Appendices

Supporting information on work outlined in the main report and technical chapters (TC 1 to TC 8)

The following Appendices provide further information on the work outlined in the main report and technical chapters (TC 1 to TC 8). Specifically, the appendices include the following:

Appendix A provides a literature review of national and international studies of drivers of hydrochemistry.

Appendix B outlines hydrochemical metrics, including water type, mineral saturation indices, redox state, recharge altitude, and soil properties.

Appendix C details the methods used for this report, including data quality assurance and quality control, and multivariate statistical methods used.

Appendix D outlines capture zone delineation for surface and groundwater sites.

Appendix E summarises the clustering of groundwater chemical data.

Appendix F provides the datasets employed in the conceptual model.

Appendix A – Literature review

A Literature review

A1 Introduction

It is well established that the dynamics of hydrochemistry are generally driven by fluctuations in natural conditions or processes, and anthropogenic impacts (Clark and Fritz, 1997; Drever, 1997; Kendall and Caldwell, 1998; McMahon and Chapelle, 2008; Jiang et al., 2009; Dinka et al., 2015). These so called 'drivers' of hydrochemical variation include climate, geomorphic setting, substrate composition (rock, soil and biological substrate type), recharge source/mechanism and land use type and management. These drivers influence solute concentrations through various reactions. These include: precipitation, sorption, complexation and ion exchange reactions as well as isotopic fractionation and redox reactions.

The assemblage of drivers varies naturally in space (some in time). Generally, different assemblages of key drivers result in distinct hydrochemical signatures: Where the assemblage of key drivers is similar, we see compositionally similar waters. Where the assemblages of drivers are different, water composition is different. Therefore, identifying the key drivers for a region and understanding why these key drivers vary in space and time is critical to explaining spatial and variation in water hydrochemistry and hence water quality.

In this section we provide a brief review of national and international literature regarding the drivers of hydrochemical variation of water. We review five key natural drivers in the context of the Southland region. These were routinely identified in literature as altitude, proximity to coast, recharge mechanism, substrate composition (i.e. soils and geology) and redox processes (Clark and Fritz, 1997; Drever, 1997; 2002; Guler et al., 2002; Salvador et al., 2010; Rissman, 2011; Rissmann et al., 2015; Daughney et al., 2015). Land use type and land management are also important drivers of hydrochemical signatures but are not directly explored in this report because their impact on water quality largely depends on the aforementioned natural drivers.

A2 Drivers of freshwater chemistry in national and international settings

The following sub-sections should not be read in isolation as it is often difficult to tease out the sole effect of a single of the aforementioned drivers on freshwater composition. Rather a mix of the drivers has unique water chemistry outcome.

A2.1 Altitude

High altitude areas (alpine, 800 m.a.s.l) are characterised by low temperatures, more extreme climatic conditions and thin soils which do not support much plant growth (carbon biomass). Erosion rates are high and hence stable areas for advanced weathering are small (Butler et al., 2001; Fitzsimons and Veit, 2001). High water volumes and limited opportunity for advanced weathering of rock and soil material equates to low concentrations of weathering derived solutes including Na and Ca. High flushing rates and moderate to low secondary clay development results in low dissolved solute concentrations including alkalinity and poor pH buffering (Borowiak et al., 2006; Bona et al., 2007; Noges, 2009). Similarly, the marine aerosolic load to alpine areas is small and due to higher altitudes the stable isotopes of water are often strongly negative (Fritz and Clark, 1993; Bundi, 2010). Strang et al. (2010) demonstrated that dilute surface and ground waters are a characteristic of alpine areas and numerous authors discuss both the isotopic and solute composition of high altitude streams and rivers (Fritz and Clark, 1993; Drever, 2002; Bundi, 2010).

The scarce vegetation in alpine areas means that the levels of organic carbon are low and hence attenuation of contaminants through biological processes (e.g. denitrification) are also limited (Lambert et al., 1985; Blaschke et al., 1992; Quinn, 2002; Bundi, 2010). Accordingly, in the absence of

land use the waters of these areas are characteristically dilute, oxidising, Ca/Mg-HCOs waters that are of low nitrogen concentration (Nichol et al., 1997; Stevenson et al., 2010). In mid altitude areas (~400-800 m.a.s.l), subalpine or Hill Country, precipitation rates are lower than Alpine areas but still elevated relative to lowland areas in temperate humid climates. High precipitation volumes equate to significant flushing of major ions and dilution of surface and groundwater (Bona et al., 2007).

The presence of vegetation in this range of mid altitude due to more favourable temperatures increases the levels of organic carbon which facilitates reduction processes and associated reduction of nitrogen (Andersen, 2003). Conversely across low altitude areas, precipitation is lower and the seasonally elevated temperatures increase water loss through evapotranspiration. Lowland areas in maritime settings also receive higher marine aerosol loads than alpine or inland hill country (Nichol et al., 1997). Across New Zealand, lowland landforms also tend to be depositional in nature both older (geomorphically) and more stable than Alpine and Hill Country areas (Summerfield, 2000). Accordingly, soils are pedogenically more evolved, thicker and tend to have higher secondary clay development and base saturation (B.S.). These factors in conjunction with higher evapotranspiration rates increase the concentration of marine aerosols, endogenous and anthropogenic salts in the soil profile during the growing season (Fritz and Clark, 1997; Goyal and Harmsen, 2013).

The hydrochemical equilibrium at different altitude is however subject to variation with proximity to the coast (Salvador et al., 2010). Sea salt advection can contribute to high saline water in alpine areas that are close to the sea (Salvador et al., 2010).

A2.2 Proximity to coast

Distance to coast determines the variation in marine aerosol load and solute load of precipitation and subsequently the hydrochemistry (Nichol et al., 1997; Legrand and Puxbaum, 2007; Salvador et al., 2010). Areas that are close to the sea are subject to advections of maritime air masses that largely comprise of Na and Cl, and relatively low concentration of Ca, K, SO_4 and Mg (Legrand and Puxbaum, 2007). Waters with a highly concentrated coastal precipitation source are therefore often of Na-Cl type. Salvador et al. (2010) demonstrated that high Na load in freshwaters of central Europe is contributed by sea salt from Atlantic Ocean. In New Zealand, Nichol et al. (1997) and Rosen (2001) showed that the precipitation is characterised by a high load of marine salts and that Na deposition drastically decreases linearly with increasing distance from coast in the first 20km. A cloud mass is likely to incorporate some land based dust as it traverses inland, and this reduces the precipitation load of Na and Cl and increases that of land derived ions such as SO_4 and K (Salavador et al., 2010; Daughney et al., 2015). Similarly, dilution of precipitation occurs when cloud mass moves over an alpine area, resulting in orographic forcing around mountain ranges and subsequent intense rainout of marine salts (Daughney et al., 2015). Therefore, coastal precipitation in low altitude areas will characteristically comprise of a much higher dissolved salt load than high altitude precipitation that has crossed over mountain ranges and/or travelled considerable distances inland (Salavador et al., 2010; Daughney et al., 2015).

A2.3 Recharge mechanism

Differences in hydrochemical signatures between alpine and low altitude areas due to precipitation source suggest that recharge contributions from these areas are likely to result in corresponding load concentration in receptor areas. Recharge sources and mechanisms such as land surface recharge (LSR), riverine recharge (RR) and a combination of both riverine and surface recharge, (mixed recharge, MR) have a significant impact on freshwater quality (Blasch and Bryson, 2007). Surface and ground waters are mainly recharged by rainfall flowing over and infiltrating through the soil (LSR) and overland flow over very steep/ non-reactive surfaces of alpine or bedrock settings. Solute concentration in LSR waters is generally higher than that of overland flow waters from alpine

areas (Gat, 1980). Ground waters are also recharged by precipitation from distant sources that travels through rivers (RR). Recharge sources of ground water are commonly identified using stable isotopes of oxygen and hydrogen (e.g. $\delta^{18}O$-H_2O and $\delta^{3}H$) and concentration of major ions such as Na and Cl (Kendall and Caldwell, 1998; Blasch and Bryson, 2007; Guggenmos et al., 2011; King et al., 2015). Rivers that are recharged by alpine areas often dilute the generally low concentrated (Na, Cl and $\delta^{18}O$-H_2O, $\delta^{3}H$) water from the alpine areas. Accordingly, aquifers that are adjacent to and recharged by rivers often have high water quality. For example Blasch and Bryson (2007) found that lower values of $\delta^{18}O$-H_2O, $\delta^{3}H$, Na and Cl in ground water of Central Arizona are associated with recharge from Verde River that is fed with cooler and high altitude precipitation. Ground recharge water from land surfaces usually has high loads of mineral ions derived from contact with the soil as the water passes through. The load concentration is generally inflated if the recharge water by-passes the soil matrix and flows through the macro-conduits (macropores) created by earthworms, large roots or wetting and drying cycles of the soil. The soil matrix can partially filter solutes, contaminants and pathogens from flowing water, hence if water bypasses this pathway, the load of contaminants and pathogens in ground water will increase. Various studies in New Zealand have demonstrated the increase in contaminants, pathogens and redox-sensitive species in by-pass flow (e.g. McLeod et al., 1998; Aislabie et al., 2001; Monaghan and Smith 2004; Houlbrooke et al., 2008, McLeod et al., 2008; Monaghan et al., 2016). Internationally, a bulk number of studies (e.g. Scott et al., 1998; Watson et al., 2000; Oliver, 2003; Martins et al., 2012; Oliver et al., 2013; Peukert et al., 2014; Zhang et al., 2015, among others) have demonstrated the importance of by-pass flow as an amplifier of water contaminants. Hence aquifers recharged from land surfaces often have degraded water quality. Generally higher concentrations of Na (relative to Cl) and $\delta^{18}O$-H_2O, $\delta^{3}H$, are indicative of land surface recharge than riverine recharge (e.g. Gat, 1980; Blasch and Bryson, 2007). The important role played by soil in determining the hydrochemistry of land surface recharged aquifers means soil chemistry is likely to confound the effect of recharge source/mechanism on hydrochemistry signature. Therefore, it is vital to take into account ancillary geologic and mineralogical and biological composition of particular soils and sub-soil layers that are in contact with water as it flows through the landscape.

A2.4 Substrate composition

Substrate composition (soil, rock and biological substrate types) plays a critical role in the evolution of hydrochemical signatures of ground and surface waters (Olson, 2012; Daughney et al., 2015; Rissmann, 2015). Substrate composition influences pH, cation exchange capacity (CEC), base saturation (BS), total exchangeable bases (TEB), cation concentrations and carbon and nitrogen content (C and N content, respectively). Soil and rock minerals are the first mediums of interaction when rain infiltrates or runoffs as it cascades over a landscape (Smith and Monaghan, 2003; Rissmann et al., 2015). Therefore the physical and chemical composition of soils and rocks determine the chemical composition of water flowing through the landscape. Thin, CO_2 rich soils are often rapidly weathered and leached of soluble minerals by infiltrating rain, resulting in acidic waters (high in carbonic acid) recharging the ground water (Daughney et al., 2015; Essington, 2015). The chemical composition of infiltrating water and its residence time also determine the rate of chemical weathering of rock and soil minerals which in turn influences the chemical composition of recharge water (Rissmann et al., 2015; Daughney et al., 2015; Essington, 2015). Generally, the pH of precipitation is slightly acidic due to interaction with atmospheric CO_2, so recharge waters become more acidic when rainfall infiltrates through the vadose zone rich in CO_2. Reactive rocks such as limestone undergo rapid weathering as acidic precipitation flows through the soil, thereby enhancing the concentration of dissolved ions in water (Essington, 2015). The acidity of groundwater is reduced when precipitation flows through slow reacting felsic or mafic rocks (Wang et al., 2014).

Infiltrating water can also leach cations, resulting in the replacement of monovalent ions (Na^+ and K^+) with divalent (Ca^{2+}, Mg^{2+}) and trivalent ions (Fe^{3+}, Al^{3+}) (Rissmann et al., 2015). This is more rapid in soils with high clay content. Alternatively, in areas of salt water intrusion, if clay soil is saturated with Ca^{2+} or Mg^{2+}, the addition of K^+ or Na^+ ions in high levels may flush these cations from the exchange sites back into solution (Rademacher et al., 2001; Daughney et al., 2015). Some soil particles store anions which enhance the immobilisation of major ions such as phosphorus (Nziguheba et al., 1998; Frossard et al., 2000; Rissmann et al., 2015). In oxidising soils with a high pH, phosphorus solubility is low and phosphorus are strongly bound to soil materials and hence the concentration in soil water is low (Frossard et al., 2010; Rissmann et al., 2015). However, in organic matter rich soils with low pH (>5) redox reactions raise the solubility of phosphorus, and increase its concentration in soil waters.

A2.5 Redox conditions

Redox is reported as one of or the main controls over the variability in water chemical data nationally and internationally (Winter et al., 1998; Daughney et al., 2005; 2010; 2015; McMahon and Chapelle, 2008; Guggenmos et al., 2011; Rissman et al., 2012; Essington, 2015). Redox evolution is governed by organic carbon (electron donor), hydrology (especially in soils) and to some extent aquifer reduction potential (Krantz and Powers, 2002; McMahon and Chaplelle, 2009; Rissmann, 2011; Rissmann et al., 2012; Killick et al., 2014). In temperate climate soils, organic carbon is seldom limited in temperate climate soils other than in areas above the treeline or 0 degree isotherm (Collins and Kuehl, 2000; Bernal, 2008; Fissore et al., 2009). Some substrates (or rock types) may have high concentrations of metabolisable organic carbon (peat, lignite, brown coal) that the majority do not (mud, silt, sandstone) (Krantz and Powers, 2002; Drever, 2002; Rissman, 2011). Redox evolution is determined by the abundance of metabolisable organic carbon, water residence time, exclusion from the atmosphere and redox buffers. The main redox buffer in most natural low temperature systems is Fe(III), the 4th largest abundance (McMahon and Chapelle, 2008).

For soils, redox is governed by soil hydrology, particularly drainage class and organic carbon content (Mitsch and Gosselink, 2007; Bernal, 2008; Stenger et al., 2013; 2014; Killick et al., 2014). Well drained soils often have low organic carbon content and are oxidising, while poorly drained soils accumulate high organic carbon content and are reducing (Mitsch and Gosselink, 2007). Accordingly, water passing through an oxidising soil will retain oxic characteristics and that passing through a reducing soil will be reduced, and the extent to which both reactions happen depends on residence time in the soil (Rissmann et al., 2015). Redox reactions require sufficient time for metabolic processes to occur (Andersen, 2003).

A3 Summary

The above gave a summary of national and international key drivers of hydrochemistry. Overall, we identified many national and international studies that consistently identify a suite of key drivers determining hydrochemical and water quality outcomes (e.g. Nichol et al., 1997; Krantz and Powers, 2002; Blasch and Bryson, 2007; Legrand and Puxbaum, 2007;McMahon and Chapelle, 2008; Salavador et al., 2010; Frossard et al., 2010; Rissmann, 2011 ; Rissmann et al., 2012; Daughney et al., 2015). The drivers include: altitude, proximity to coast, recharge mechanism, substrate composition (i.e. soils and geology) and redox processes.

The specific combination of the above key drivers, in conjunction with land use intensity, determines the natural variability in hydrochemistry. The latter equates to distinct spatial variation hydrochemical indicators and water quality metrics.

Appendix B – Hydrochemical metrics

This appendix details hydrochemical metrics used in this study.

B1 Water type

Water type is assigned as described by Freeze and Cherry (1979) on the basis of measured concentrations of the major ions Ca, Mg, Na, K, HCO_3, Cl and SO_4. The water type is determined from the dominant dissolved cation(s) and the dominant dissolved anion(s), based on concentrations expressed in milliequivalents per litre[1]. For example, seawater has Na as the dominant cation and Cl as the dominant anion and so is referred to as a Na-Cl type water.

B2 Mineral saturation indices

Mineral dissolution reactions can play an important role in governing the hydrochemistry of surface water and groundwater. Consider the generic reaction where the mineral A_nB_m dissolves to produce the ions A^+ and B^-:

$$A_nB_m \rightarrow nA^+ + mB^- \quad \text{(B1)}$$

The solubility product K_{SP} is a constant that is specific for the mineral in question and that defines the relationship between A^+ and B^- at chemical equilibrium:

$$K_{SP} = [A^+]^n[B^-]^m \quad \text{(B2)}$$

where square brackets represent the activities[2] of the enclosed ions and n and m are the stoichiometric coefficients in the dissolution reaction. The saturation index (SI) provides an evaluation of whether a given solution is at chemical equilibrium with respect to a given mineral:

$$SI = \log\left(\frac{[A^+]^n[B^-]^m}{K_{SP}}\right) \quad \text{(B3)}$$

If SI = 0 then the system is at chemical equilibrium and the solution is saturated with respect to the given mineral; if SI < 0 the solution is under-saturated and the mineral is expected to dissolve; and if SI > 0 the solution is oversaturated and the mineral is expected to precipitate.

