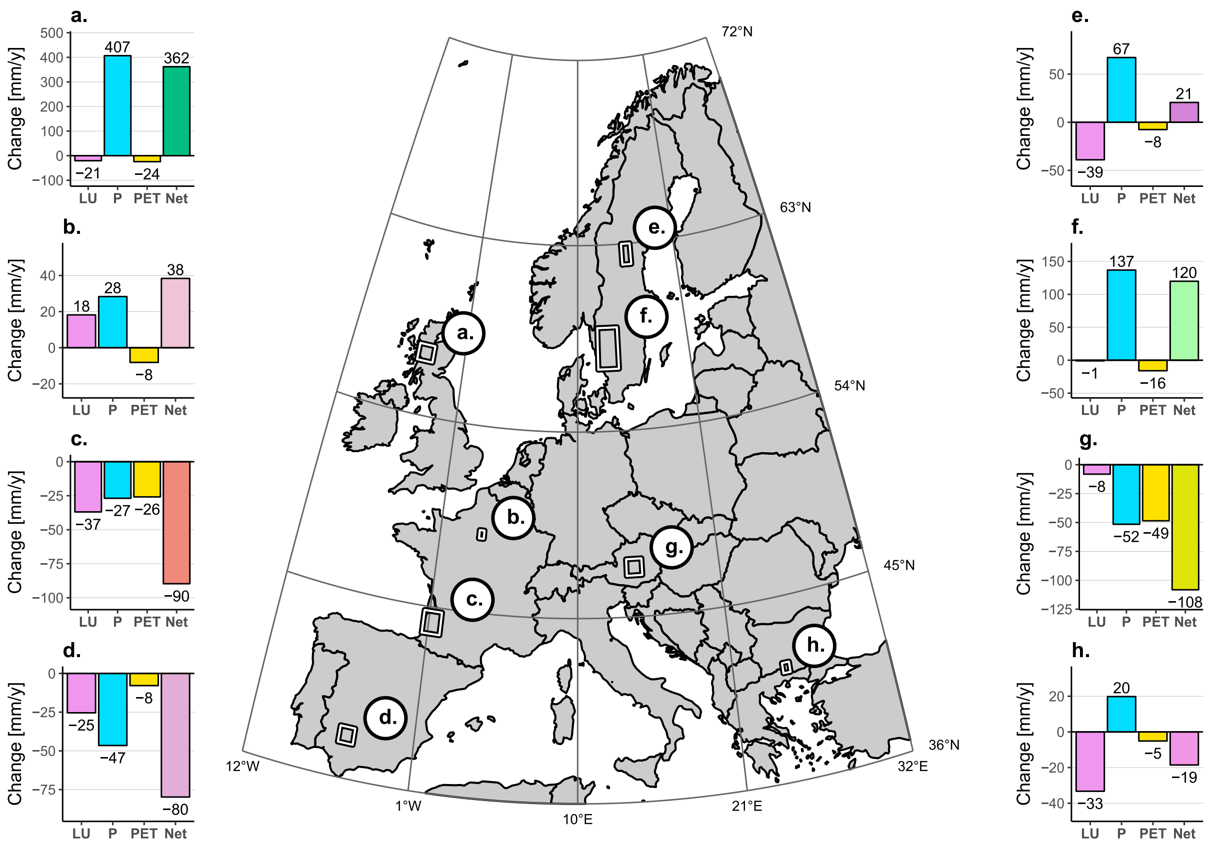


Figure 10. Distribution of absolute contribution of climate (P and PET) and land use (LU) changes on changes in streamflow over the period 1960–2010. See caption Fig. 8 for explanation of colors. Side panels show the absolute contribution of LU, P, and PET and the net change for selected regions. **a** Southern Highlands (Scotland), **b** Paris metropolitan area (France), **c** Landes forest region (France), **d** Seville region (Spain), **e** Central Sweden, **f** Southern Sweden, **g** Styria region (Austria), **h** Smolyan Province (Bulgaria). Domain-averages are listed in Table 2.

Table 1. Data used in the Budyko-analysis. Units of fluxes are in mm/y.

