

Supplementary Materials for

Managing salinity in upper Colorado River basin streams: Selecting catchments for sediment control efforts using watershed characteristics and random forests models

Fred D Tillman^{1*}, David W. Anning², Julian A. Heilman¹, Susan G. Buto³, and Matthew P. Miller⁴

¹ United States Geological Survey, Arizona Water Science Center, Tucson, Arizona, United States of America; jahr@usgs.gov

² United States Geological Survey, Arizona Water Science Center, Flagstaff, Arizona, United States of America; dwaning@usgs.gov

³ United States Geological Survey, Nevada Water Science Center, Carson City, Nevada, United States of America; sbuto@usgs.gov

⁴ United States Geological Survey, Utah Water Science Center, Salt Lake City, Utah, United States of America; mamiller@usgs.gov

* Correspondence: ftillman@usgs.gov; Tel.: 1-520-670-3312

Introduction

The following Supplementary Materials includes a description of watershed characteristics that were investigated for potential contribution to the relation between suspended-sediment and dissolved-solids concentrations in streams of the upper Colorado River basin (UCRB). Text in [brackets] indicates field titles in Supplementary Materials Tables S1, S3, S4, and S6, and Figure S2. Following the description of watershed characteristics are the following two figures:

Figure S1. Out-of-box and SM+ class error rates as a function of the number of trees in initial random forests classification simulations using default argument values

Figure S2. Boxplots showing median and interquartile range (box) and maximum and minimum values not including outliers (whiskers) for standardized watershed characteristic data. * denotes p-value<0.05 for Wilcoxon rank-sum test on SM+ and N classes for that characteristic. See Supplementary Materials Table S4 for all standardized values and Wilcoxon rank-sum results.

Description of watershed characteristics investigated in this study

Topography and area

Attributes: Catchment elevation, slope, and area

Description: Minimum [min_elev], maximum [max_elev], mean [mean_elev], median [med_elev], and range [range_elev] in catchment elevation in meters; mean catchment percent slope [mean_pct_slope]; catchment area in km² [area_km2]

Source and Processing of Geospatial Data: Ten meter (1/3 arc-second) digital elevation model (DEM) data were obtained from the USGS National Map (http://nationalmap.gov/3DEP/3dep_prodserv.html#) and mosaicked together to cover the upper Colorado River basin study area. Catchment elevation and percent slope statistics were computed using the ArcGIS zonal statistics and slope tools.

Land cover and land use

Attributes: Land cover

Description: Fraction of catchment area classified as open water [NLCD_11]; perennial ice/snow [NLCD_12]; developed – open space [NLCD_21]; developed – low intensity [NLCD_22]; developed – medium intensity [NLCD_23]; developed – high intensity [NLCD_24]; barren land [NLCD_31]; deciduous forest [NLCD_41]; evergreen forest [NLCD_42]; mixed forest [NLCD_43]; shrub/scrub [NLCD_52]; grassland/herbaceous [NLCD_71]; pasture/hay [NLCD_81]; cultivated crops [NLCD_82]; woody wetlands [NLCD_90], or emergent herbaceous wetlands [NLCD_95] in the National Land Cover Dataset (NLCD). For descriptions of NLCD categories, see http://www.mrlc.gov/nlcd11_leg.php

Source and Processing of Geospatial Data: The 2011 NLCD dataset for the conterminous United States was downloaded from http://www.mrlc.gov/nlcd11_data.php. The ArcGIS tool “zonal statistics as table” was used to determine the area of each land cover type within each catchment, which was then converted to a fraction of the catchment area in Excel.

Attributes: Irrigation method

Description: Fraction of catchment area irrigated by flood [flood_irr] or sprinkler [sprinkler_irr] methods in 2010.

Source and Processing of Geospatial Data: Dataset derived from Buto et al. (2014). The ArcGIS tool “intersect” was used to determine the area of each irrigation type within each catchment, which was then converted to a fraction of the catchment area in Excel.

Attributes: Rangeland

Description: Fraction of catchment area designated as rangeland using the U.S. Forest Service Forest Inventory and Analysis (FIA) classification [FIA_rangeland].

Source and Processing of Geospatial Data: Dataset described in Reeves and Mitchell (2011) and raster dataset obtained from Matt Reeves of the U.S. Forest Service (personal communication). The ArcGIS tool “zonal statistics as table” was used to determine the area of rangeland within each catchment, which was then converted to a fraction of the catchment area in Excel.