Measurement of the dissolved concentrations of several different elements, such as those listed in Table A 1, allows the saturation indices for a wide variety of minerals to be calculated by comparison to published tables of mineral solubility products (Langmuir, 1997). This information on mineral saturation state can provide useful information on what types of minerals natural water has reacted with, which in turn can provide insight into hydrological processes and flow paths (Hem, 1985).

In this study, saturation indices are calculated with respect to a limited set of minerals. The minerals of interest for saturation calculations and their corresponding solubility products are listed in Table A1. These minerals are selected because they are most likely to reach chemical equilibrium as a result of in natural water-soil-rock interaction (Lasaga, 1984; Langmuir, 1997), and so differences in saturation state can be used to infer origin and flow pathways of water (Lambrakis et al., 2004; Daughney and Reeves, 2005; Cloutier et al., 2008). All calculations of saturation indices were performed using AquaChem 2014 software (Schlumberger Water Services) and the phreeqc

[1] A milliequivalent (meq) is a measurement of the molar concentration of the ion multiplied by its ionic charge. For example, the Ca ion has a formula weight of 40 grams per mole and an ionic charge of +2, so a concentration of 10 mg/L is equivalent to 0.5 meq/L.

[2] In thermodynamics, activity defines the amount of a substance that is available for reaction. For most natural waters, activity is approximately equal to concentration in moles per litre of solution.

thermodynamic database (Parkhurst and Appelo, 2013) and are based on either calculated median concentrations or sample-specific analytical results for the relevant elements.

Table A 1: Solubility products for selected minerals (Fitz, 2002).

Mineral	Log K_{SP}	Dissolution Reaction
Quartz	-3.98	$SiO_2 + 2H_2O \leftrightarrow Si(OH)_4$
Chalcedony	-3.55	$SiO_2 + 2H_2O \leftrightarrow Si(OH)_4$
Amorphous silica	-2.71	$SiO_2 + 2H_2O \leftrightarrow Si(OH)_4$
Calcite	-8.48	$CaCO_3 \leftrightarrow Ca^{2+} + CO_3^{2-}$
Dolomite	-17.1	$CaMg(CO_3)_2 \leftrightarrow Ca^{2+} + Mg^{2+} + 2CO_3^{2-}$
Aragonite	-8.34	$CaCO_3 \leftrightarrow Ca^{2}+ + CO_3^{2}-$
Siderite	-10.9	$FeCO_3 \leftrightarrow Fe^{2+} + CO_3^{2-}$
Rhodochrosite	-10.39	$MnCO_3 \leftrightarrow Mn^{2+} + CO_3^{2-}$
Magnesite	-8.03	$MgCO_3 \leftrightarrow Mg^{2+} + CO_3^{2-}$
Gypsum	-4.58	$CaSO_4 + 2H_2O \leftrightarrow Ca^{2+} + SO_4^{2-} + 2H_2O$
Anhydrite	-4.36	$CaSO_4 \leftrightarrow Ca^{2+} + SO_4^{2-}$
Goethite	-41.5	$FeOOH + H_2O \leftrightarrow Fe^{3+} + 3OH^-$
Manganite	-18.26	$MnOOH + H_2O \leftrightarrow Mn^{3+} + 3OH^-$

B3 Redox state

Assignment of redox state is based on the methodology of McMahon and Chapelle (2008), as implemented in the excel spread sheet of Jurgens et al. (2009). This approach involves comparison of the concentrations of selected redox-sensitive substances to the thresholds listed in Table A 2. It is recognised that redox reactions are usually not at chemical equilibrium in low temperature groundwaters. Hence this approach for assignment of redox state is based on the widely observed ecological succession of electron-accepting processes. Redox assignment for each sample includes identification of the dominant redox state (i.e., oxic, suboxic, or mixed (oxic-anoxic)) and the principal terminal electron accepting process operating within that groundwater. If the required data are available, iron reducing conditions are differentiated from SO_4 reducing conditions by measuring the mass ratio of Fe(II) to total sulphide (Chappelle et al., 2009).

Table A 2: redox category and dominant redox process for groundwater and surface water as determined from measured concentrations of water quality parameters (after Jurgens et al., 2009)

Redox Category	Redox Process	Criteria for Assigning Redox State					
		D.O. ≥0.5 mg/L	NO_3-N ≥0.5 mg/L	Mn(II) ≥0.05 mg/L	Fe(II) ≥0.01 mg/L	SO_4 ≥0.5 mg/L	Fe/H_2S ≥0.3 (mass ratio)
Oxic	O_2	Y	NA	N	N	NA	NA
Suboxic	O_2	N	N	N	N	NA	NA
Mixed (oxic-anoxic)	O_2-Mn(IV)	Y	NA	Y	N	NA	NA
Mixed (oxic-anoxic)	O_2-Fe(III)	Y	N	NA	Y	Y	Y
Mixed (oxic-anoxic)	O_2-Fe(III)-SO_4	Y	N	NA	Y	Y	Y
Mixed (oxic-anoxic)	O_2-SO_4	Y	N	NA	Y	Y	N
Mixed (oxic-anoxic)	O_2-CH_4 gen.	Y	N	NA	Y	N	NA
Mixed (anoxic)	NO_3-Mn(IV)	N	Y	Y	N	NA	NA
Mixed (anoxic)	NO_3-CH_4 gen.	N	Y	NA	Y	N	NA
Anoxic	NO_3	N	Y	N	N	NA	NA
Anoxic	Mn(IV)	N	N	Y	N	NA	NA
Anoxic	Fe(III)	N	N	NA	Y	Y	Y
Anoxic	SO_4	N	N	NA	Y	Y	N
Anoxic	CH_4 gen.	N	N	NA	Y	N	NA

B4 Recharge altitude

Recharge altitudes were calculated for each site based on the median data. The method for this is as follows. Firstly, the mean $\delta^{18}O$ composition of sea-level precipitation in Southland is derived utilising the following equation from Dansgaard (1964):

$$\delta^{18}O = 0.695T_m - 13.6 \quad \text{(B4)}$$

Where T_m is mean air temperature, 9.3°C for Invercargill, from which a mean sea-level $\delta^{18}O$ value of -7.1‰ is derived. This value is similar to the mean $\delta^{18}O$ composition of -6.98 ‰ (n = 32) for low altitude coastal groundwaters in Southland (ES unpublished data). Given some confidence about the mean $\delta^{18}O$ composition of sea level precipitation, it is possible to estimate the mean recharge altitude assuming a depletion factor of -0.23‰ per 100 m rise in altitude as recommended by Stewart and Taylor (1979) for New Zealand settings.

B5 Soil properties (BS%, CEC, TEB)

Cation Exchange Capacity (CEC) is defined as a measure of the number of negatively-charged binding sites in the soil. For example, a high CEC of soil indicates that the soil has a high ability to bind or hold exchangeable cations.

Total Exchangeable Bases (TEB) is defined as the sum of cations in the soil that act as bases. These include K, Mg, Ca and Na (B5)

$$TEB = [Ca]+[Mg]+[Ca]+[Na] \quad \text{(B5)}$$

Base saturation (BS%) is defined as the fraction of the negative binding sites occupied by bases, such as K, Mg, Ca and Na (B6). Cations which are alkaline/act as base raise the soil pH

$$BS\% = TEB/CEC \quad \text{(B6)}$$

Appendix C - Methods

This appendix details data analysis methods used in this study. Specifically, data quality assurance and use of self organizing maps to estimate missing data, as well as multivariate statistical methods (HCA and PCA) are detailed.

C1 Data quality assurance and quality control

GW, SW, soil water, soil and precipitation samples were analysed by Hills Laboratories following standard procedures detailed in each section. Careful examination and grooming of the hydrochemical dataset as an essential first step in any investigation such as this was carried out.

In this study, five approaches were used for quality assurance and quality control (QA/QC) of the raw data. First, the continuity and length of the hydrochemical dataset were evaluated. Second, coverage of the hydrochemical parameters was assessed. This involved comparison of results for dissolved vs. total concentrations and field vs. lab measurements to determine if any parameters were sufficiently highly correlated that they could be dropped from consideration without significant loss of information. Where small populations of censored data were present these were processed using recommendations of Hornung and Reed (1990) (halved). Where small populations of "greater than detection" values existed these were processed by using the value given as the upper limit of detection. Where data had been entered into the database under pseudonyms i.e. "conductivity lab (µs/cm)" and "conductivity lab" the two columns were combined and the analyte renamed "conductivity (combined)". Third, standard QA/QC was applied, including calculation of charge balance error and assessment of the robustness of estimated results derived from the SOM approach. Fourthly, cumulative probability plots were used for identification of outliers (and sub-populations within the dataset). Finally, the characteristics and representativeness of the monitoring sites were evaluated. For example, sites with apparent contamination were identified and not considered in the data analysis, because their chemistry does not reflect natural processes. Specific methodological details for all of the aforementioned techniques are provided in (Daughney et al., 2015), except for the cumulative probability plots, which are described subsequently.

One notable advance in this study was to apply a modified version of the self-organising map (SOM) technique to estimate the values of every hydrochemical parameter that had not been specifically measured in the lab or the field. This approach greatly expanded the amount of information that could be used for the hydrochemical evolutions. A series of QA/QC checks confirmed that these hydrochemical datasets are fit for the purpose of this investigation and would permit a high level of analysis. The use of SOMs is also described subsequently.

C1.1 Cumulative probability plots for data QA/QC

Cumulative probability plots were used to identify outliers and to assess the robustness of estimation methods used for missing hydrochemical data. Cumulative probability plots were also used to support the identification of distinct populations within a datasets e.g. populations that are likely to be representative of direct contamination or a result from differing processes. All are detailed below.

To assess the robustness of estimation methods for missing hydrochemistry data, the measured and estimated results were compared using the Kolmogorov-Smirnov test for each parameter. The test evaluates the maximum distance between the cumulative distributions of a single hydrochemical parameter between the population of measured vs. estimated results. The Kolmogorov-Smirnov test was performed using $\alpha = 0.05$ using Statgraphics version 15.2.06 (Manugistics Inc., USA). The results of the Kolmogorov-Smirnov test were visualised using quantile-quantile plots as shown below.

Using Mg as an example, Figure A 1 compares the distribution of the lab-measured concentrations for all groundwater and surface water samples collected from 2010 onwards (n = 5785) to the distribution of lab-measured plus estimated concentrations (n = 26399, of which 78% are estimated values). Quantile-quantile plots for estimation of missing hydrochemical data are based on 1) other available lab-measured hydrochemical data or 2) available lab-measured hydrochemical data, plus the date of sample collection and whether the sample was a groundwater or surface water. This comparison indicates that, for Mg, the former estimation approach gives slightly greater conformance to the 1-to-1 line, indicating a closer match to the distribution of the lab-measured concentrations. These results for Mg were found to be typical of other hydrochemical parameters. Hence the estimation approach based on lab-measured hydrochemical data alone, i.e. without inclusion of any additional ancillary data as input, is applied for the remainder of this study. For example, this meant that a missing value for Mg concentration in a sample would be estimated from the concentrations of other parameters that were measured in that sample, e.g. perhaps Ca, Cl, etc.

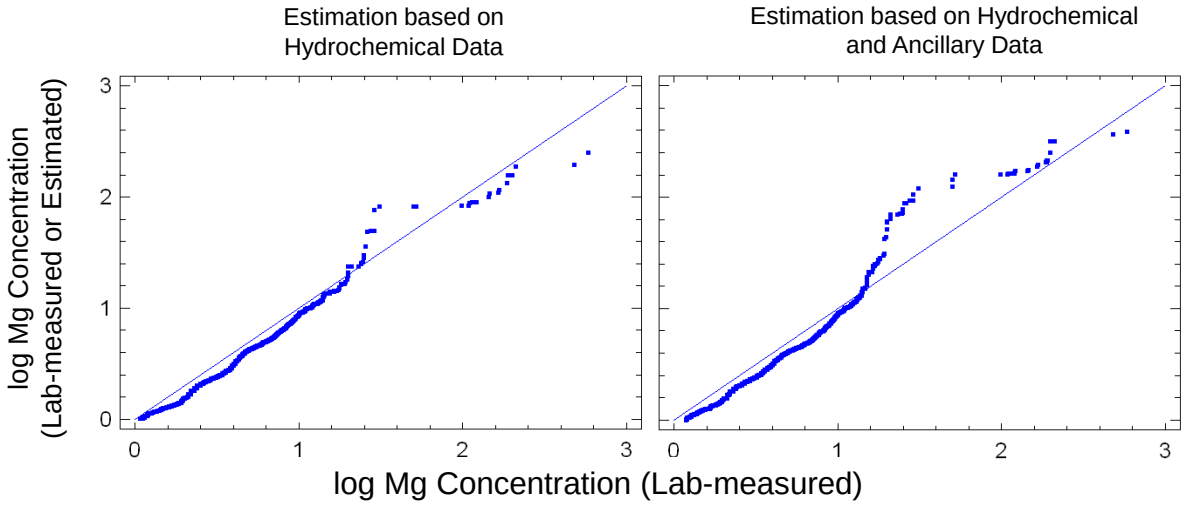


Figure A 1: Comparison of the distributions of lab-measured and estimated concentrations using two different estimation approaches. If the estimation method generates a distribution that is similar to that of the lab measurement, the points should be close to the diagonal 1-to-1 line. A general pattern of points on either side of the line indicates a difference between the two distributions.

To identify outliers/populations that are likely to be representative of direct contamination, cumulative probability plots were calculated for analytes that are indicative of direct anthropogenic source contamination (Na, Cl, K, SO_4 and *E. Coli*). These samples were subsequently removed from the dataset (following methods of Sinclair, 1974). An example is illustrated in Figure A **2**.

Two major infliction points in the cumulative probability identify three distinct populations within the data. The infliction point/threshold between population 1 and 2 30 mg/L corresponds with the upper limit for natural background for Southland (Rissmann et al., 2012). The infliction point/threshold between population 2 and 3 (40 mg/L) corresponds to the threshold beyond which a significant anthropogenic Cl input is evident (Rissmann et al., 2012) supporting that population 3 has been effected by direct contamination and can be considered as outlier population.

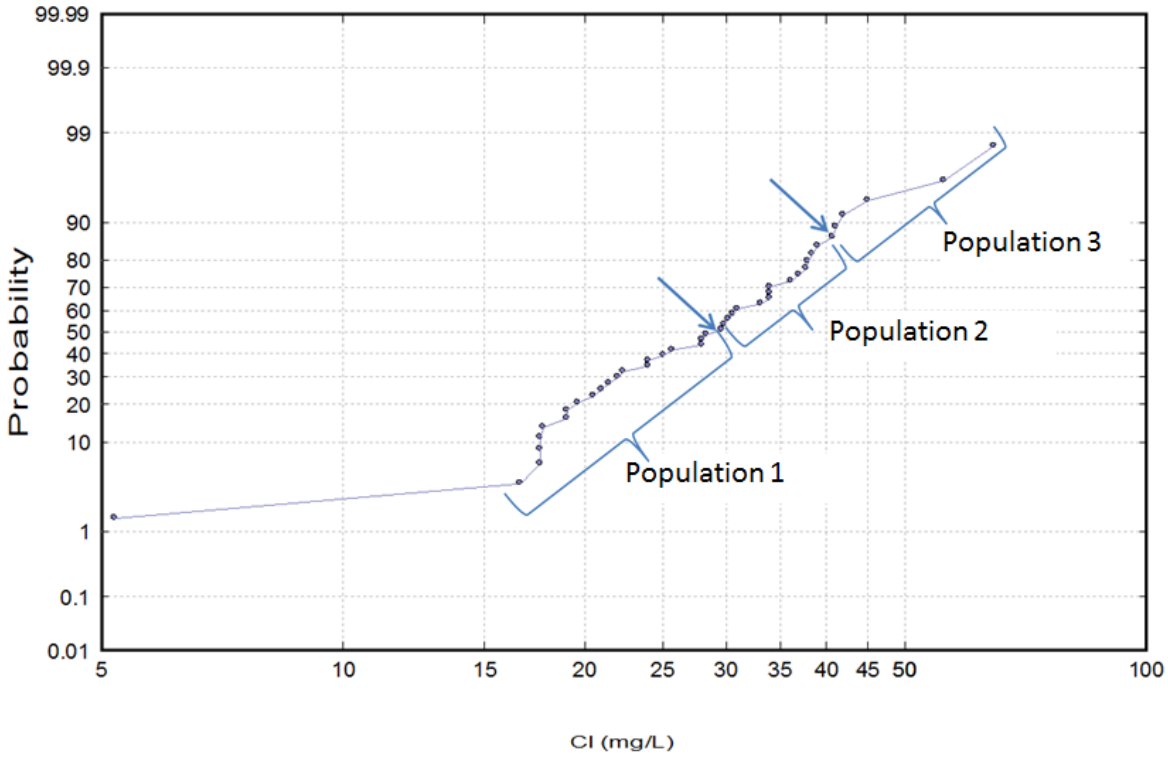


Figure A 2: Cumulative probability plot of Cl (mg/L) for tile drain waters identifies 3 distinct populations within the dataset. Two major inflection points are depicted by arrows occur at c. 30 and c. 40 mg/L, respectively. The 30 mg/L inflection point reflects the upper limit for natural background for Southland whereas beyond 40 mg/l a significant anthropogenic Cl input is evident (Rissmann et al., 2012).