Site	Lat.	Lon.	Land use	Period	P	PET*	ET	Site reference/Source
Castricum	52.55	4.64	Bare soil	1941–1952	825	554	201	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Bare soil	1957–1966	893	554	205	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Bare soil	1972–1981	805	574	202	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Bare soil	1987–1996	887	588	192	Tollenaar and Ryckborst (1975)
Arnhem	51.98	5.91	Urban	2012–2013	781	668	281	Jacobs et al. (2015)
Basel	47.57	7.59	Urban	2001–2002	800	660	300	Christen and Vogt (2004)
Rotterdam	51.93	4.47	Urban	2012	700	693	175	Jacobs et al. (2015)
St. Arnold	52.21	7.39	Grassland	1969–1978	687	558	343	Harsch et al. (2009)
St. Arnold	52.21	7.39	Grassland	1982–1991	765	585	332	Harsch et al. (2009)
St. Arnold	52.21	7.39	Grassland	1995–2004	834	604	427	Harsch et al. (2009)
Brandis**	51.53	12.10	Cropland	1981–1994	654	706	556	Haferkorn and Knappe (2002)
Rheindahlen	51.14	6.37	Grassland	1983–1994	795	660	532	Xu and Chen (2005)
Rietholzbach	47.38	8.99	Grassland	1976–1985	1416	598	573	Seneviratne et al. (2012)
Rietholzbach	47.38	8.99	Grassland	1986–1995	1456	633	559	Seneviratne et al. (2012)
Rietholzbach	47.38	8.99	Grassland	1996–2005	1430	634	543	Seneviratne et al. (2012)
Rietholzbach	47.38	8.99	Grassland	2006–2015	1449	664	583	Seneviratne et al. (2012)
St. Arnold	52.21	7.39	Forest (coniferous)	1969–1978	687	558	497	Harsch et al. (2009)
St. Arnold	52.21	7.39	Forest (coniferous)	1982–1991	765	585	582	Harsch et al. (2009)
St. Arnold	52.21	7.39	Forest (coniferous)	1995–2004	834	604	662	Harsch et al. (2009)
St. Arnold	52.21	7.39	Forest (deciduous)	1969–1978	687	558	364	Harsch et al. (2009)
St. Arnold	52.21	7.39	Forest (deciduous)	1982–1991	765	585	485	Harsch et al. (2009)
St. Arnold	52.21	7.39	Forest (deciduous)	1995–2004	834	604	638	Harsch et al. (2009)
Castricum	52.55	4.64	Forest (coniferous)	1941–1952	825	554	386	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Forest (coniferous)	1957–1966	893	554	680	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Forest (coniferous)	1972–1981	805	574	688	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Forest (coniferous)	1987–1996	887	588	764	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Forest (deciduous)	1941–1952	825	554	336	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Forest (deciduous)	1957–1966	893	554	519	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Forest (deciduous)	1972–1981	805	574	533	Tollenaar and Ryckborst (1975)
Castricum	52.55	4.64	Forest (deciduous)	1987–1996	887	588	534	Tollenaar and Ryckborst (1975)
Eberswalde	52.89	13.81	Forest (deciduous)	1978–1984	633	680	341	Müller (2009)

*Derived from E-OBS (P) and CRU (PET).

**Mean of 24 lysimeters listed, minimum value 478 mm/y and maximum 614 mm/y also shown as errorbar in Fig. 1.

***Values digitized from Calder (1976)

Table 2. Data used in the Budyko-analysis (table continued because of double line-spacing). Units of fluxes are in mm/y.

Site	Lat.	Lon.	Land use	Period	P	PET*	ET	Site reference/Source
Eberswalde	52.89	13.81	Forest (deciduous)	1985–1989	625	706	455	Müller (2009)
Eberswalde	52.89	13.81	Forest (deciduous)	1990–1998	633	704	489	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1978–1984	633	680	299	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1985–1989	625	706	417	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1990–1998	633	704	580	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1978–1984	633	680	363	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1985–1989	625	706	476	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1990–1998	633	704	584	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1978–1984	633	680	443	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1985–1989	625	706	537	Müller (2009)
Eberswalde	52.89	13.81	Forest (coniferous)	1990–1998	633	704	625	Müller (2009)
Plynlimon***	52.47	-3.73	Forest (coniferous)	1974–1975	2300	552	999	Calder (1976)

*Derived from E-OBS (P) and CRU (PET).

**Mean of 24 lysimeters listed, minimum value 478 and maximum 614 also shown in Fig. 1.

***Values digitized from Calder (1976)

Table 3. Climate and land use contributions to changes in evapotranspiration and streamflow over the period 1960–2010. All units in km^3 y^{-1}. For reference, 1 km^3 y^{-1} corresponds to an average discharge of 32 m^3 s^{-1}. The total area with available data is 4,312,807 km^2.

	Evapotranspiration			Streamflow		
factor	positive	negative	net	positive	negative	net
			Whole study domain			
Land use	~~53.6~~ 54.0	−~~9.26~~9.0	~~44.3~~ 45.0	~~9.3~~ 9.0	−~~53.6~~ 54.0	−~~44.3~~ 45.0
Precipitation	~~89.2~~92.4	−~~48.2~~58.2	~~41.0~~ 34.4	~~165.3~~ 162.3	−~~76.0~~ 66.0	~~89.3~~ 96.3
Potential evapotranspiration	~~42.1~~ 60.6	−0.2	~~41.9~~ 60.4	0.2	−~~42.1~~ 60.6	−~~41.9~~ 60.4
			Near-natural catchments (Stahl et al., 2010, and Fig. 6c,d)			
Land use	~~2.2~~ 2.4	−0.3	~~2.0~~ 2.1	0.3	−~~2.2~~ 2.4	−~~2.0~~ 2.1
Precipitation	~~2.8~~ 3.0	−~~1.3~~ 1.9	~~1.4~~ 1.2	~~7.4~~ 7.1	−~~3.4~~ 2.9	~~4.0~~ 4.2
Potential evapotranspiration	~~2.6~~ 3.7	0	~~2.6~~ 3.7	0	−~~2.6~~ 3.7	−~~2.6~~ 3.7
			Rhine basin			
Land use	~~2.1~~ 2.2	−0.4	~~1.6~~ 1.8	0.4	−~~2.1~~ 2.2	−~~1.6~~ 1.8
Precipitation	~~1.6~~ 1.8	−~~0.9~~ 1.1	0.7	~~4.1~~ 3.9	−~~2.8~~ 2.6	1.3
Potential evapotranspiration	~~2.7~~ 3.6	0	~~2.7~~ 3.6	0	−~~2.7~~ 3.6	−~~2.7~~ 3.6