Climate and climate-related

Attributes: Basin Characteristic Model (BCM) input and calculated parameters

Description: Mean total annual value for actual evapotranspiration [BCM_AET], climatic water deficit [BCM_CWD], excess water [BCM_EXC], snowmelt [BCM_MLT], snowpack [BCM_PCK], potential evapotranspiration [BCM_PET], precipitation [BCM_PPT], sublimation [BCM_SBL], and snowfall [BCM_SNW] for each catchment. The BCM groundwater recharge model (Flint and Flint, 2007) estimates groundwater recharge at a basin scale using a water-balance approach. Several climate and climate-related datasets are required for BCM simulations or are computed as part of the water-balance calculations.

Source and Processing of Geospatial Data: BCM parameters for the UCRB averaged over the 1985-2012 time period were obtained from Alan Flint of the USGS (personal communication). The ArcGIS tool “zonal statistics as table” was used to determine the mean total annual value of each attribute for each catchment.

Geology and rock chemistry

Attributes: Lithologic groupings

Description: Fraction of catchment area classified as the following lithologic groupings: crystalline and volcanic rocks [Xline], high-yield sedimentary Cenozoic rocks [HSedCz], low-yield sedimentary Cenozoic rocks [LSedCz], high-yield sedimentary Mesozoic rocks [HSedMz], low-yield sedimentary Mesozoic rocks [LSedMz], high-yield sedimentary Paleozoic and Precambrian rocks [HSedPzPc], and low-yield sedimentary Paleozoic and Precambrian rocks [LSedPzPc].

Source and Processing of Geospatial Data: Geologic units derived from 1:500,000-scale state geologic maps were grouped into 34 defined units based on the 1:2,500,000-scale King and Beikman (1974) geology of the conterminous United States and further aggregated into seven source groups that were used to define the geologic units. Details on the grouping methods and corresponding King and Beikman and 1:500,000-scale state geologic units which comprise each geologic source group are presented in Kenney et al. (2009). The ArcGIS tool “zonal statistics as table” was used to determine the area of each lithologic group within each catchment, which was then converted to a fraction of the catchment area in Excel.

Attributes: Rock chemistry

Description: Area weighted mean for each catchment of the following rock chemistry types: calcium oxide [Rock_CAO], iron oxide [Rock_FE], potassium oxide [Rock_K], magnesium oxide [Rock_MGO], phosphorus [Rock_P], sulfur [Rock_S], and silicon dioxide [Rock_SI]. Area weighted mean of rock hydraulic conductivity [Rock_PERM] and uniaxial compressive strength [UCS].

Source and Processing of Geospatial Data: Rock chemistry data obtained from Geochemical and Geophysical Characteristics of the Conterminous United States (Olson and Hawkins, 2014). The Geospatial Modeling Environment (GME) tool “isectpolypoly” was used to compute the area weighted mean of each of the rock chemistry attributes for each catchment.

Attributes: Subsurface evaporite deposits

Description: Fraction of catchment area underlain by gypsum/anhydrite or halite [hal+gyp-anhyd] and gypsum/anhydrite [gyp-anhyd].

Source and Processing of Geospatial Data: Datasets derived from digitized map in “Evaporite-karst problems and studies in the USA” (Johnson, 2008). The ArcGIS tool “zonal statistics as table” was used to determine the area of each evaporite deposit type within each catchment, which was then converted to a fraction of the catchment area in Excel.

Soils

Attributes: STATSGO soil parameters

Description: Slope gradient of the dominant component in the map unit [slopegraddcp], slope gradient of all components in the map unit [slopegradwta], fraction of area for each hydrologic soil group [HSG_A, HSG_B, HSG_C, HSG_D, HSG_B/D]. For upper soil horizon only and weighted average for all soil horizons: erodability factor [ALLHZkfact and UPHZkfact], horizon thickness [ALLHZhzthk and UPHZhzthk], total clay [ALLHZclaytotal and UPHZclaytotal], total silt [ALLHZsilttotal and UPHZsilttotal], total sand [ALLHZsandtotal and UPHZsandtotal], total organic matter [ALLHZom and UPHZom], saturated hydraulic conductivity [ALLHZksat and UPHZksat], and available water capacity [ALLHZawc and UPHZawc].