C1.2 Self-organising maps for estimation of missing data

For various reasons, data sources often times are missing measurement observations. Given that traditional hydrochemical data analysis relies on continuous measurements, much of the available information associated to a data set goes unused. One means of expanding the usefulness of measurement is to estimate missing values in a sparse data set using an artificial adaptive system technique. Artificial adaptive system techniques complete the missing data sets by rebuilding the hyper-surface based on underlying spatiotemporal multivariate relations. Some examples include the auto-associative neural network (Buscema et al., 1998) and self-organising map (Kohonen, 2001) techniques.

In this study, a modified version of the self-organising map (SOM) technique is used to complete the hydrochemical data set. Application of the self-organizing map (SOM) technique to hydrochemical data involves training, diversity, and estimation. The SOM training process provides a way of representing multidimensional data in a lower dimensional space than the original data set. Reducing the dimensionality is based on a two-step process that is performed each time an input pattern is presented to the map: competition to determine the best matching unit (BMU) vector and cooperative learning (spreading information contained in the current input vector across the map).

In the first training step, a weight vector W_i with the same dimensionality as the input data vectors V_j is assigned to a grid neurons in the SOM. Following an iterative procedure, the SOM is constructed considering the differences between the normalized input vector V_j and the weights W_i of the neurons given by Eqn. C1:

$$D_{ij} = (W_i - V_j)^T (W_i - V_j) \quad \text{(C1)}$$

where T is the transpose. Normalisation of the input vectors is conducted with respect to their standard deviations.

In the second step, a weight update is determined as a function of the distance to the current BMU, expressed through the Gaussian neighbourhood function, $\varphi(\Delta, n)$. The rate used to adjust the weight of neurons decreases with distance between each neuron and BMU. Updates of the weights are adjusted according to Eqn C2:

$$W_i\ (n+1) = W_i\ (n) + \alpha(n)\varphi(\Delta, n)[V_i(n) - W_i\ (n)], \qquad \text{(C2)}$$

where *α(n)* is a scalar value called the learning rate. The BMU ensures that the largest weight correction is adjusted in the direction of the input vector. The association effect takes place at the neighbouring nodes but to a lesser degree because of the Gaussian shape. This adaption procedure stretches the weight vectors of the BMU and its topological neighbours towards the input vector. Presenting similar input vectors to the map provides further activations in the same neighbourhood and thereby tends to produce clustering of data in the feature space. Association between neurons decreases during the learning process (the width of the neighbourhood function *Φ(n)* is forced to decrease with n preserving large clusters of data while enabling the separation of clusters that are closely spaced).

Model diversity is evaluated in the underlying density function using component planes (Vesanto, 1999). The component-planes representation provides information on the distribution of component (variable) values according to a temperature scale (where red are high values, and blue are low values). In this way, an individual component plane is analogous to a histogram except the same value can be present in multiple places of the SOM when it relates to different groups of variables. The simultaneous inspection of multiple component planes allows for the visualization of correlated variables despite their disparate and sparse nature. The presence of relations will appear as similar patterns among component plane variables. In the case of a strong positive correlation along the sampling gradient, the colour patterns will be identical among variables meaning that as one variable increases (or decreases) the others do the same. Conversely, a strong negative correlation among variables along the sampling gradient will appear as the same patterns but opposite colour distribution, meaning that as values in one variable increase the other variable decreases. Principal component analysis (Eigenanalysis of the correlation matrix) is applied to component planes to visually reveal the relative strength among correlated groups of variables. One extension of this analysis is cross-component planes in which pairs of component planes map values are used to compute the correlation among monotonically increasing (or decreasing) trends.

In the traditional SOM-based estimation approach, estimates of missing values are taken directly from the prototype vectors of the BMUs (Fessant and Midenet, 2002). Often times certain training data sets result in biased estimates (Dickson and Giblin, 2007; Malek, et al., 2008) requiring a modified scheme that incorporates bootstrapping (Breiman, 1996), ensemble average (Rallo et al., 2002), or nearest neighbour (Malek et al., 2008). This study uses an alternative scheme that uses the associated BMUs as the initial values. The final values are arrived at iteratively (3-5 iterations) based on the simultaneous minimisation of the topographical error and quality error vectors (Figure A 3). The topographical error vector is defined as the proportion of all data vectors for which first and second BMUs are not adjacent units, and the quality error vector is defined as the average distance between each data vector and its BMU and is a measure of map resolution. The estimation of missing values (often referred to as imputation) is done simultaneously for all variables (Kalteh and Hjorth, 2009). In addition to missing data in sparse files, the modified SOM can be used as spatiotemporal interpolator of training variables to nodes (or elements) associated with groundwater model grid (mesh) (Friedel, 2014; Friedel, in review).

In using the SOM as an estimator of missing values, it may be necessary to cope with the phenomenon of over-fitting. Over-fitting occurs when the degrees of freedom (complexity) associated with the problem are too high to be constrained in a stable way by the available data set (Iwashita et al., 2011) and (Friedel et al., 2012). To test for over-fitting, a 'leave one out' cross-validation strategy is used in which there is one test value, and the training data set consists of all but that one test value (Efron and Tibshirani, 1993). Training and testing are carried out N times in a round-robin manner, where N is the number of values for a dependent variable each representing a new SOM. This strategy guarantees a minimum bias of the estimated prediction error (Hastie et al., 2002) and enables nonlinear statistics to be computed for the trained SOM. Experience finds cross-validation to be suitable for relatively small (hundreds of values) data sets; for larger data sets the traditional split-sample validation approach is used. In using the split-sample validation approach, the original data set are randomly split into two equal parts: 50% known observations for training and 50% independent observations for bias evaluation. For more details about SOM training and estimation, the reader also is referred to (Kohonen, 2001; Vesanto and Alhoniemi, 2000).

Further detail on the approach and QA/QC of the SOM estimated data can be found in Daughney et al., (2015).

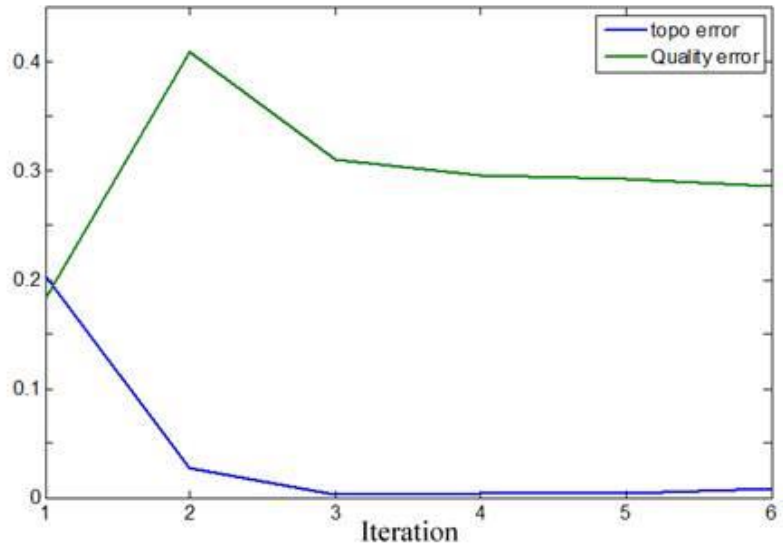


Figure A 3: Minimization of the topographic and quality error vectors (iteration 3) are used to estimate missing values in sparse data sets.

C2 Multivariate statistical methods

Multivariate statistical methods are usefully applied to evaluate relationships between the hydrochemical parameters, for example to determine whether certain parameters or groups of parameters are typically correlated (positively or negatively) or uncorrelated. Relationships between hydrochemical parameters were evaluated using Principal Components Analysis (PCA) and Hierarchical Cluster Analysis (HCA).

For both PCA and HCA log transformed (expect pH and the isotopic indicators) and z-scored site specific median values of selected hydrochemistry parameters were used. PCA and HCA were conducted using StatGraphics version 15.2.06 (Manugistics Inc., USA) and unistat version 6.0nz (unistat Ltd.).

C2.1 PCA

PCA is essentially a mathematical manipulation that reduces the dimensionality of a dataset. PCA was performed following established methodologies (Güler et al., 2002; Lambrakis et al., 2004; Daughney and Reeves, 2005; Cloutier et al., 2008; Daughney et al., 2012). The results of PCA are the eigenvalue of the principal components (an Eigenvalue > 1 is considered significant) and the variance in the data that each PC can describe (example Table A 3). PCA also provides principal component weightings or loadings of the considered variables (Table A 4). These indicate whether a parameter is significantly correlating with PC with highest correlation with eigenvectors closest to -1 or +1 (similar to conventional correlation techniques). A positive value indicates a positive correlation and vice versa for a negative correlation.

Table A 3: PCA results for Southland groundwater. PCA confirms one significant component (eigenvalue > 1) which explains 44% of the data variance. Component 2 and 3 are significantly less significant than component 1 (eigenvalue close to 1).

Component No.	Eigenvalue	Percent	Cumulative %
1	3.492	43.65	43.65
2	1.439	17.99	61.64
3	1.114	13.85	75.49

Table A 4: Eigenvectors for each of the variables considered within the PCA for Southland's groundwater indicating Northing and elevation are strongly positively correlating with PC1, Chloride and Sodium are strongly negatively correlating with PC1

Variables	Component 1	Component 2	Component 3
Northing	0.493	0.195	-0.086
Elevation	0.491	0.094	-0.029
Calcium	-0.161	0.400	-0.256
Chloride	-0.487	-0.067	0.055
Magnesium	-0.153	0.428	-0.606
Sodium	-0.444	-0.202	-0.080
Sulphate	-0.157	0.641	0.128
Potassium	-0.090	0.392	0.730

C2.2 HCA

HCA was applied as a complementary technique to understand the relationships amongst the hydrochemical variables and identify related samples. Hierarchical Cluster Analysis (HCA) is a multivariate statistical method used to group individual samples into distinct hydrochemical facies (Güler, 2002; Daughney, 2005; Rissmann et al., 2015). HCA works by quantifying the degree of similarity between waters in terms of their chemistry and unambiguously assigns different sites to distinct categories or 'clusters.' Because HCA is a purely data-driven approach, it operates without any assumptions about the geographical, hydrogeological or land use setting. The HCA algorithm automatically excludes any sites that are missing values for one or more of the specified input variables. Therefore, if too many variables are selected for HCA, the subset of sites considered may be very small and/or regionally biased. On the other hand, if too few variables are considered, the results may not provide an accurate perspective on groundwater quality or the processes that control it.

HCA was carried out in UNISTAT® 6.0 in two steps. In the first step a squared Euclidian distance and single linkage were used to identify any potential outlying data points. In the second step Wards linkage was used, as it has been found in previous studies to result in smaller and more distinct clusters (Guggenmos et al., 2011).

The square of the Euclidean distance (E) is a measure of similarity (Eqn. A3). Ward's method is based on an analysis of variance, and produces smaller distinct clusters than other linkage rules, such that each hydrochemical variable in a cluster is more similar to other variables in the same cluster than to any variable assigned to a different cluster.

$$E^2(i,k) = \sum_{j=1}^{n} (z_{ij} - z_{kj})^2 \qquad \text{(C3)}$$

where $z_{i,j}$ and $z_{i,k}$ represent the z-score for variable j at sites i and k. The summation is performed over all n variables included in HCA.

Results from HCA are presented as dendrograms, an example of which is shown in Figure A 4. In this application of HCA the terminus of each vertical line represents a single hydrochemical parameter (for clarity only four of the 32 parameters are labelled). The Y axis is a dissimilarity measure: parameters or groups of parameters are joined together by horizontal lines, and the position of any horizontal line indicates how similar (small values on the Y axis) or dissimilar (large values on the Y axis) the parameters or groups it joins actually are. Accordingly, the example in Figure 45 shows that the concentrations of Ca and Mg are strongly correlated in the dataset, as are the concentrations of Fe(II) and Mn(II); however, concentrations of Ca and Fe(II) are uncorrelated in the dataset. The horizontal red line represents a separation threshold at which the hydrochemical parameters are partitioned into five distinct clusters. A higher value for the separation threshold would result in recognition of fewer clusters, whereas a lower value would result in definition of more clusters. The actual value for the separation threshold must be selected by the analyst and is dictated by the aim of the investigation.

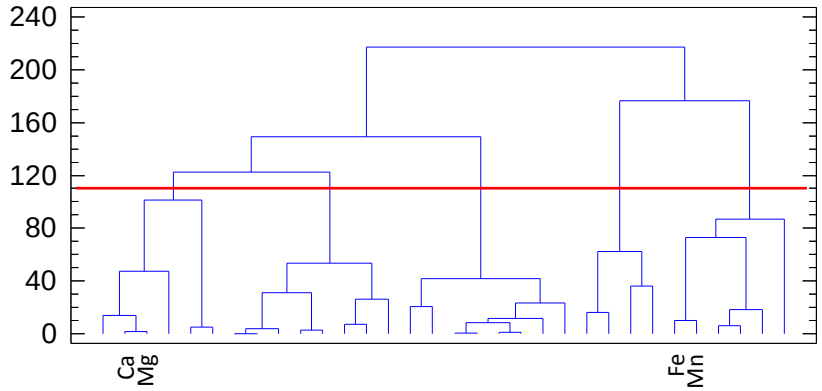


Figure A 4: Example dendrogram that might be produced by HCA when clustering hydrochemical parameters. For clarity only four of the hydrochemical parameters are labelled on the X axis.

Both PCA and HCA were carried out on z-scored log-transformed median hydrochemistry data. Median statistics per site per variable were generated in UNISTAT® 6.0. Median values were used as these provide a value that is less influenced by temporal variability within the data. Median data were log transformed (with the exception of pH) and z scores calculated (Güler et al., 2002; Guggenmos et al., 2011; Daughney, 2010; Rissmann et al., 2015). Therefore, the analysis does not directly consider the seasonal or episodic variation in water types that may exist as a result of changing flow conditions or associated contributing areas (Daughney, 2010).

Appendix D – Capture zone delineation for surface and groundwater sites

This appendix details capture zone delineation for surface and groundwater sites employed in this study.

D1 Surface water

Capture zones (also referred to as water sheds) for surface waters are derived from DEM (REC, River Environment Classification) derived catchment boundaries for all catchment area contributing (upstream) of the monitoring location (example in Figure A 5). Each site has a unique polygon, many will share perimeter location because they were derived from the same parent data (REC polygon) but differences will arise where the combination of polygons will have been edited so the polygon closes at the monitoring site location rather than the confluence with the next contributing sub catchment.

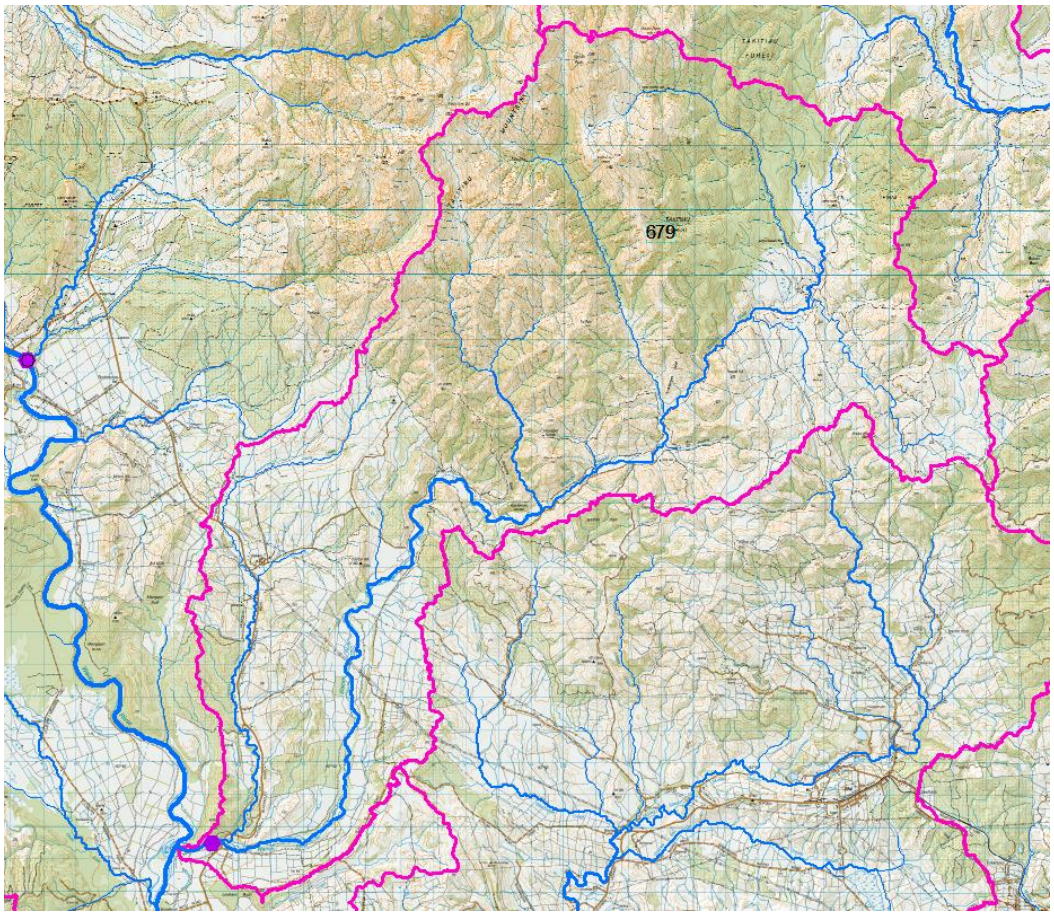


Figure A 5: example of capture zone/catchment area (pink outline) of a surface water site (pink dot) and stream network (blue lines)

D2 Groundwater

In this section, we describe how we delineated capture zones (CZs) for 173 unconfined groundwater sites of the 197 groundwater sites that were used for stratification and empirical modelling. The exclusion of groundwaters sites that tap into confined aquifers ensures that we only look at groundwaters that are connected to surface processes.

