Supporting Information for

Key landscape and biotic indicators of watersheds sensitivity to forest disturbance identified using remote sensing and historical hydrography data

Buma B¹, Livneh B²

1. University of Alaska Southeast, Juneau
2. University of Colorado, Boulder

Contents of this file

Supplementary Methods S1
Figures S1 to S4

Additional Supporting Information (Files uploaded separately)

Supplementary Data Table 1

Introduction

The following supplementary information provides additional detail on the runoff ratio detrending methodology, model creation and variable selection, and final grouping outcomes.

Figures S1-S4 provide further elaboration on points in the main manuscript.

Supplementary Data Table 1, provided as an Excel table, lists all the variables and their associated statistical summaries (e.g., mean, median) along with reference to their derivation and the original dataset.
Supplementary Methods S1

There are 671 complete, gaged watersheds [Newman et al., 2015], distributed across the continental United States (CONUS); 228 of those are undammed (as of 2009). For each watershed, the forest area disturbed each year was totaled [Hansen et al., 2013], which is built from MODIS and 30m Landsat imagery. This dataset identifies all forested pixels which are disturbed (regardless of disturbance type) and has been validated in several published studies [Hansen et al., 2013; Buma and Barrett, 2015] with an approximate error rate of 10%. The disturbed area per watershed was calculated for each year (2001-2010), and the year of the largest disturbance was identified; subsequent analyses focused on changes in water yield in the water year (Oct. – Sept.) following that largest event.

The initial metric for water yield response was the runoff ratio: the fraction of precipitation that becomes streamflow, observed at the USGS gauge. Post-disturbance change in runoff ratio was chosen (as opposed to streamflow itself, e.g., m³/s) because runoff ratio is normalized by climate enabling comparison across basins. However, as a second order effect, the runoff ratio is influenced by inter-annual precipitation variability [Karl and Riebsame, 1989; Vano et al., 2012], whereby the efficiency of runoff production may be disproportionately high in a wet year as opposed to a dry year. This could create unexplained response if the year following the disturbance was abnormally dry or wet. To control for second order precipitation effects, observed yearly runoff ratio values were detrended (via a linear model) with respect to annual precipitation for the years prior to the largest disturbance event. This model was then used to create a predicted runoff ratio for the year following the major disturbance using observed precipitation for that year. Importantly, this prediction represents an estimate of the expected runoff ratio as if the disturbance had not occurred. The predicted runoff ratio was then subtracted from the observed (observed – predicted), giving the detrended

Figure M1. Flowchart of datasets and processing steps used in the analysis.
residual runoff ratio ($dRRR$) for the water year after the disturbance. This was done for each watershed independently (Fig. M1). In summary, this new metric is a measure of how much water yield deviates from expected, given observed precipitation. For example, if the detrended residual runoff ratio (post-disturbance) was positive, it would indicate that yield increased for a given precipitation amount (Fig. 1).

To quantify runoff timing, the centroid of observed runoff was calculated. The centroid is day of the year on which 50% of the annual streamflow has passed the gauge. To control for differences in centroid timing driven by differences in watershed characteristics (e.g., cold, snowmelt watersheds vs. monsoonal watersheds) we again applied the same detrending methodology as detailed above. This has the effect of controlling for between basin variation by isolated deviance from expected timing, rather than deviance from absolute timing.

Watersheds were then grouped into three categories for statistical analyses by percentiles. Residual runoff ratios were grouped into the lowest 25% (“Group 1”: negative residual indicating less water yield than expected given precipitation; for timing differences negative indicates earlier centroid timing), the middle 25-75% (“Group 2”: residual near zero, indicating little change post-disturbance), and the top 75% (“Group 3”: positive residual, indicating greater water yield than expected post-disturbance, or later centroid timing). Finally, watersheds with less than 1% disturbed were excluded from the initial analysis, as they would not be expected to follow similar patterns.

Figure M2. Distribution of watersheds in the GAGES-II database and their proportion forested.

Figure M3. Group responses and LDA results for the yield change analysis. Groups are significantly different ($p < 0.05$). For the explanatory variables used in this classification, see main text.
However, these watersheds were included in the final application of the predictive analysis to provide a broad estimate of expected responses.

**Explanatory Variable Selection:**
The GAGES-II dataset contains extensive descriptive data for each watershed (see Supplementary Data 1 for the 75 utilized in this study). Yet descriptors are highly cross-correlated and potentially irrelevant to the processes of interest. To reduce the number of potential explanatory variables and reduce cross-correlation, relative variable importance was estimated using random forests. Random forests [Brieman 2001] is a widely used statistical tool, often used to determine the importance of variables in datasets with highly correlated explanatory variables.

First, all predictor variables were standardized (mean = 0, standard deviation = 1). Second, a random forest model was applied to classify watersheds into response group (water yield or streamflow timing treated independently) containing all potential predictor variables. A list of variable importance (calculated via reduction in mean squared error (MSE)) was created for the entire dataset. The top 25 most important variables, as determined by their individual reduction in MSE, were retained. If potential variables were cross-correlated ($r > 0.75$), the variable of lower importance was discarded and the more important variable retained. The reasoning to this step was that including multiple, strongly correlated variables is less useful than a set of more independent variables when identifying the characteristics of sensitive vs. insensitive watersheds. The result of this step is the creation of a list of 25 independent watershed variables best able to distinguish between the response groups for both yield and timing.

While useful for complex problems, random forests are difficult to interpret directly. In contrast, linear discriminant analysis (LDA) is a useful technique when the identity of predictor variables needs to be retained for interpretation and group membership is known (as here). An LDA was then applied to determine which variables best describe watershed response at the group level using the top 25 scaled, centered, and uncorrelated variables. Classifications as a function of the first two LDA axes are shown in Fig. M3 (yield change) and Fig. M4 (timing change).

**Cited:**

![Figure M4. Group responses and LDA results for the timing change analysis. Groups are significantly different ($p < 0.05$). For the most explanatory variables used in this classification, see main text.](image)


Figure S1. (a) Percent forest loss in the largest disturbance event (2001-2010) and the original percent forested for undammed watersheds with a range of disturbance percentages (228 watersheds), (b) post-disturbance response group distribution in terms of deviations in flow timing, and (c) deviations in water yield. Boxplots show 25th and 75th percentiles, whiskers extent to 1.5 the interquartile range beyond those percentiles. Outliers shown as points. See text for information on groups.
Figure S2. Proportional distribution of cover types within the undammed watersheds used in the study (n=228). Cover types are from the National Landcover Database 2006 dataset, 30m resolution.
Figure S3. Proportion plot comparing watershed response in terms of yield and timing. Watersheds with increased water yield post-disturbance (higher runoff ratio than expected given precipitation inputs) were disproportionally likely to also have yield occur earlier in the year. The differences in proportions are significant (p < 0.05).
Figure S4. Distribution of two edaphic variables in the US West, from Washington and going to California: soil contact time (days) and soil bulk density (grams/cm$^3$). Topography shown in background. Generally low bulk densities are found in the Northwest, but contact time varies considerably as a function of topography, soil type, texture, and other variables. For a full list of edaphic variables and derivation methods and citations, see Supplementary Data Table 1.
Data Set S1. Supplementary Data Table 1 contains all the potential predictor variables considered in the methodology, along with summary statistics for each. It is available as an Excel file. A general description of their creation is given along with reference to the original dataset where specific reference for each variable is available.