Source and Processing of Geospatial Data: The National Resources Conservation Service (NRCS) State Soil Geographic (STATSGO) database was obtained from the USDA Geospatial Data Gateway (<http://datagateway.nrcs.usda.gov/>). Queries were run for the listed attributes and resulting tables joined to state maps. The Geospatial Modeling Environment (GME) tool “isectpolypoly” was used to compute the area weighted mean of each STATSGO parameter for each catchment, except hydrologic soil group for which the catchment fraction was calculated for each HSG value.

Hydrology and water quality

Attributes: Basin Characteristic Model (BCM) calculated parameters

Description: Mean total annual groundwater recharge [BCM_RCH] and runoff [BCM_RUN] for each catchment. The BCM groundwater recharge model (Flint and Flint, 2007) estimates groundwater recharge at a basin scale using a water-balance approach.

Source and Processing of Geospatial Data: BCM output for the UCRB averaged over the 1985-2012 time period were obtained from Alan Flint of the USGS. The ArcGIS tool “zonal statistics as table” was used to determine the mean total annual recharge and runoff for each catchment.

Attributes: Base-flow index

Description: Mean base-flow index [mean_BFI] for each catchment. Base-flow index is the relative amount of baseflow compared with total flow in a stream.

Source and Processing of Geospatial Data: Base-flow index was derived from U.S. Geological Survey Digital Data Series 491-04 (Wieczorek and LaMotte, 2010). The ArcGIS tool “zonal statistics as table” was used to determine the mean base-flow index for each catchment.

Attributes: Presence of saline groundwater

Description: Fraction of catchment area underlain by saline groundwater less than 500 feet below land surface [saline_GW].

Source and Processing of Geospatial Data: Presence of saline groundwater was derived from the georeferenced map in U.S. Geological Survey Fact Sheet 075-03 (Alley, 2003). The ArcGIS tool “zonal statistics as table” was used to determine the area of each catchment underlain by saline groundwater, which was then converted to a fraction of the catchment area in Excel.

Attributes: Rainfall-runoff erosivity factor (R-factor)

Description: Mean rainfall-runoff erosivity factor [Rf30] for each catchment. The R-factor is a measure of the cumulative erosive force of individual precipitation events (see Daly and Taylor, 2002).

Source and Processing of Geospatial Data: This tabular dataset was created by the USGS National Water Quality Assessment Program from the source data of Daly and Taylor (2002). The Geospatial Modeling Environment (GME) tool “isectpolypoly” was used to compute the area weighted mean of the R-factor for each catchment.

Attributes: Specific conductance information

Description: 90th percentile [90thSC] and median [medianSC] specific conductance values at the site in units of microsiemens per centimeter at 25 degrees Celsius.

Source and Processing of Geospatial Data: Specific conductance data were obtained from the USGS National Water Information System (NWIS; <http://waterdata.usgs.gov/nwis>).

Attributes: Streamflow

Description: Mean daily flow at site [mdf_cfs] in cubic feet per second.

Source and Processing of Geospatial Data: Mean daily flow for the sites was obtained from NWIS (<http://waterdata.usgs.gov/nwis>).

References

Alley, William M., 2003, Desalination of ground water—earth science perspectives: U.S. Geological Survey Fact Sheet 075-03, 4 p., <http://pubs.usgs.gov/fs/fs075-03/>

Buto, S.G., Gold, B.L., and Jones, K.A., 2014, Development of a regionally consistent geospatial dataset of agricultural lands in the Upper Colorado River Basin, 2007–10: U.S. Geological Survey Scientific Investigations Report 2014-5039, 20 p., <http://dx.doi.org/10.3133/sir20145039>.

Daly, C., and Taylor, G.H., 2002, Development of new spatial grids of R-factor and 10-yr EI30 for the conterminous United States: Report submitted to the U.S. Environmental Protection Agency, November 2001, accessed September 30, 2014, at <ftp://ocid.nacse.org/pub/prism/maps/Precipitation/final.doc>

Flint, A.L., and Flint, L.E., 2007, Application of the basin characterization model to estimate in-place recharge and runoff potential in the Basin and Range carbonate-rock aquifer system, White Pine County, Nevada, and adjacent areas in Nevada and Utah: U.S. Geological Survey Scientific Investigations Report 2007–5099, 20 p

Johnson, K.S., 2008, Evaporite-karst problems and studies in the USA, *Environmental Geology*, v. 53, no. 5, p. 937–943, doi: 10.1007/s00254-007-0716-8

Kenney, T.A., Gerner, S.J., Buto, S.G., and Spangler, L.E., 2009, Spatially referenced statistical assessment of dissolved-solids load sources and transport in streams of the Upper Colorado River Basin: U.S. Geological Survey Scientific Investigations Report 2009-5007, 50 p. <http://pubs.usgs.gov/sir/2009/5007/>.