Methods for delineation of capture zones (CZs) are relatively well established for surface water, where one makes use of topographical data (see previous section for detail on the delineation of CZs

for surface waters). For groundwater (GW), currently no standard method for delineation of CZs exists. For simplicity often the properties of the entire aquifer are assumed to represent those of a GW's capture zone, but this can lead to misleading interpretations/predictions. Available methods for GW CZ delineation include (sorted from simplest to most complex):

1) Desktop review = arbitrary fixed radius or hydrogeological mapping, where the CZ is represented by a circle of fixed radius (Figure A 6). The radius is chosen based on expert knowledge, existing studies or guidelines (if applicable).

2) Manual methods, where the CZ is represented by a circle of calculated fixed radius or simplified variable shapes, or the CZ is estimated using a uniform flow equation method (**Error! Reference source not found.**). The extent (e.g. radius) is determined based on hydraulic/geologic/topographic boundaries.

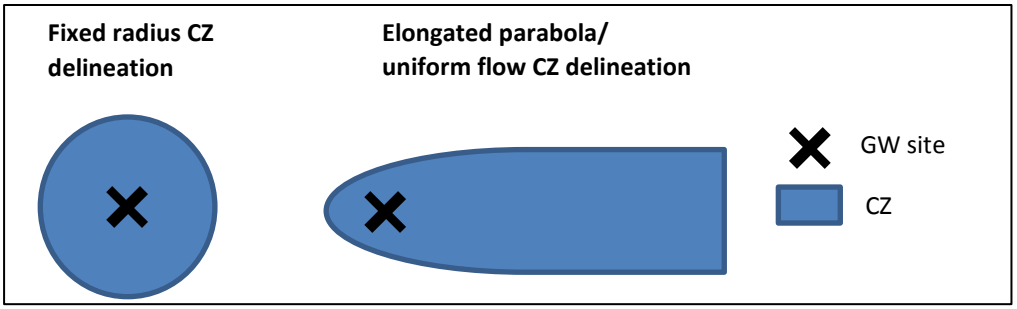


Figure A 6: Example of desktop review and manual delineation of groundwater capture zones (CZ).

3) AEM (analytical element modelling), where the CZ is delineated using analytical element modelling based on a AEM groundwater model.

4) Numerical models, where CZs are delineated using very detailed (3D) GW flow modelling.

The above methods require different model inputs (i.e. data) and resources, and provide CZ estimates of different levels of accuracy (Figure A 7, further detailed in Moreau et al., 2014 and 2014a). Generally, the least complex desktop review methods need the least data and resources, but also provide the least accurate GW CZs. The choice of model is therefore dependent on available data and resources and the required level of accuracy. To normalize GW CZ delineation in New Zealand, GNS has put some guidelines together (Moreau et al., 2014 and 2014a).

For the purpose of this study, we chose a manual method to delineate CZs for Southland GW wells. Specifically, we used the elongated parabola method assuming uniform flow and an ArcGIS tool developed by GNS Science (Toews et al., 2013 see http://www.gns.cri.nz/Home/Our-Science/Environment-and-Materials/Groundwater/Database-and-tools/Groundwater-capture-zone-GIS-toolkit). A CZ example established using the elongated parabola method is illustrated in Figure A 8. To use the method/run the ArcGIS tool, one requires the following data inputs:

1) Flow path of each well
2) Hydraulic conductivity
3) Saturated aquifer thickness
4) Hydraulic gradient
5) Pump or discharge rate
6) Effective porosity

Data of the above parameters are often lacking and need to be estimated as described subsequently.

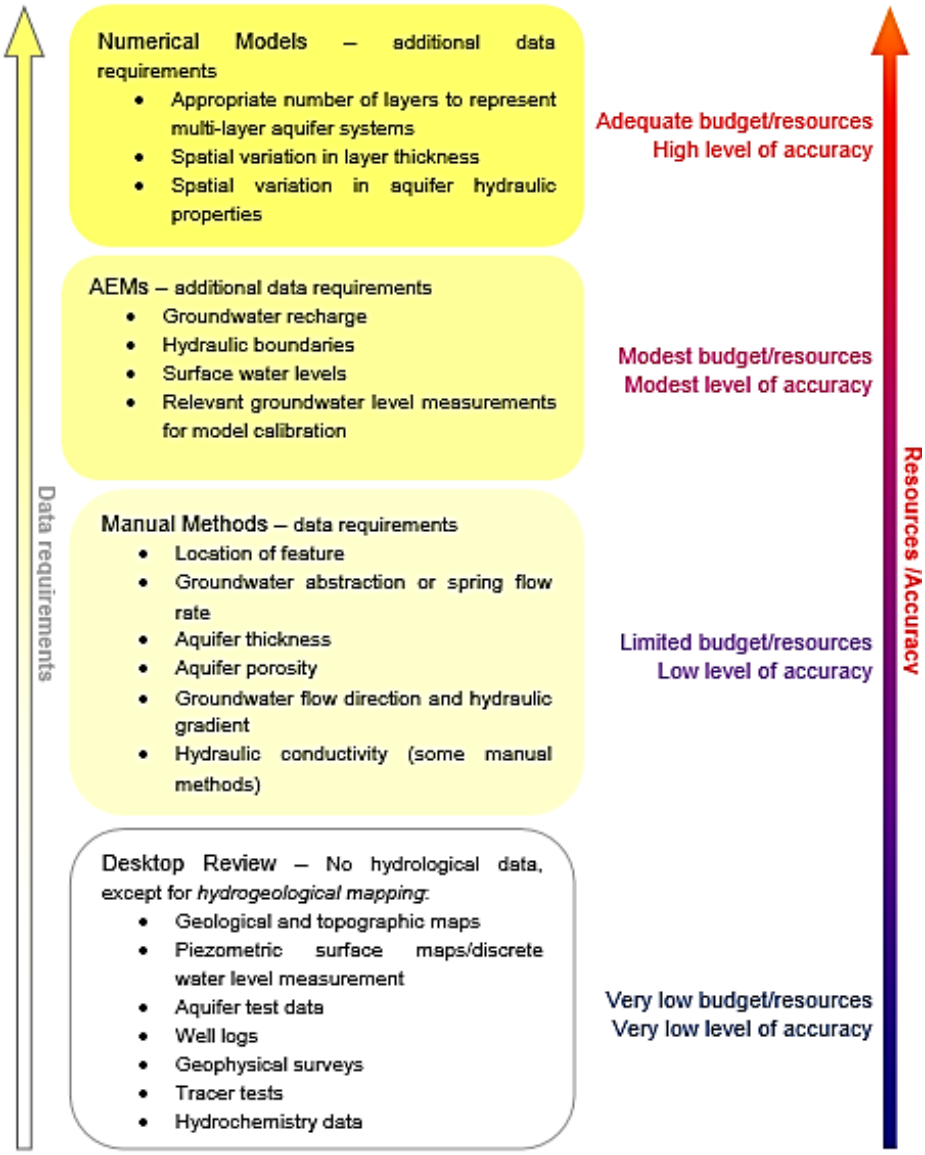


Figure A 7: data requirements, resources and accuracy of four capture zone delineation methods (Moreau et al., 2014)

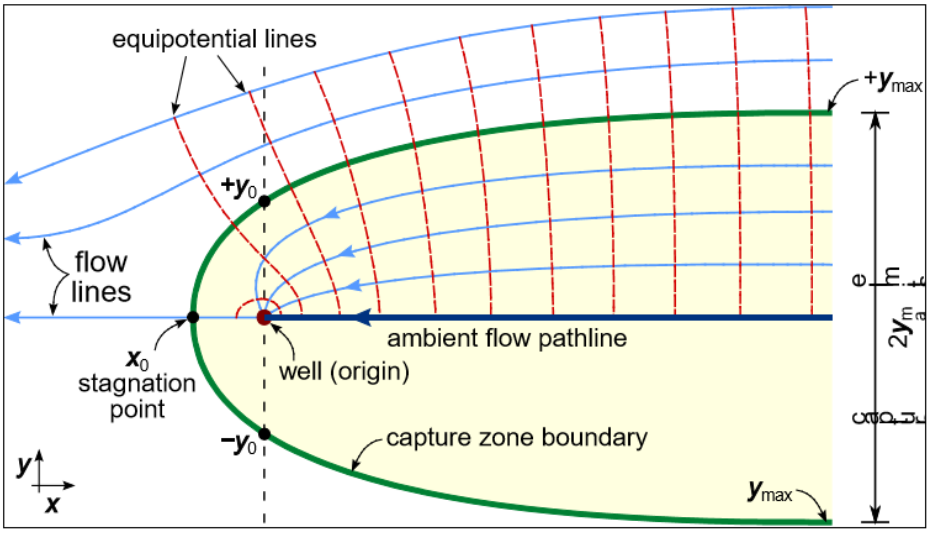


Figure A 8: CZ established using the elongated parabola method (Toews et al., 2013)

The groundwater flow paths of the wells were established using general flow direction provided by Brydon Hughes (an expert in groundwater). Normally, one would use Piezometric head contours to establish flow lines for each groundwater site (as illustrated in Figure A 9). However, this process is very time consuming as it is a very manual process. We chose to do a more straight forward and less time consuming process and assume that the delineated CZs can be used to assess the assemblage of drivers that impacts each groundwater well reasonably well. For further study, we suggest delineation of GW CZ using flowlines established through head contour interpretation.

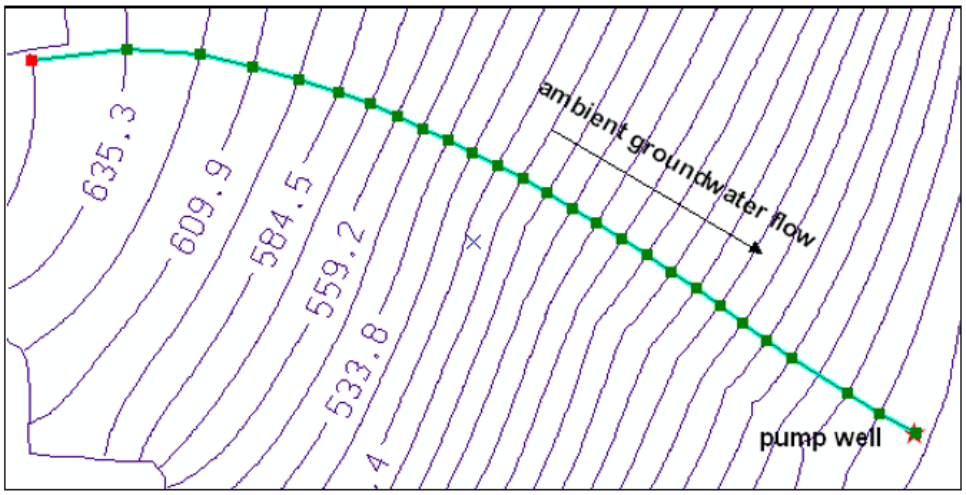


Figure A 9: Creation of a GW flowline perpendicular to the groundwater head contours {Toews et a., 2013].

Hydraulic conductivity, porosity and saturated aquifer thickness have been estimated for each groundwater management zone as presented in Gusyev et al, (2011), Morgan and Evans (2003), SKM (2005); Burbery et al. (2013), Phreatos (2007) and Hughes (2013), summarized in Table A 5.

The hydraulic gradient of a well is normally estimated from the length of the flow line and the head difference between the well and starting point of the flowline (Eqn. D1). However, in our case piezometric head data were not available. We assume the hydraulic gradient for all wells is 0.01 m/m (as a reasonable estimate, person. Comm. Brydon Hughes)

$$\Delta h = {h_1 - h_2}/{l} \qquad \text{(D1)}$$

where l is the length of the flow line, h1 is the head at the start of the GW flowline (usually aquifer boundary), h2 is the head at the end of the GW flowline (the GW well).

Table A 5: Estimated hydraulic conductivity (K_H), porosity (n), aquifer thickness (d) and abstraction rate (Q) for Southland aquifer zones

Aquifer	Average thickness [m]	porosity	porosity effective	K_H Hydraulic conductivity [m/d]
Southland Fm	10	0.2	0.15	10
Forest Hill Fm	10	0.2	0.15	10
Pomahaka Fm	10	0.2	0.15	10
Makarewa	20	0.2	0.15	10
Middle GLM	20	0.2	0.15	10
lower Oreti	15	0.2	0.15	15
Central Plains	15	0.2	0.15	15
Knapdale	10	0.2	0.15	25
Orepuki	10	0.2	0.15	25
Longridge	20	0.2	0.15	15

Waihopai	20	0.2	0.15	15
Waimea Plain	10	0.2	0.15	30
Upper GLM	30	0.2	0.15	10
Winton Hill Fm	30	0.25	0.19	10
Wendonside	20	0.2	0.15	15
Chatton Fm	35	0.25	0.19	10
Waimatuku	20	0.2	0.15	20
lower Aparima	15	0.2	0.15	30
Castlerock	15	0.2	0.15	30
Catlins	50	0.1	0.05	10
Hokonui	50	0.1	0.05	10
Upper Aparima	10	0.2	0.15	50
Five Rivers	10	0.15	0.12	50
lower Mataura	10	0.2	0.15	50
Te Anau	10	0.2	0.15	50
Wendon	20	1.2	0.92	25
lower GLM	50	0.2	0.15	10
Oreti	10	0.2	0.15	50
lower Waiau	20	0.2	0.15	40
Cattle Flat	10	0.2	0.15	100
Upper Mataura	15	0.2	0.15	100
Croyden	10	0.2	0.15	150
Edendale	15	0.2	0.15	100
Tiwai	25	0.2	0.15	75
Waipounamu	15	0.2	0.15	250
Riversdale	20	0.2	0.15	200

Pump or discharge rate data are not available (only available if groundwaters are consented). We therefore assumed worst case, i.e. that the wells have a maximum pump rate of 20L/s (limit for consent application).

Following the delineation of hydromorphic capture zones (HCZ) for each groundwater site it was possible to constrain the following hydromorphic setting: (i) groundwater zone (Te Anau, Waihopai); (ii) aquifer type (i.e., riparian, terrace etc); (iii), aquifer confinement (unconfined, semi confined, etc), (iv) up-gradient proximity to a given geomorphic domain (i.e., Alpine, Bedrock 1, Bedrock 2 and Lowland); (v) up-gradient proximity to an alpine, bedrock and/or lowland sourced stream and associated stream order, and (vi); soil hydrology (i.e. drainage class). Steps i – iii were intersected with the existing framework of Hughes (2003, 2016); iv – v through geomorphic domain and REC, and; vi through the TopoClimate South Soil Map.

Although these relatively simplistic 'capture zones' help constrain the general hydromorphic setting of a groundwater site they are still crude relative to evolved capture zone analysis (e.g. developed using complex groundwater models). Limitations include:

- 1) it is likely that we over- or under-estimate the size of the capture zone and
- 2) the shape of the capture zone may differ from an elongated parabola.

Despite these limitations, we assume that approach meets the purpose of this study, i.e. the method can give reasonable estimates on the proportions of geology/soil types that groundwater interacted with.

From here on, groundwater capture zones are referred to as hydrodynamic capture zones (HCZ) so as to not be confused with more evolved capture zone analysis that factors in aquifer properties, pumping rate and flow line analysis.