King, P.B., and Beikman, H.M., 1974, Geologic map of the United States: U.S. Geological Survey Professional Paper 901, 40 p., 2 pl., scale 1:2,500,000. <http://pubs.usgs.gov/dds/dds11/kb.html>.

Olson, J.R. and Hawkins, C.P., 2014, Geochemical Characteristics of the Conterminous United States: U.S. Geological Survey data release, <http://dx.doi.org/10.5066/F7X0653P>.

Reeves, M.C., and Mitchell, J.E., 2011, Extent of conterminous US rangelands; Quantifying implications of differing agency perspectives: *Rangeland Ecology Management*, v.64, p. 585–597, doi: 10.2111/REM-D-11-00035.1.

Wieczorek, M.E., and LaMotte, A.E., 2010, Attributes for MRB_E2RF1 catchments by major river basins in the conterminous United States—Base-flow index, 2002: U.S. Geological Survey Digital Data Series 491–04, http://water.usgs.gov/GIS/metadata/usgswrd/XML/mrb_e2rf1_bfi.xml

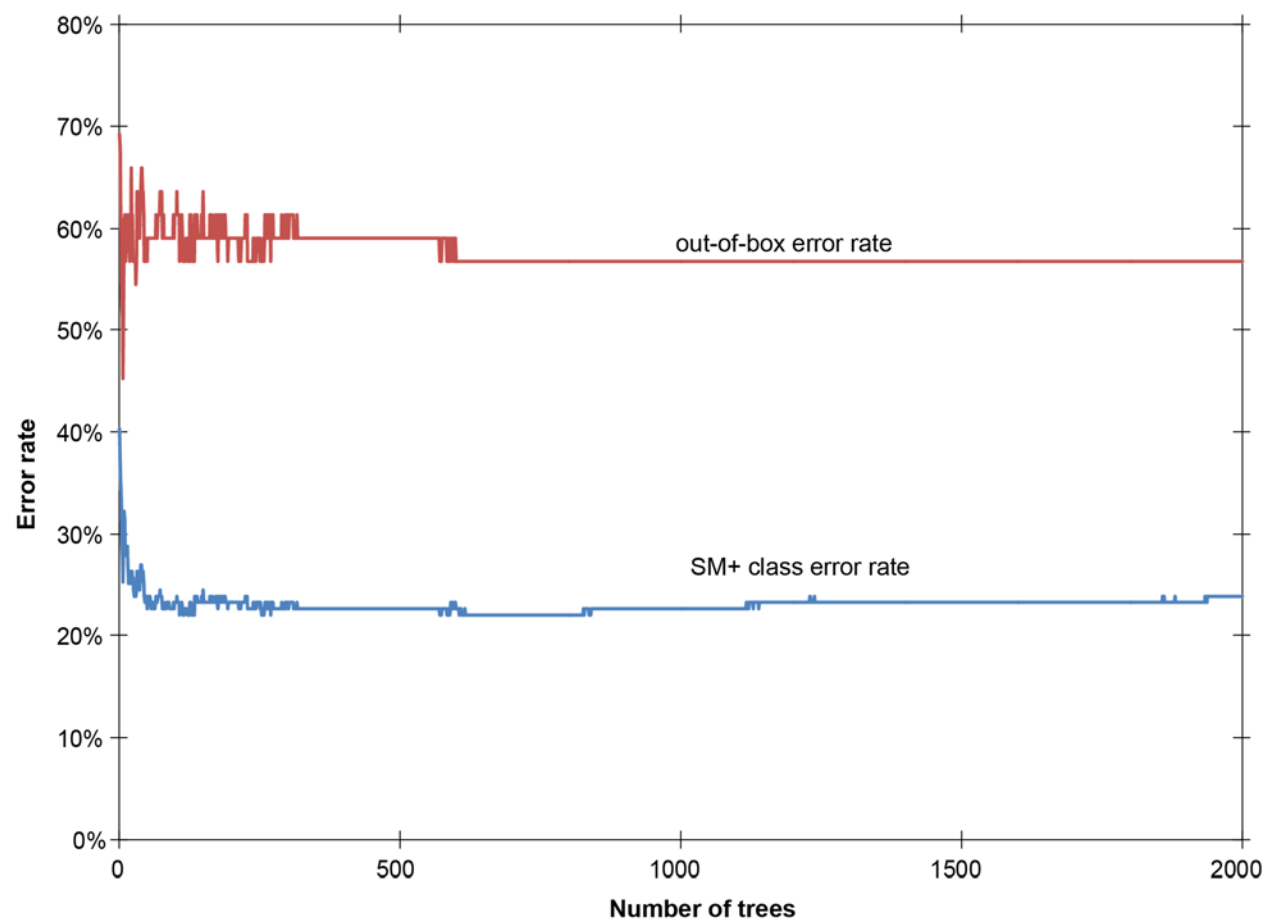


Figure S1. Out-of-box and SM+ class error rates as a function of the number of trees in initial random forests classification simulations using default argument values

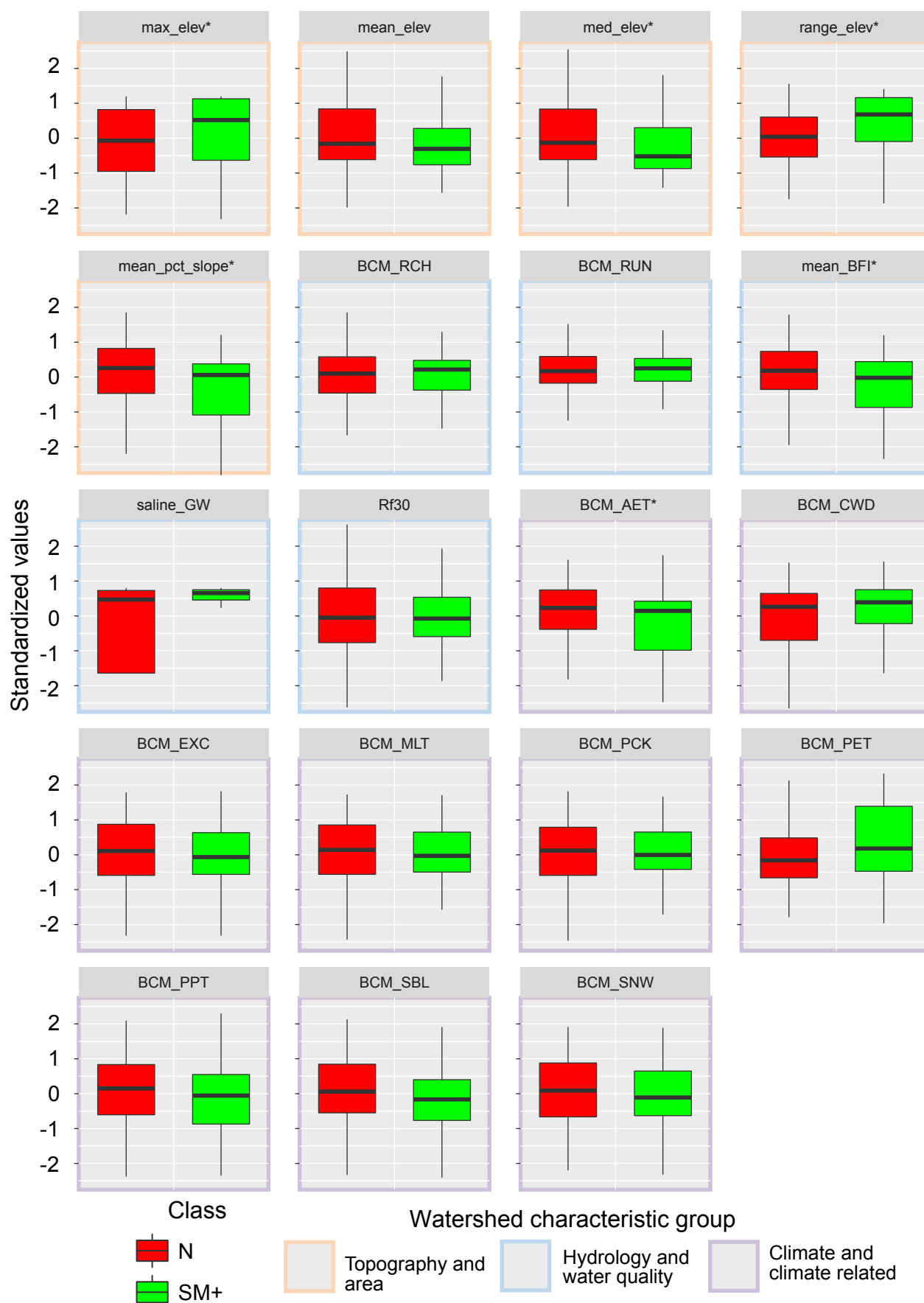


Figure S2. Boxplots showing median and interquartile range (box) and maximum and minimum values not including outliers (whiskers) for standardized watershed characteristic data. * denotes p-value<0.05 for Wilcoxon rank-sum test on SM+ and N classes for that characteristic. See Supplementary Materials Table S4 for all standardized values and Wilcoxon rank-sum results.