Appendix E – Groundwater clustering

We initially used HCA clusters defined in Daughney et al. (2015). However, the clustering by Daughney et al. (2015) was run on a combined surface and groundwater data set, and was not specific to groundwater and did not provide sufficient resolution over groundwater origin and consequently hydrochemical setting. The poorer resolution of HCA clusters for groundwater reflects the combination of surface and groundwater data set for the HCA run and the use of only 10 variables (hydrochemical analytes) (Daughney et al., 2015). For these reasons we reclustered the groundwater data separately and included an additional 5 variables including the stable isotopes of water (Variables Selected: pH, Cl, Cl/Br, TON, TN, Fe(II), Mn(II), Na, K, Ca, Mg, SiO_2, SO_4, Alk(total), $\delta^{18}O\text{-}H_2O$). Clustering was carried out using HCA, Euclidian distance and Wards linkage method on log transformed z-scored data.

Reclustering of 193 groundwater sites with a greater number of key variables identified 6 general groundwater clusters with distinct characteristics that correspond to 6 general hydrochemical settings for regional unconfined aquifer systems across Southland. Inclusion of $\delta^{18}O\text{-}H_2O$ was particularly important as it provided greater resolution around water source and recharge altitude and hence cluster membership than solutes alone. Assessment of drivers against new cluster memberships gave much clearer, hydrochemically/genetically sensible results and a good to strong, fitted miss-classification rate (see section 4.3.1.2 Main Report).

The HCA dendrogram is depicted in Figure A 10. We identified 6 clusters (Figure A 11). The cluster characteristics are detailed in the following and summarised in Table A 6.

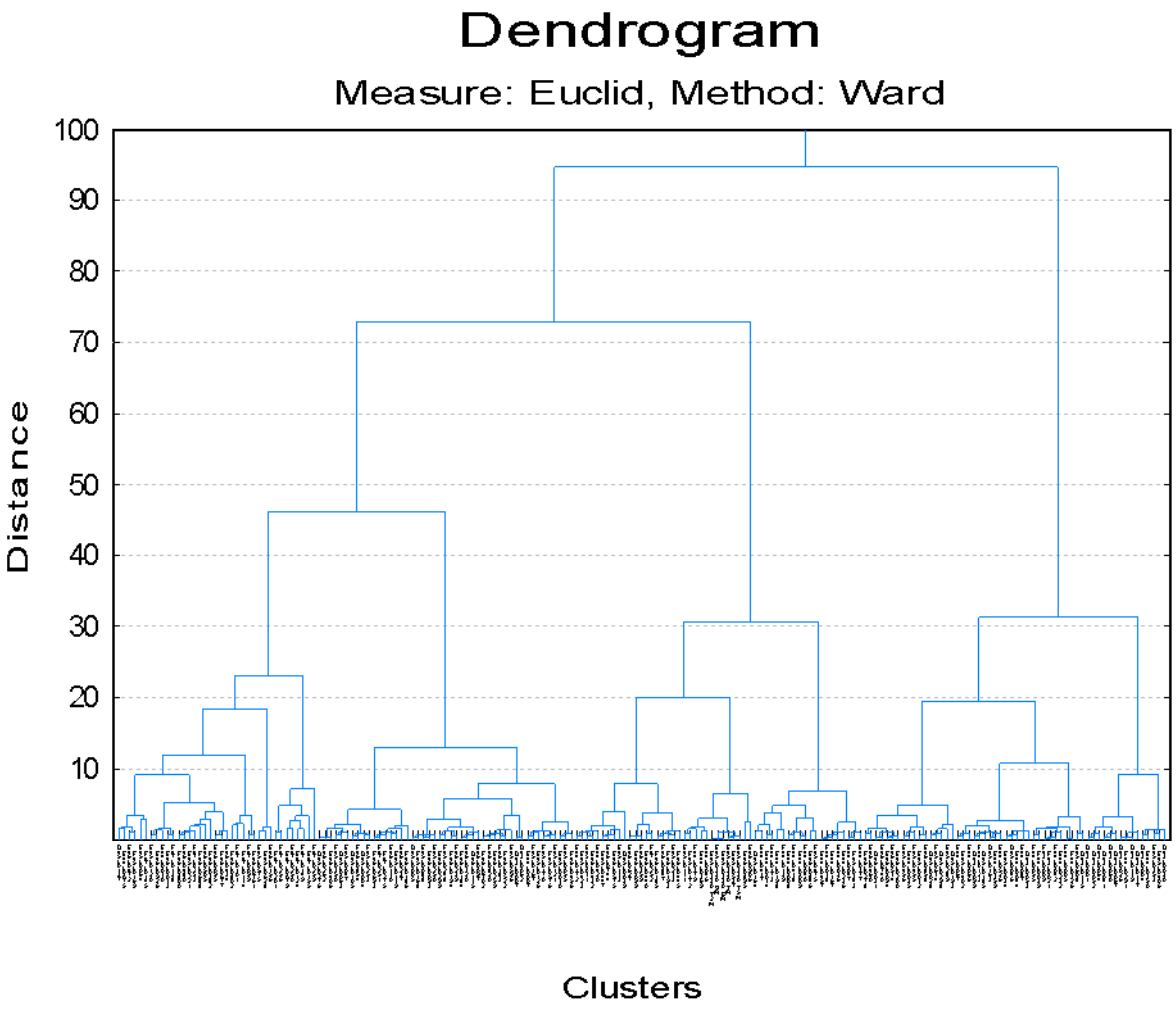


Figure A 10: dendrogram of HCA clustering of groundwater data using 26 analytes (Variables Selected: pH, Cl, Cl/Br, TON, Alk(total), Ca, Na, K, TON, $\delta^{18}O$-H_2O, SO_4, Mn(II), Fe(II), TN), Euclidian distance and Wards linkage method. Phenon lines for 6 and 4 clusters are shown.

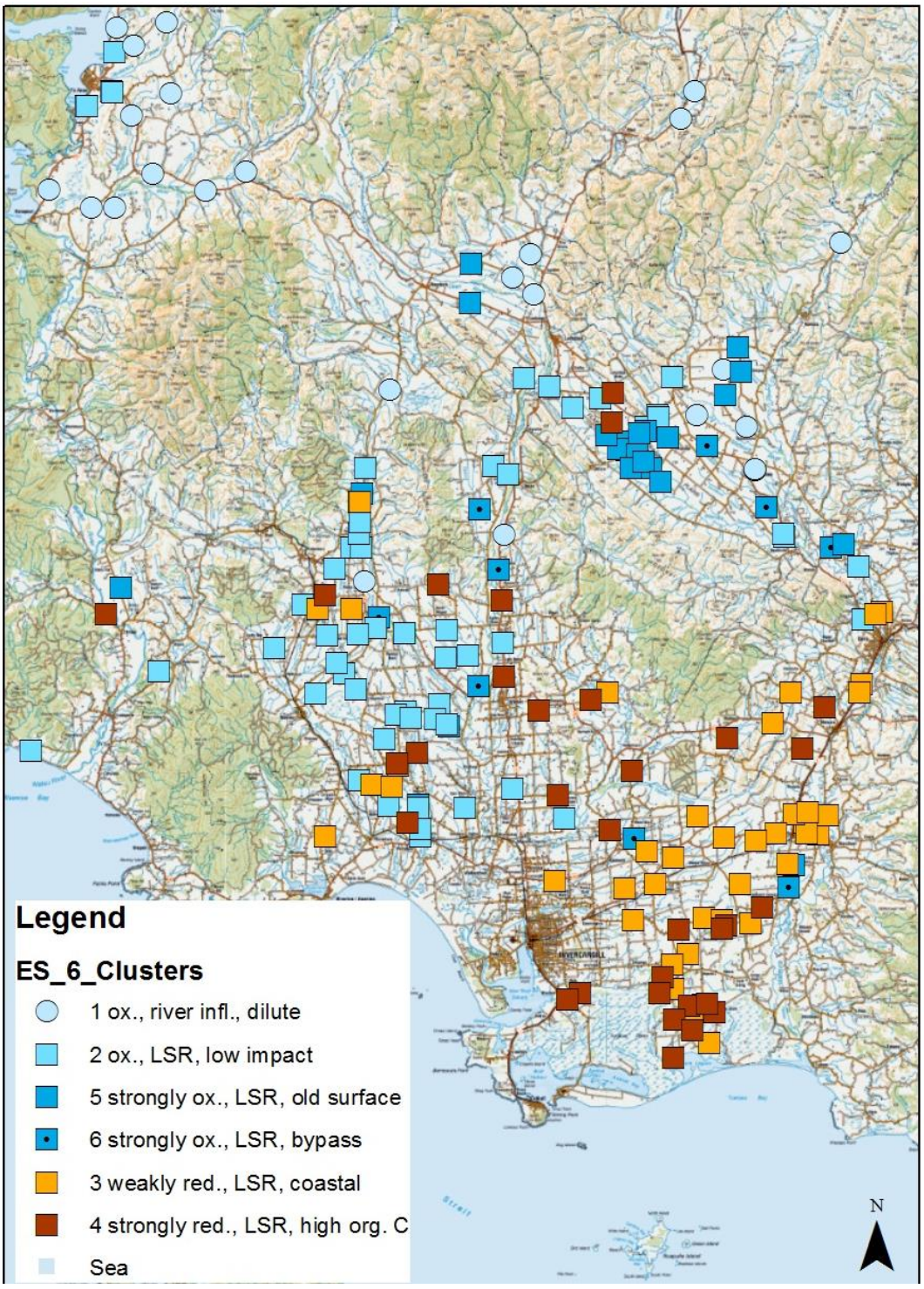


Figure A 11. Spatial occurrence of 6 clusters (regional hydrochemical signatures) for Southland groundwaters (unconfined).

The 6 general hydrochemical settings for regional unconfined aquifer systems resolved from HCA correspond to 6 regionally significant hydrochemical settings for regional unconfined aquifers across Southland and are characterised as:

Cluster 1 = Oxidising, River Influenced, Dilute,

Very dilute waters, predominantly oxic redox state, northern and southern groundwaters, occur in riparian aquifers, inland or alpine ppt, derived from high altitude, primarily alpine river recharge. Low EC, oxidising; low alkalinity, low TON, Ca-HCO_3 and Mg-HCO_3 waters of low Cl; very high D.O. (n = 25).

Cluster 2 = Oxidising, Land Surface Recharge, Low Impact

Groundwaters of elevated conductivity and mod-low TON (n = 56) occurring across northern eastern and south central (west of Oreti) Southland in association with areas of reducing soils. High TOC and large coefficients of variation for time series redox sensitive species indicative considerable temporal variation in redox status. Dominantly, Ca-HCO_3 waters.

Cluster 3 = Weakly Reducing, Land Surface Recharge, Na-Cl waters.

Mainly southeastern, Kamahi and Waikiwi Tce Formations with fewer south-western occurrences. Slightly elevated Fe(II) and TOC and low TON consistent with weak reduction in soil zone and/or aquifer. Predominantly Na-Cl waters (n = 43).

Cluster 4 = Strongly Reducing, Land Surface Recharge, with high Organic Carbon and/or Carbonate

Most prevalent at distal end of Mataura Catchment in conjunction with marine terraces, peat wetlands and variably lignite measures. Also occurring in places across southeast of Mataura River to Coastal Longwoods. Few occurrences in northern Southland. These waters have very low TON, high dissolved Fe, relative elevated TAN (ammonification of organic matter) and organic carbon (n = 32).

Cluster 5 = **Strongly Oxidising, Land Surface Recharge, Old Surfaces**

Occurring predominantly across northern Southland in association with old remnant surfaces including Balfour, Wendonside and Knapdale areas with a long history of elevate groundwater nitrate. Few occurring in Five Rivers, Lower Waiau and Central Plains. Groundwaters characterised by very high TON (median 13.5 mg/L) but low SO_4 due to anion retention in aluminium oxide rich soils associated with the oldest geomorphic surfaces in Southland (n = 19). Mainly, Na-Cl waters.

Cluster 6 = Strongly Oxidising, Land Surface Recharge, Bypass

Occurring across north-eastern and southern lowland these groundwaters are associated with areas of bypass flow. Waters are strongly oxidising with strongly elevated TON, Ca, SO_4, K (n = 18) consistent with soil zone contamination due to natural vertical bypass.

An ability to estimate the general hydrochemical setting for groundwater is important as redox in conjunction with water source and recharge mechanism are the two chief determinants of water quality outcomes (variation) in areas of intensive landuse in Southland and New Zealand (TC1-8 and Daughney, 2005; Daughney and Reeves, 2005; Daughney et al., 2010; Rissmann et al., 2012; Daughney et al., 2015). As defined above, general hydrochemical setting is a factor of recharge mechanism and water source and includes discrimination between river influence and land surface recharged aquifers and redox setting. Modifiers such as geomorphic age and substrate composition

along with variation and marine aerosol load introduce extra resolution of hydrochemical metrics such as major ion facies (i.e., whether a water is a Ca-HCO_3 or Na-Cl water), which although useful for understanding water provenance are not essential for explaining spatial variation in water quality. Accordingly, a more generalised set of 3 broad hydrochemical settings based purely on recharge and redox were resolved and are defined as:

1. Oxidising, Land Surface Recharge (n = 93): clusters 2,5,6
2. Reducing, Land Surface Recharge (n = 75): clusters 3,4
3. Oxidising, River Influenced (n= 25): cluster 1

Centroid statistics for the three general hydrochemical settings for Southland's unconfined aquifer systems are presented in Table A 7 and depicted spatially in Fig. A 12. The pattern of regional hydrochemical settings for Southland's unconfined aquifer systems shows dominance by oxidising, river-influenced groundwaters across Southland's northern, inland basins (Te Anau, Fiver Rivers, Upper Mataura, Upper Waikaia and parts of the Waimea valley that are hydrologically connected to alpine rivers) and associated alpine range fronts as well as riparian aquifers adjacent to main stem rivers. Land surface recharged groundwaters showing signs of reduction occur predominantly in the southeast, south of the Waimea Valley and across the historical floodplain of the Matura River (Kamahi Formation and associated reworked surfaces). Reducing groundwaters occur interspersed with oxidising groundwaters across south-western Southland and with minor occurrences across northern Southland. Land surface recharged groundwaters that are oxidising occur across both northern and southern Southland but are consciously absent across large areas of southeastern Southland. This classification provides a regional spatial hydrochemical framework for Southland's unconfined aquifer systems.

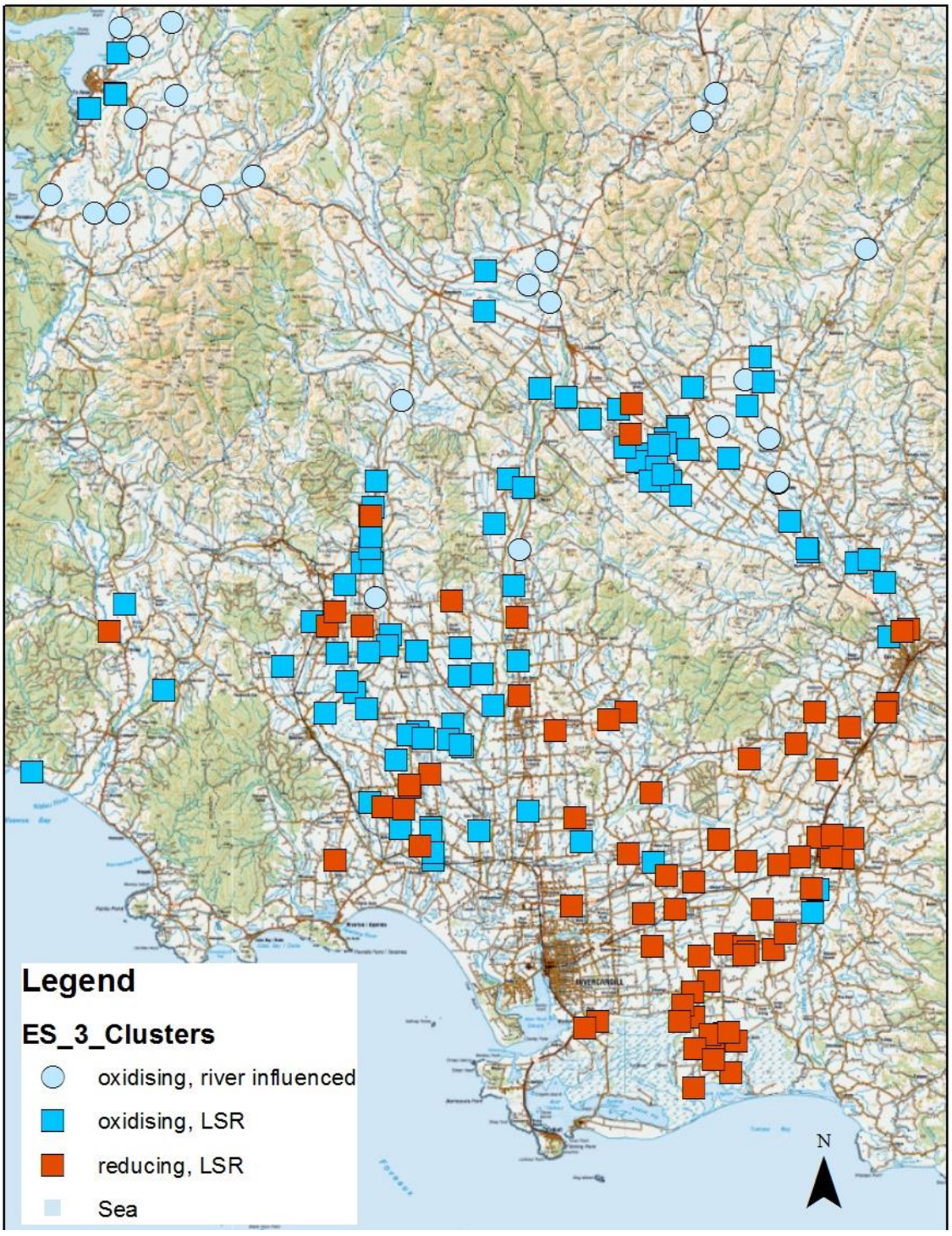


Figure A 12. General hydrochemical settings for Southland unconfined groundwaters as defined from redox and recharge signatures in hydrochemistry.