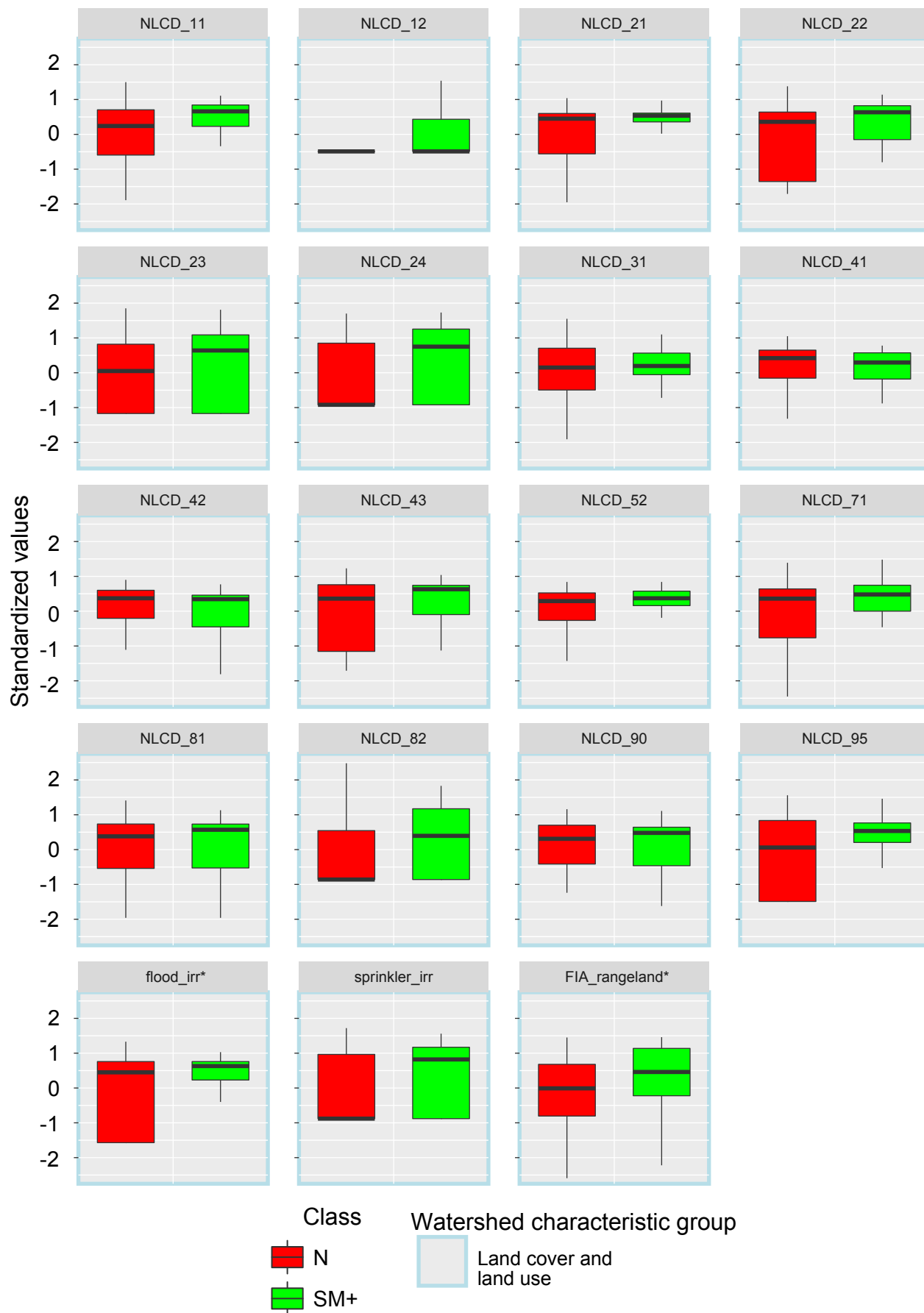


Figure S2. — Continued.

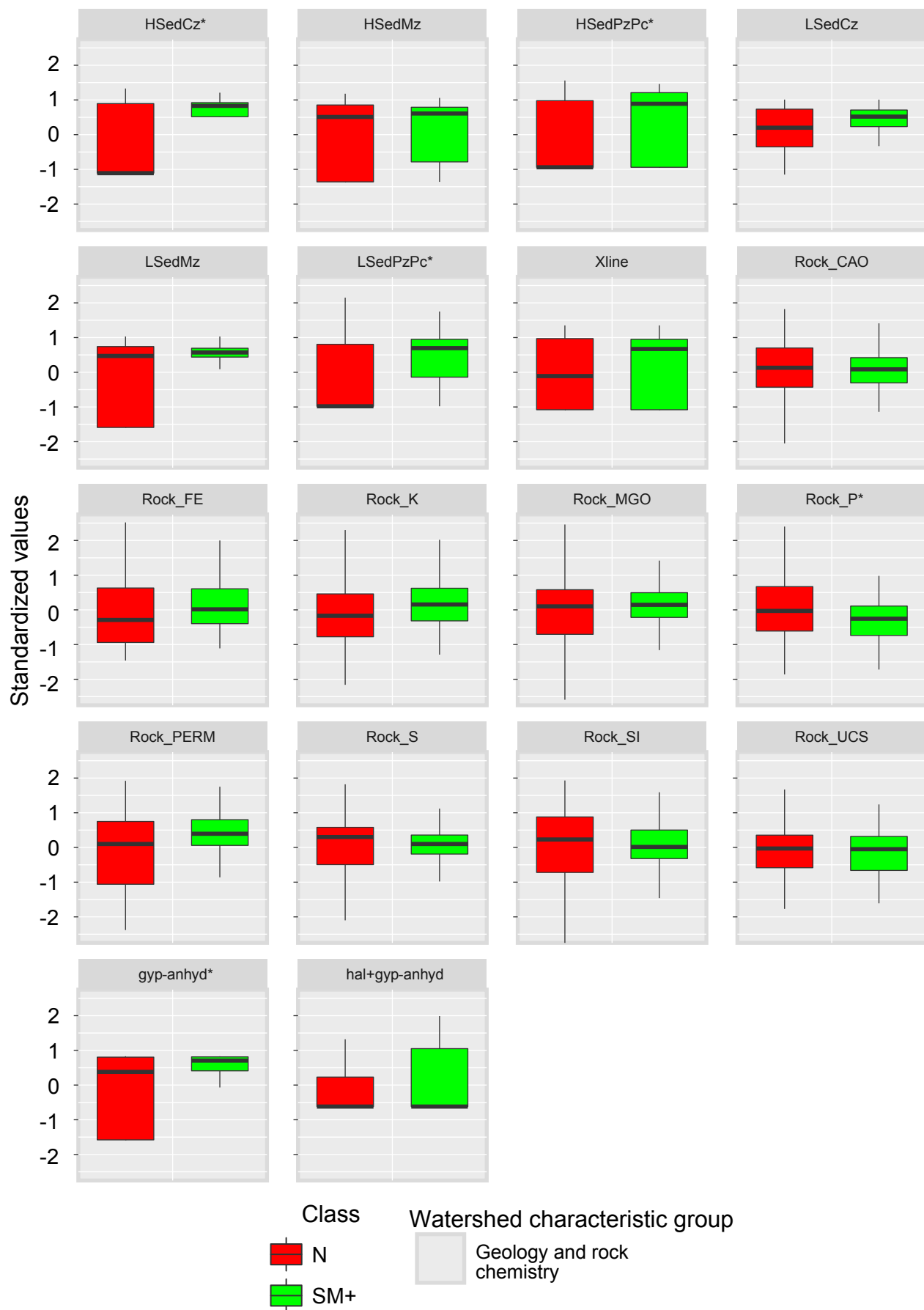


Figure S2. — Continued.