Table A 6. Centroid statistics for hydrochemical variables (analytes) by groundwater cluster

Cluster	Statistic	pHField	DOField	Cond	Cl	Cl/Br	TON	Mn^{II}	Fe^{II}	SO_4	TAM	DOC	TOC	AlkTot	$AlkHCO_3$	Ca	K	SiO_2	Na	Mg	TN	DRP	δ^{13}C-DIC	$\delta^{18}O$-H_2O	$\delta^{2}H$-H_2O
1	Geometric Mean	6.5	5.7	106.0	4.4	147	1.6	0.001	0.07	3.6	0.02	0.8	1.8	33.5	33.4	9.6	0.7	14.4	5.2	3.7	1.6	0.007	****	****	****
2		6.6	5.2	221.6	18.6	270	2.5	0.002	0.11	6.4	0.05	1.2	9.4	58.2	57.7	16.1	0.9	26.1	15.8	8.0	2.7	0.017	****	****	****
3		5.9	4.4	188.0	25.2	228	1.1	0.007	0.32	7.1	0.03	1.6	8.3	29.7	29.5	10.0	1.1	18.4	19.9	3.6	1.8	0.007	****	****	****
4		6.5	0.5	277.1	28.2	238	0.1	0.199	3.24	6.6	0.12	1.6	7.3	74.7	73.2	14.6	1.0	31.7	23.2	6.7	0.5	0.018	****	****	****
5		6.0	9.3	206.3	16.8	173	12.2	0.001	0.12	2.4	0.01	0.8	8.3	19.9	19.6	11.9	1.0	17.8	15.3	5.3	11.1	0.016	****	****	****
6		6.1	6.1	316.7	28.7	300	12.0	0.005	0.06	24.7	0.04	1.6	5.1	33.7	33.4	23.0	1.3	18.1	18.3	9.9	11.3	0.011	****	****	****
1	Median	6.4	7.1	115.0	4.4	131	2.2	0.001	0.07	3.5	0.02	0.9	0.9	34.0	34.0	9.8	0.8	14.9	5.9	4.0	2.0	0.009	-20.6	-9.7	-69.0
2		6.5	6.7	222.8	21.0	271	3.1	0.002	0.09	7.0	0.03	1.2	11.9	58.0	58.0	15.6	0.9	25.5	17.8	7.7	3.6	0.019	-20.6	-8.2	-56.6
3		5.8	6.7	190.1	25.0	234	2.2	0.007	0.23	8.3	0.03	1.6	10.6	29.1	29.0	9.3	1.0	18.9	18.2	3.5	1.8	0.007	-19.8	-7.4	-49.5
4		6.4	0.7	277.2	29.5	250	0.0	0.203	4.68	6.2	0.11	1.5	10.1	72.1	70.5	14.6	1.1	32.0	24.0	7.0	0.4	0.028	-19.8	-7.5	-50.4
5		5.9	9.5	197.2	17.7	177	13.5	0.002	0.12	1.9	0.01	0.8	8.2	18.8	18.3	10.7	1.0	18.1	16.0	5.1	12.4	0.016	-21.4	-8.4	-58.0
6		6.1	7.2	318.7	27.3	304	12.8	0.005	0.07	24.6	0.02	1.7	4.0	33.4	32.7	23.6	1.2	18.8	19.0	10.2	13.1	0.011	-22.4	-7.7	-53.6

Data reported in mg/L for chemical species, µS/cm for conductivity, and δ (permil) for stable isotopes relative to relevant standards (i.e., V-PDB for δ^{13}C-DIC, V-SMOW for $\delta^{18}O/\delta^{2}H$ of H_2O). Where TAM = Total Ammoniacal Nitrogen.

Table A 7. Centroid statistics for hydrochemical variables (analytes) by general hydrochemical setting

General Hydrochemical Setting	Statistic	pHField	DOField	Cond	Cl	Cl/Br	TON	Mn^{II}	Fe^{II}	SO_4	TAM	DOC	TOC	AlkTot	$AlkHCO_3$	Ca	K	SiO_2	Na	Mg	TN	DRP	δ^{13}C-DIC	δ^{18}
Oxidising LSR (n = 93)	Geometric Mean	6.3	234.0	6.1	19.8	252	4.7	0.002	0.10	6.8	0.04	1.2	8.2	42.1	41.6	16.2	1.0	22.5	16.2	7.7	4.7	0.015	****	*
Oxidising, River Influenced (n = 25)		6.5	106.0	5.7	4.4	147	1.6	0.001	0.07	3.6	0.02	0.8	1.8	33.5	33.4	9.6	0.7	14.4	5.2	3.7	1.6	0.007	****	*
Reducing LSR (n = 75)		6.2	221.9	1.7	26.4	232	0.3	0.029	0.86	6.9	0.06	1.6	7.9	44.0	43.4	11.7	1.1	23.2	21.3	4.7	1.1	0.010	****	*
Oxidising LSR	Median	6.3	234.0	7.5	21.0	253	7.1	0.002	0.10	7.1	0.02	1.2	8.2	47.0	47.0	17.0	0.9	20.4	16.9	7.5	6.1	0.016	-21.3	-
Oxidising, River Influenced		6.4	115.0	7.1	4.4	131	2.2	0.001	0.07	3.5	0.02	0.9	0.9	34.0	34.0	9.8	0.8	14.9	5.9	4.0	2.0	0.009	-20.6	-
Reducing LSR		6.2	216.6	3.7	26.0	238	0.7	0.036	0.85	6.9	0.05	1.6	10.4	43.0	43.0	12.4	1.1	24.0	21.0	5.0	1.0	0.010	-19.8	-
Oxidising LSR	Coefficient of Variation	0.1	0.3	0.4	0.5	0.4	0.8	2.1	1.3	0.9	3.3	0.6	0.8	0.5	0.5	0.4	0.4	0.4	0.3	0.3	0.7	0.7	-0.1	-
Oxidising, River Influenced		0.1	0.3	0.4	0.5	0.5	0.7	1.7	0.6	0.7	1.1	0.6	1.3	0.3	0.3	0.4	0.4	0.3	0.4	0.4	0.7	1.2	-0.1	-
Reducing LSR		0.1	0.3	0.8	0.4	0.3	1.4	1.5	1.7	0.8	2.2	1.3	0.7	0.8	0.8	0.8	0.4	0.4	0.4	0.5	1.1	1.7	-0.1	-

Data reported in mg/L for chemical species, S/cm for conductivity and δ (permil) for stable isotopes relative to relevant standards (i.e., V-PDB for δ^{13}C-DIC, V-SMOW for $\delta^{18}O/\delta^{2}H$ of H_2O). Where TAM = Total Ammoniacal Nitrogen.

Appendix F – further detail on validation of conceptual model

F1 Statistical Modelling

This section details the validation of the conceptual model through statistical modelling.

F1.1 Groundwater results

Categorical responses

The various ground water groupings, including the four and six class levels of the ground water HCA, the redox clusters and the general categorisations (**Error! Reference source not found.**) comprised clusters with an average of between 32 and 64 sites per cluster with generally at least 15 sites in a class, with the exception of the General Redox Category that had a class with only three sites (**Error! Reference source not found.**A 8).

Table A 8: Details of the categories and sites for the 5 groundwater categorical responses. The first value in the number of sites per category is the mean the minimum and maximum are shown in the parentheses.

Categorical response	Number of categorical levels	Number of sites per category
Cluster4	4	48 (25 - 88)
Cluster6	6	32 (18 - 56)
General Redox Category	4	48 (3 - 107)
General Category1	6	32 (18 - 56)
General Category2	3	64 (25 - 93)

The sites varied appreciably between clusters in terms of their hydro-chemical characteristics (Table A 9). For example, relative to the other classes, class 4 of the six class level of the HCA had high conductivity, chloride and sodium, low dissolved oxygen and total organic nitrogen and high values of d18OH2O and d2HH2O (**Error! Reference source not found.**).

Table A 9 Characteristic values of selected hydro-chemical variables in each class defined by the eight class level of the ground water HCA. The values are the mean of the median values of the sampling sites in each class.

Class	Cond	Cl	Na	DOField	TON	DOC	d18OH2O	d2HH2O
1	112	5	6	6	2	0.99	-10	-68
2	231	21	17	6	4	1.46	-8	-56
3	193	27	21	6	3	2.82	-7	-49
4	286	30	24	2	0	2.59	-8	-52
5	210	18	16	9	13	0.84	-8	-57
6	322	30	19	7	13	1.68	-8	-53

The reduced random forest models of the classifications had a fitted miss-classification rates of between 18% and 28% and fitted Kappa values ranged from 0.4 to 0.7, indicating the models had fair to good fits (Table A 10). Out-of bag (OOB, i.e. independent predictions) miss-classification rates

ranged between 31% and 50% and OOB Kappa values ranged between 0.2 and 0.5. These results indicate consistency in the relationships between drivers and classes but a weak ability the predict class for a new (i.e. independent data) as a function of the drivers.

The reduced models of the categorical responses retained between eight and 13 predictors whose importance is shown in **Error! Reference source not found.**Table A 10. This indicates that membership of classes was significantly associated with several predictors and that multiple drivers (i.e. driver assemblages) are associated with particular hydro-chemical outcomes.

Table A 10. The order of importance of the predictors used in the RF classification models of the five categorical response variables and model performance. The values associated with the predictors indicate the rank order of importance. NA values indicate the predictor was not included. Model performance is the out-of-bag (OOB) misclassification rate and the misclassification rate of the fitted model and the OOB and fitted Kappa values.

Response	Predictors													Model Performance			
	CRPdomain	PPTSource	SubSurface	SoilRP	GeoRP	Bypass	GeomorphicAge	Recharge	SurfaceC	AquiferType	RiverInfluence	RiverConnectivity	VerticalBypassRedox	OOB Misclassification	Fit Misclassification	OOB Kappa	Fit Kappa
Cluster 4	3	1	2	6	-	-	4	-	7	-	5	-	8	40	26	0.4	0.6
Cluster 6	3	1	2	6	-	-	4	-	5	7	9	8	10	50	22	0.4	0.7
General Redox Category	1	-	2	-	-	-	5	-	3	4	-	-	6	42	28	0.2	0.4
General Category 1	3	1	2	7	-	-	4	-	5	6	9	8	10	50	22	0.4	0.7
General Category 2	-	1	2	-	-	-	4	-	6	3	7	5	8	31	18	0.5	0.7

The partial dependence plots for each model indicate many associations between driver gradients and hydro-chemical outcomes that are consistent with prior expectations indicated by the conceptual model. An example of partial dependence of Class 4 of the six-class HCA classification (Cluster 6) on the 6 most important predictors is shown in **Error! Reference source not found.**A 13. The hydro-chemical characteristics of class 4 are shown in **Error! Reference source not found.**A 10. The plots show the effect of the 6 most important predictors (X-axis) on the probability a site belongs to Class 4 (Y-axis). The values on the Y-axis of these plots are logit transformed probabilities.

They should be interpreted as the marginal effect of the predictor (here always a category) on the response (which here is the logit transformed probability of belonging to a particular class).

Error! Reference source not found. A 13's top left plot indicates there is a marked difference in the probability of membership associated with the 5 category predictor precipitation source (PPTSource). This indicates that, all other things being equal, a site has a low probability of belonging to Class 4 if PPTSource is in the *Alpine* and *Alpine2* categories. And that a site has a higher probability of belonging to Class 4 if PPTSource is *Coastal* or *Coastal2* category. Similarly, a site has lower probability of belonging to Class 4 if SubSurface is *UndiffClasstics*, GeomorphicAge is *Q1,* SoilRP is *Low* and CRPdomain is *Low/Low.*

Where predictor categories have similar values this indicates that there is little difference in the response probability for these different values of the predictor. For example, in Figure A 13, the probability of membership of Class 4 associated with the predictor "PPTSource", indicates that membership probability is similar when PPTSource is *Alpine*, *Alpine2*.

Partial plots for all classes of each of the five groundwater categorical responses are supplied separately with the following naming convention "GWPartialPlotsCatResponse_*Classification_-me*.pdf" (e.g., "GWPartialPlotsCatResponse_Cluster6.pdf").

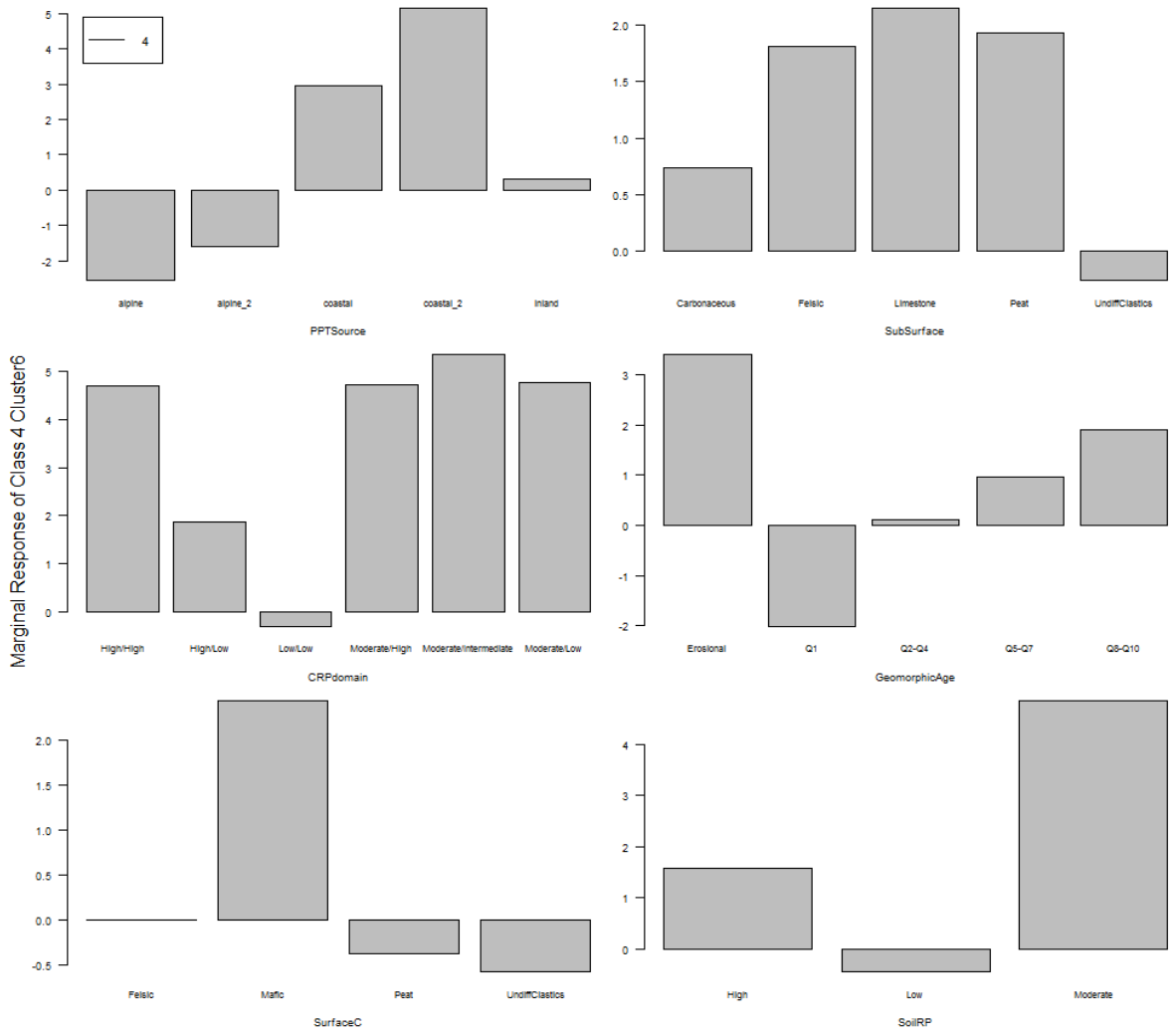


Figure A 13. Partial Dependence of Class 4 of the six-class HCA classification of the groundwater sites on the 6 most important predictors. The bars indicate the marginal effects of each variable's category on the class logit transformed probability.

Continuous groundwater hydro-chemical variables

The performance of the models fitted to 23 individual hydro-chemical continuous variables varied from reasonable (e.g r^2> 40%) to poor (r^2 < 10%) (**Error! Reference source not found.** A 11). The number of predictors included in the reduced models for individual groundwater hydro-chemical variables and the order of importance of the predictors varied from one (e.g. NH4 and DRP; **Error! Reference source not found.**A 11) to 13 (e.g. TON, TN, AlkTot and AlkHCO3; **Error! Reference source not found.**A 11) depending on the variable (**Error! Reference source not found.**A 11).

Examples of partial dependence two hydro-chemical variables (Cl and TON) are shown in **Error! Reference source not found.** and **Error! Reference source not found.**15. The plots show the effect of the 6 most important predictors (X-axis) on the absolute value of the response variable (Y-axis). They should be interpreted as the margi-l effect of the predictor (here always a category) on the absolute value of the response.

The value of the response variable Cl varies over all sites from 0.9 to 41 with a mean value of 19.6 (**Error! Reference source not found.**11). **Error! Reference source not found.**A 14's top left plot indicates there is a marked difference in the value of Cl associated with the 5 category predictor PPTSource. This indicates that, all other things being equal, a site has a Cl value of greater than 25 if PPTSource is in the *Coastal* or *Coastal2* category. If PPTSource is the *Alpine* and *Alpine2* categories, all other things being equal, values of Cl are less than 20. Similarly, a site has higher values of Cl if RiverConnectivity is *Low*, AquiferType is *Lowland* and VerticalBypassRedox is *High/Low.*

The value of the response variable TON varies over all sites from 0 to 6.5 with a mean value of 1.3 (**Error! Reference source not found.**A 11). **Error! Reference source not found.**A 15**Error! Reference source not found.**'s top left plot indicates there is a marked difference in the value of TON associated with the 5 category predictor PPTSource. This indicates that, all other things being equal, a site has a TON value of greater than 5 or 6 if PPTSource is in the *Coastal* or *Inland* categories respectively. If PPTSource is the *Alpine* and *Alpine2* categories, all other things being equal, values of TON are less than 4. Similarly, a site has higher values of TON if SoilRP is *Low*, GeomorphicAge is *Q2-Q4* and RiverConnectivity is *Low.*

Partial plots for all of the 26 individual hydro-chemical variables are supplied separately in the file "GWPartialPlots_ContinuousVariables.pdf"

Table A 11. Order of importance of the individual predictors included in the reduced random forest models for the 23 individual hydro-chemical continuous variables and OOB and fitted performance (r^2) of the models.

Response	Predictors														
	CRPdomain	PPTSource	SubSurface	SoilRP	GeoRP	Bypass	GeomorphicAge	Recharge	SurfaceC	AquiferType	RiverInfluence	RiverConnectivity	VerticalBypassRedox	**OOB r^2**	**Fitted r^2**
pHField	2	1	3	-	-	-	5	-	4	6	8	7	9	15	17
Cond	-	2	3	-	-	-	6	9	5	4	7	8	1	31	31
Cl	-	1	-	-	-	-	-	-	-	2	-	4	3	50	51
Cl_Br	-	1	-	-	-	-	-	-	2	-	-	-	3	26	26
DOField	3	1	2	5	13	10	6	11	7	4	8	9	12	6	9

TON	3	1	2	4	12	11	6	13	9	8	5	10	7	26	26
Mn	-	-	-	-	-	-	-	-	-	-	-	1	-	0	0
Fe	1	2	-	-	3	7	5	-	4	6	-	8	9	24	25
SO4	2	1	4	6	9	12	3	13	8	5	10	11	7	19	19
NH4	-	1	-	-	-	-	-	-	-	-	-	-	-	-2	1
DOC	3	1	2	5	4	11	6	13	7	10	8	12	9	6	11
TOC	-	1	-	-	-	-	-	-	-	-	-	-	-	11	12
AlkTot	1	-	3	-	-	-	-	-	2	4	5	6	-	15	16
AlkHCO3	2	3	1	5	13	10	6	12	4	7	8	9	11	11	14
Ca	-	-	2	-	6	-	3	-	1	-	-	4	5	19	19
K	-	-	-	-	-	-	-	-	1	-	-	-	-	6	6
SiO2	1	2	-	-	-	-	-	-	-	-	3	-	-	24	24
-	-	1	-	-	-	-	-	-	-	-	-	-	-	45	45
Mg	-	3	1	-	-	5	4	-	2	6	7	9	8	27	27
TN	3	1	2	4	13	11	7	12	8	6	5	10	9	29	30
DRP	-	-	-	-	-	1	-	-	-	-	-	-	-	1	1
d13CDIC	1	3	-	-	5	-	2	-	-	4	-	-	6	11	11
d18OH2O	-	1	-	-	-	-	-	-	-	-	-	-	-	38	38

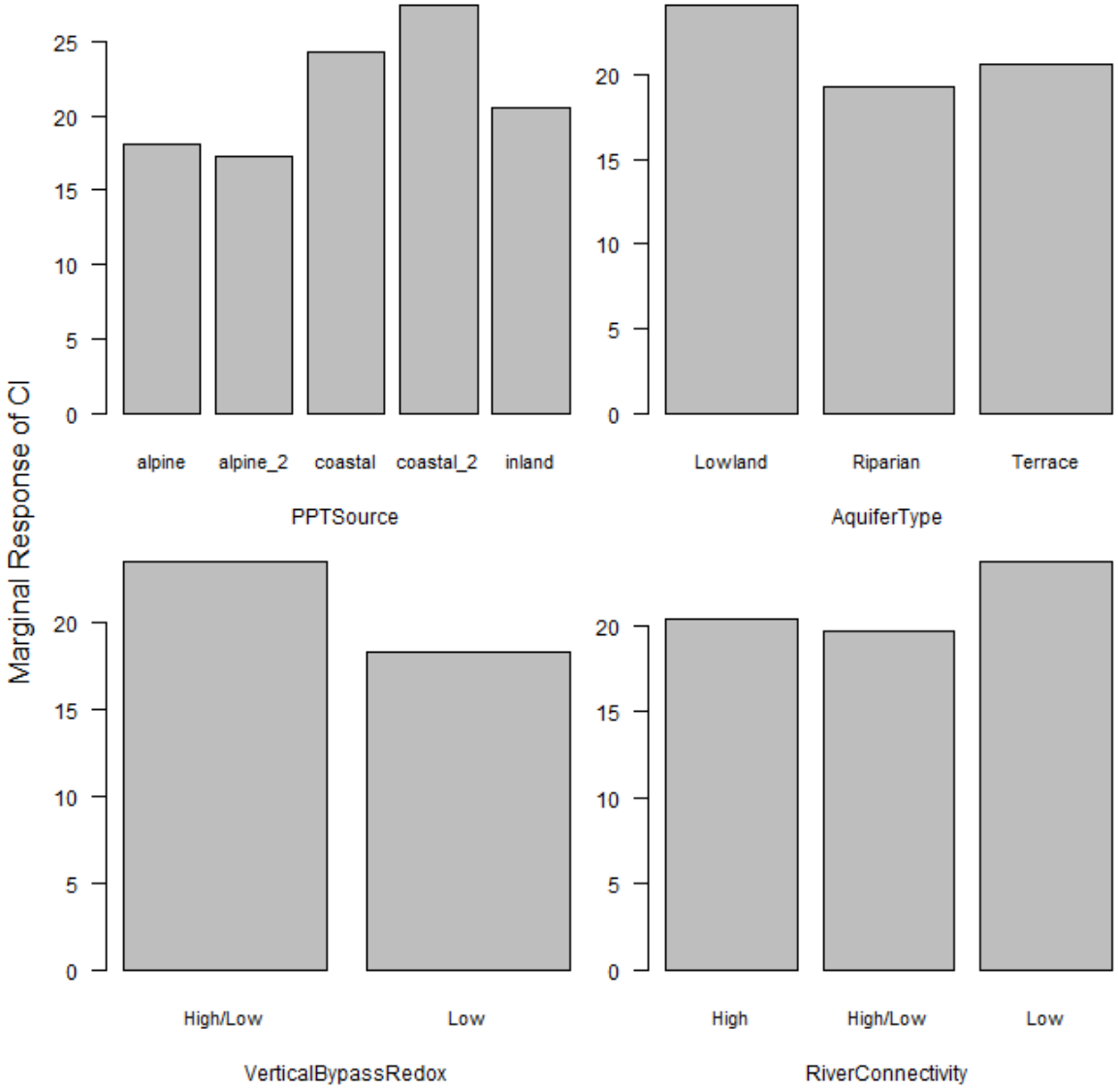


Figure A 14. Partial Dependence of one of the 26 individual groundwater hydro-chemical variables (Cl) on the four predictors that were included in the reduced model (note that these predictors are all categorical). The bars show the marginal effects of each variable's category on the absolute value of the response.

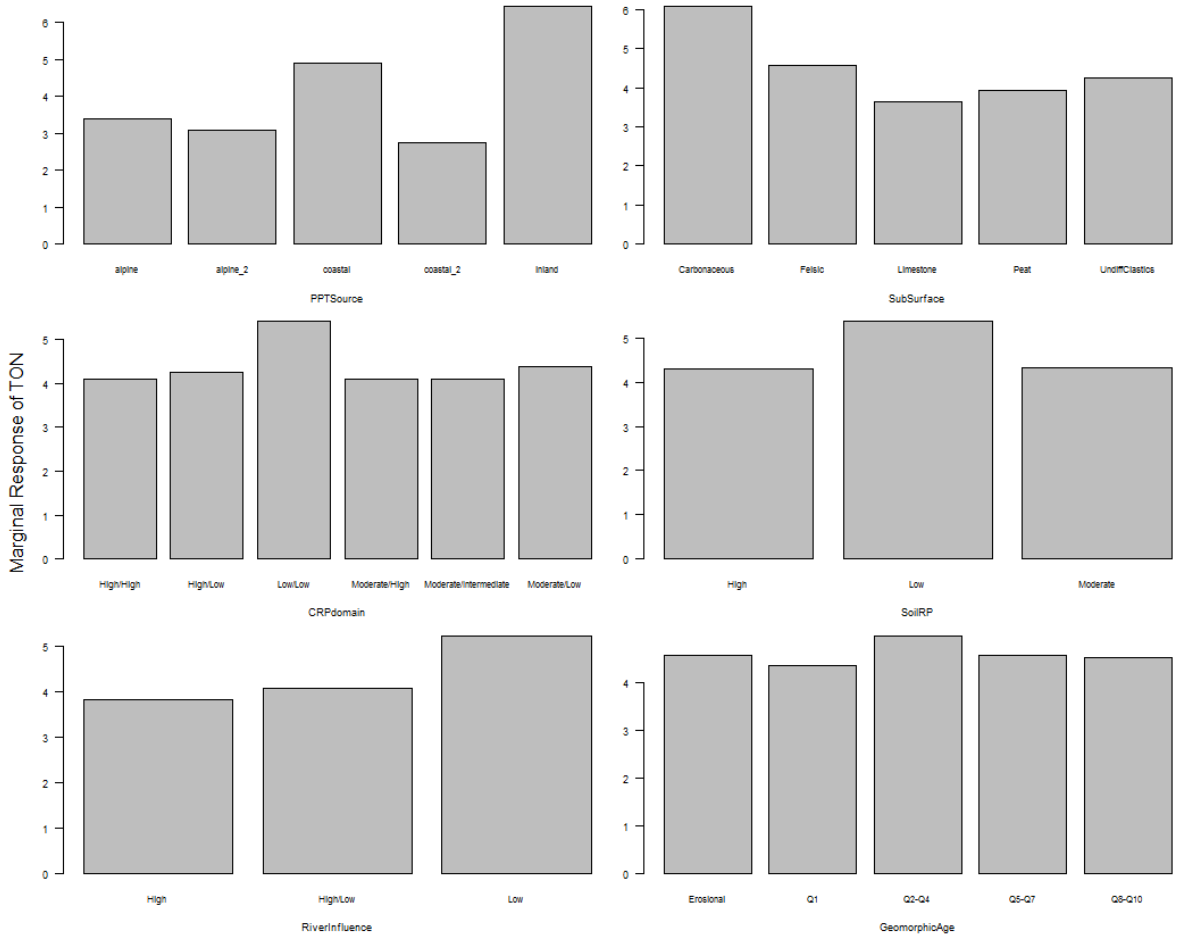


Figure A 15. Partial Dependence of one of the 31 individual groundwater hydro-chemical variables (TON) on the six most important predictors that were included in the reduced model (note that these predictors are all categorical). The bars show the marginal effects of each variable's category on the absolute value of the response.

F1.3 Surface water results

Categorical responses

The categorical surface water responses comprised classes with an average of between 5 and 46 sites per category, depending on the number of categories represented by the response (**Error! Reference source not found.**A 12**Error! Reference source not found.**).

Table A 12. Details of the number of sites in each category for the 5 surface water categorical responses. The first value in the number of sites per category is the mean the minimum and maximum are shown in the parentheses.

Categorical response	Number of categorical levels	Number of sites per category
HCA_Thresh	8	12 (1 - 24)
General_Redox	2	46 (33 - 60)
Redox_Proc	3	31 (17 - 43)
WaterTypeS	5	19 (1 - 40)
Water_Type	17	5 (1 - 19)

The sites varied appreciably between clusters in terms of their hydro-chemical characteristics (Table A 13). For example, relative to the other classes, class 1A0 of the HCA classification had low conductivity, high chloride and sodium, low dissolved oxygen and total organic nitrogen and high values of d18OH2O and d2HH2O (TableA 13**Error! Reference source not found.**). Class 2B0 was characterised by high conductivity, high chloride and sodium, very high TON and medium values of d18OH2O and d2HH2O (TableA 13**Error! Reference source not found.**).

Table A 13. Characteristic values of selected hydro-chemical variables in each class defined by the eight class level of the surface water HCA classification. The values are the mean of the median values of the sampling sites in each class.

Class	Cond	Cl	-	DOField	TON	DOC	d18OH2O	d2HH2O
1A0	166	39	22	8	0.01	29	-6	-38
1B0	54	2	3	11	0.21	2	-10	-66
1C2	100	11	9	11	0.65	3	-8	-55
2A1	303	25	18	12	2.97	4	-7	-49
2A2	187	22	16	11	1.66	5	-7	-48
2B0	285	27	18	9	5.85	4	-7	-49
2C0	208	33	19	11	1.46	11	-6	-42
3B2	215	23	15	10	1.42	8	-7	-47

The reduced model of HCA classes had a fitted miss-classification rate of 2% and a OOB miss-classification rate of 27% (Table A 13). These results indicate a strongly consistent relationship between drivers and HCA classes and reasonable ability to predict HCA class (i.e. general hydro-chemical composition) as a function of the drivers.

The reduced model retained nine variables whose importance is shown in TableA 13**Error! Reference source not found.**. This indicates that membership of HCA classes is significantly associated with all nine predictors. Partial dependence of each of the eight HCA classes on the 8 most important predictors are shown in **Figure A 16**A 16 and their hydro-chemical characteristics are shown in TableA 13**Error! Reference source not found.**. The plots show the effect of the 8 most important predictors (X-axis) on the probability a site belongs to Class 1 (Y-axis). The values on the Y-axis of these plots are logit transformed probabilities. They should be interpreted as the margi-l effect of the predictor on the response (which here is the probability of belonging to a particular class).

The plot indicates (for example) that the probability of a site belonging to HCA class 1A0 increases with increasing values of *Land_Surface_Recharge* and generally decreases with increasing values of the other 7 predictors (Figure **A 16**A 16). Similarly, the probability of a site belonging to HCA class 2B0 increases with increasing values of *Land_Surface_Recharge, HVLL,* and *Q1_Q4_Mafic* but decreases with increasing values of the other five predictors (**Figure A 16**A 16).

The models of the four other categorical (General_Redox, Redox_Proc, WaterTypeS and Water_Type) all performed reasonably well with the Water_Type model having the poorest cross validated performance (Table4). These models retained between two and nine predictors (Table A 13).

Partial plots for all classes of each of the five surface water categorical responses are supplied separately with the following naming convention "SWPartialPlotsCatResponse_*Classification_name*.pdf" (e.g., " SWPartialPlotsCatResponse_HCA_Thresh.pdf").

Table A 14. The order of importance of the predictors used in the RF classification models of the five categorical response variables and model performance. The values associated with the predictors indicate the rank order of importance. Missing values (-) indicate the predictor was not included. Model performance is the out-of-bag (OOB) misclassification rate and the misclassification rate of the fitted model and the OOB and fitted Kappa values.