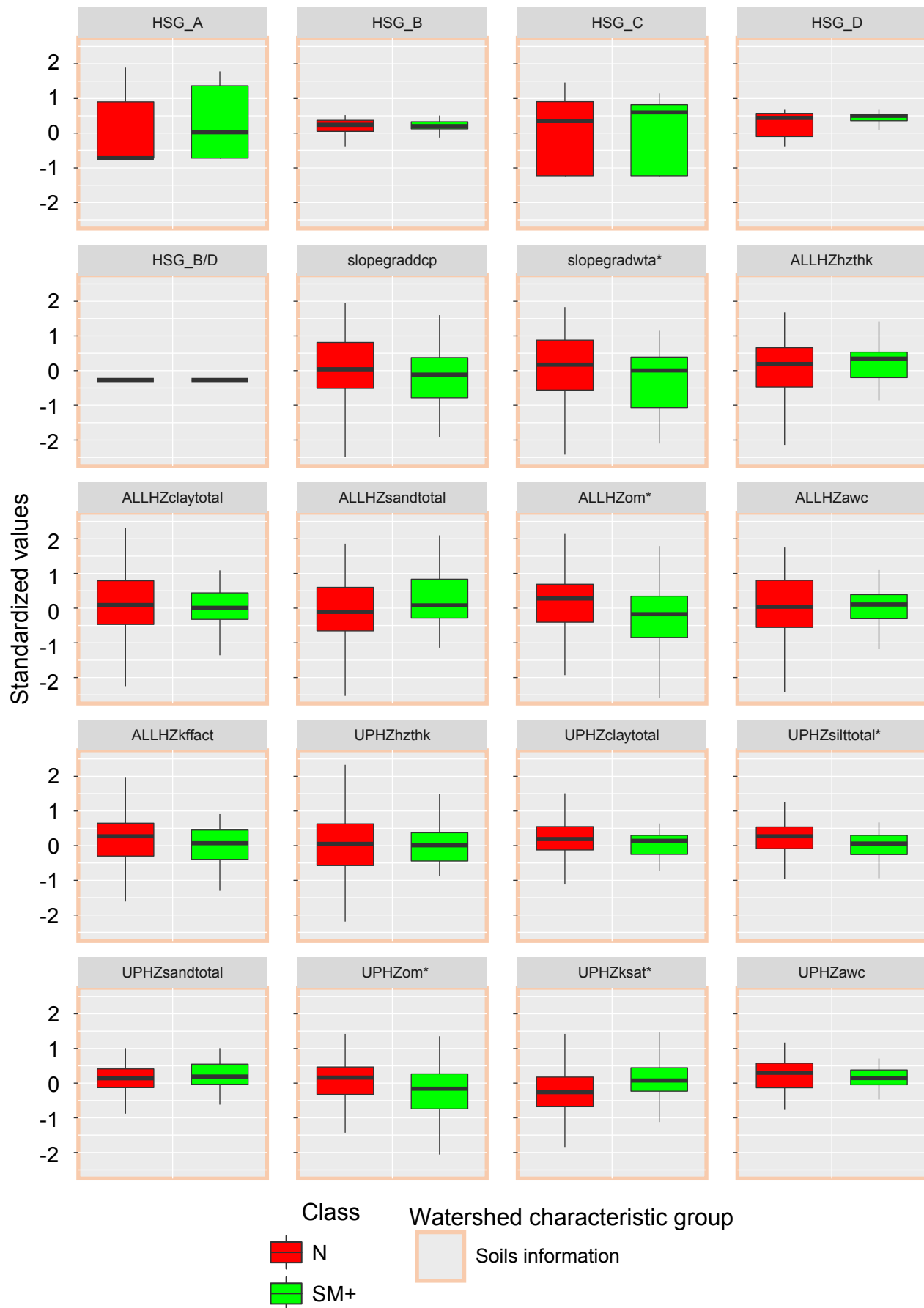


Figure S2. — Continued.