Response	Predictors																										Model Performance			
	Alpine_River_Recharge	Land_Surface_Recharge	HVLL	Mixed_Alpine_Land_Surface_Recharge	Q1_Q4_Mafic	Q6_eQFelsic	MVLL	LVLL	Hill_Bedrock_River_Recharge	RIVER_SOURCE	HVHH	alpine	Clastics	Peat	coastal	coastal_2	Erosio-l_Felsic	inland	Bypass_Flow	Carbo-te	Q4_Q6Felsic	Q6_eQMafic_Ultramafic	Q1_Q4_Feslic	alpine_2	Erosio-l_Mafic_Ultramafic	Lignite	Q4_Q6Mafic_Ultramafic	OOB Misclassification	Fit Misclassification	OOB Kappa	Fit Kappa
HCA_Thresh	4	1	2	-	3	-	-	-	5	9	-	-	7	-	-	-	8	-	-	-	-	6	-	-	-	-	-	27	2	0.7	1.0
General_Redox	1	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	9	9	0.8	0.8
Redox_Proc	3	-	-	2	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	26	20	0.6	0.7
WaterTypeS	3	-	-	2	1	-	-	-	-	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	20	14	0.7	0.8
Water_Type	7	-	1	9	4	-	3	2	-	-	-	8	-	-	6	-	-	5	-	-	-	-	-	-	-	-	-	45	1	0.5	1.0

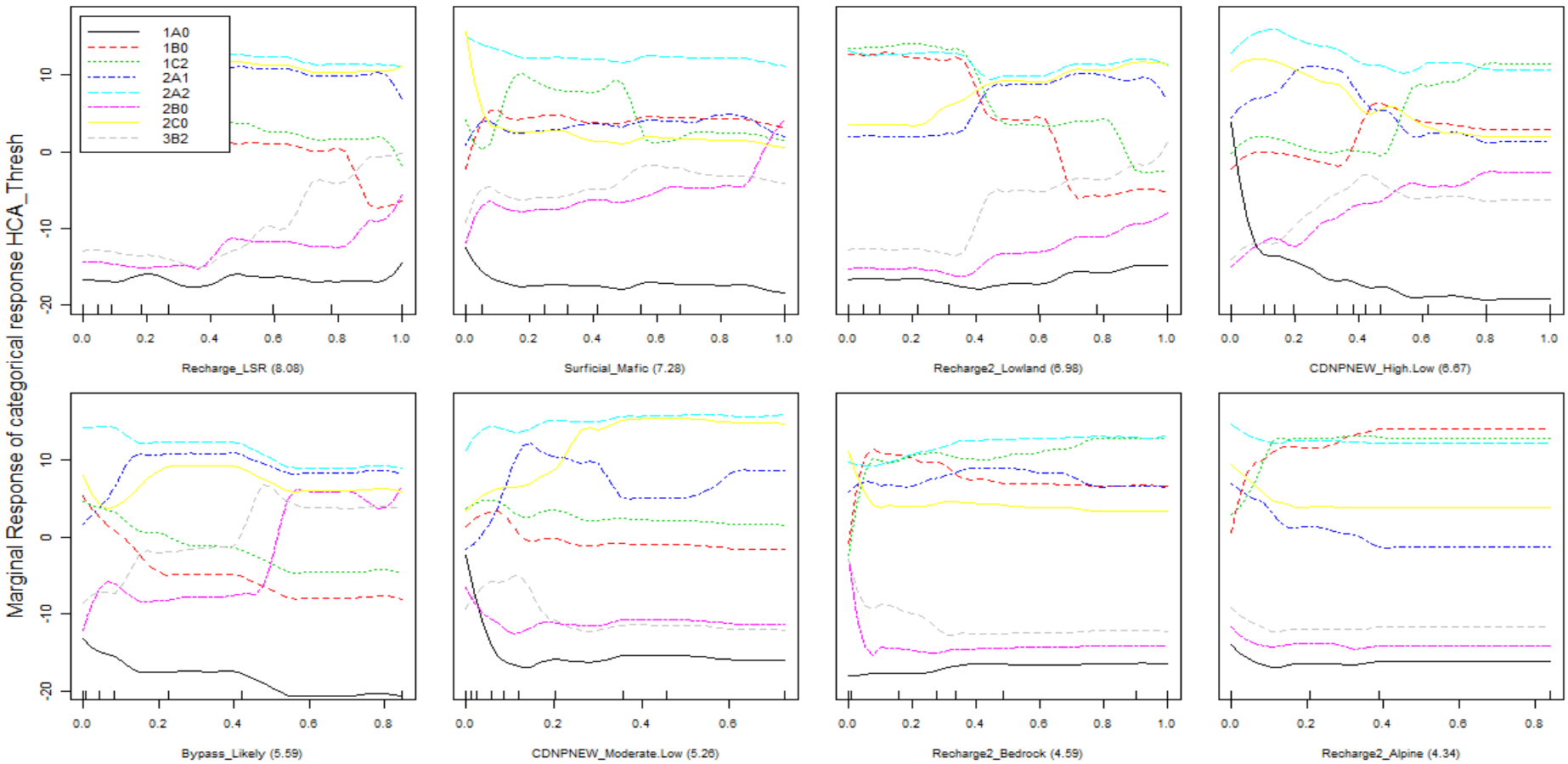


Figure A 16. Partial Dependence of each of the eight surface water HCA classes on the 8 most important predictors. The lines show the probability of class membership for given values of each predictor (i.e. the change in probability along the predictor gradient when all other predictors in the model are held at their respective mean values). The ticks on the x-axis represent centiles in increments of ten of the predictor data. The lines have been smoothed to reduce noise to aid interpretation.

Individual surface water hydro-chemical variables

Thirty continuous hydro-chemical response variables were modelled (Table A 15). Model performance varied across the modelled responses from reaso-ble to very good. The number of predictors included in the reduced models and the order of importance of the individual predictors varied for each variable. Partial plots for all models are provided as supplementary files.

Partial plots for the selected responses are shown in Figure A 17. Partial plots for all of the continuous individual hydro-chemical variables are supplied separately.

The plot indicates (for example) that the value of d13CDIC decreases with increasing values of *Alpine_River_Recharge* and *HVLL* (Figure A 17)*,* which are the only two predictors that were included in the reduced model for this variable (Table A 15). Similarly, the Conductivity decreases with increasing with increasing values of *Alpine_River_Recharge* and *Mixed_ Alpine_River_Recharge* and increases with increasing values of *HVLL* and *Q1_Q4_Mafic* (Figure A 17).

Table A 15. Performance of the RF models for the individual surface water hydro-chemical variables. Model performance is indicated by cross validated and fitted model r^2 values. The values associated with the predictors indicate the rank order of importance. Missing values (-) indicate the predictor was not included. Model performance is the cross validated misclassification rate and the misclassification rate of the fitted model.

Response	Predictors																										Model Performance	
	Alpine_River_Recharge	Land_Surface_Recharge	HVLL	Mixed_Alpine_Land_Surface_Recharge	Q1_Q4_Mafic	Q6_eQFelsic	MVLL	LVLL	Hill_Bedrock_River_Recharge	RIVER_SOURCE	HVHH	alpine	Clastics	Peat	coastal	coastal_2	Erosio-l_Felsic	inland	Bypass_Flow	Carbo-te	Q4_Q6Felsic	Q6_eQMafic_Ultramafic	Q1_Q4_Feslic	alpine_2	Erosio-l_Mafic_Ultramafic	Lignite	Q4_Q6Mafic_Ultramafic	Cross validated r^2	Fitted r^2
d13CDIC	-	1	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	68	68
d18OH2O	1	6	11	3	10	16	15	18	8	4	14	2	13	-	17	12	9	7	19	-	-	20	-	5	-	-	92	93
d2HH2O	1	6	12	2	9	17	14	15	11	4	19	3	13	20	16	7	10	8	18	-	-	-	-	5	-	-	93	93
Cl	4	2	-	6	10	-	13	11	3	1	12	5	-	-	9	7	-	8	14	-	-	-	-	-	-	-	94	94
Br	2	3	10	6	8	1	11	12	5	4	18	7	15	-	20	9	-	13	16	-	-	-	17	19	14	-	86	86
SO4	4	1	2	-	-	-	5	-	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	81	82
-	1	5	-	2	14	-	10	11	6	4	13	3	-	-	12	-	7	8	-	-	-	-	-	-	9	-	91	91
Cond	3	1	2	6	5	-	10	-	9	8	-	7	11	-	12	14	4	-	-	-	-	-	-	13	-	-	79	80
Ca	14	4	3	19	1	7	5	8	15	16	13	17	11	12	6	-	9	18	10	2	-	-	-	20	-	-	58	59
Mg	7	4	2	-	1	9	-	-	10	8	-	-	-	-	6	-	5	-	3	-	-	-	-	-	-	-	86	86
K	11	2	3	12	9	1	4	6	5	7	-	16	14	8	18	-	13	19	17	-	10	15	-	20	-	-	67	67
SiO2	6	-	3	-	1	5	-	8	7	4	-	-	-	-	2	-	-	9	-	-	-	10	-	-	-	-	74	75
AlkTot	14	-	3	-	1	8	4	7	12	11	-	-	-	9	5	-	10	13	6	2	-	-	-	-	-	-	54	55
AlkHCO3	14	8	3	-	1	6	-	9	12	10	7	-	-	11	4	-	-	13	5	-	-	2	-	-	-	-	51	51
DOC	5	-	-	-	-	-	-	4	-	-	2	-	-	1	-	3	-	-	-	-	-	-	-	-	-	-	81	84
pHField	-	9	-	20	7	4	6	12	8	18	1	16	13	5	-	3	10	17	19	14	2	11	-	-	15	-	48	48
DOField	-	-	-	-	-	5	-	4	-	-	2	-	-	1	-	-	3	-	-	-	-	-	-	-	-	-	-	38	38
TON	-	-	1	-	3	5	-	2	-	-	-	-	-	-	-	-	-	-	4	-	-	-	-	-	-	-	76	77
Mn	9	1	-	6	11	7	3	8	-	-	2	10	12	-	14	5	4	-	13	-	-	-	-	-	-	-	71	71
Fe	6	-	-	5	-	-	-	-	-	-	2	-	-	1	-	4	-	-	7	-	3	-	-	-	-	-	75	76
NH4	9	2	-	7	-	1	-	-	4	8	-	10	6	-	-	3	5	-	-	-	-	-	-	-	-	-	46	46
SS	7	9	4	3	-	-	1	10	-	15	5	8	2	12	16	13	11	18	-	-	-	-	6	19	14	17	20	33	33
DRP	-	5	-	8	6	-	4	-	-	-	1	9	7	2	-	-	-	-	-	-	3	-	-	10	-	-	-	32	33
SSV	13	5	16	14	18	-	4	7	17	15	3	6	8	2	11	1	9	19	-	-	12	-	10	20	-	-	-	62	62
Clarity	3	5	-	1	-	8	-	-	10	7	12	4	2	-	-	13	9	11	14	-	-	-	-	6	-	-	-	72	73
Ecoli	-	5	3	7	-	2	4	1	-	-	-	-	-	-	6	-	-	-	-	-	-	-	-	-	-	-	-	38	38

ORP	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-	37	38
I	6	2	4	8	12	1	-	5	3	9	10	14	11	-	-	7	-	-	13	-	-	-	-	-	-	-	-	76	76
F	8	-	-	-	-	6	4	7	-	9	5	-	3	2	-	-	-	1	-	-	10	-	-	-	-	-	-	64	64
Turb	11	10	3	5	13	2	1	-	-	14	6	9	4	-	-	-	-	-	12	-	8	-	7	-	-	-	-	50	50
Temp	-	-	-	-	-	2	-	1	4	-	-	-	-	-	5	-	-	-	-	-	-	3	-	-	-	-	-	36	36

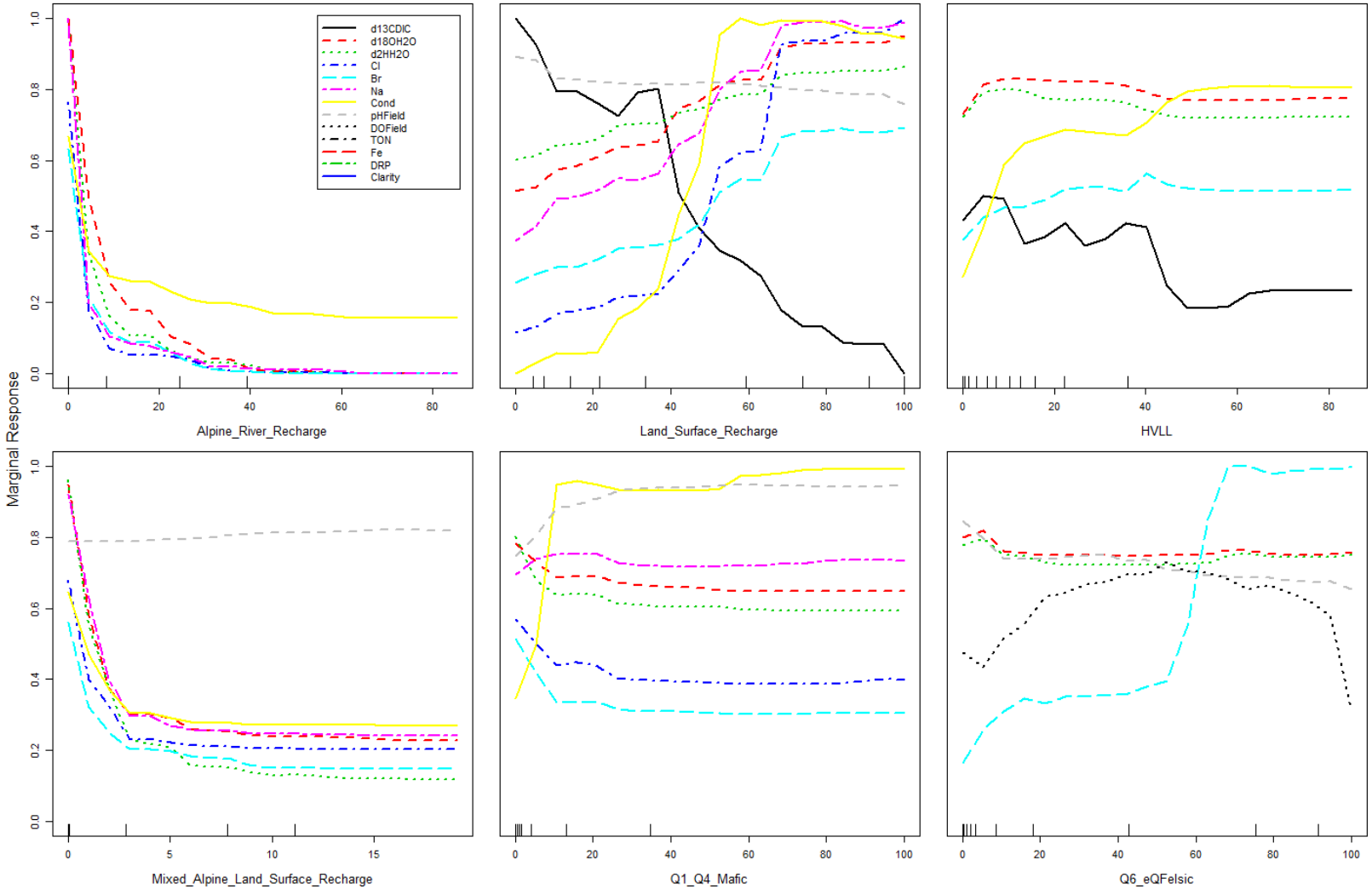


Figure A 17. Partial plot for reduced model for selected individual surface water hydro-chemical variables. The Y-axis represents the margi-l effects of each predictor on the absolute value of the response (i.e. the change in probability along the predictor gradient when all other predictors in the model are held at their respective mean values). The ticks on the x-axis represent centiles in increments of ten of the predictor data. The lines have been smoothed to reduce noise to aid interpretation

F1.4 Summary

In this analysis we tested the relationships between observed hydro-chemistry and drivers postulated by the conceptual model. Despite the relatively small test datasets, all the models achieved statistical significance, however, the models performance was variable (i.e. models explained differing amounts of variation). In general, model performance was better for the surface water than the groundwater data.

Model fitting used variable reduction to reduce the predictors selected by the models to the most parsimonious set. While some models selected only a subset of predictors, the reduced models generally included a number of predictors. In particular, various models predicted classes that represented groups of sites with similar hydro-chemical characteristics. These models always included many predictors, which supports an underlying hypothesis postulated by the conceptual model that aspects of hydro-chemical character are associated with a complex mix of drivers (i.e. driver assemblages).

We used partial plots to reveal the associations between surface and groundwater hydro-chemical characteristics and the model predictors. The model predictors represented the drivers of hydro-chemistry postulated by the conceptual model. We found that the partial plots described associations between hydro-chemistry and drivers that are consistent with the conceptual model.

F2 Stratification

To support findings of stratification presented in Chapter 4 and our conceptual understanding of the chemical evolution of Southland's freshwater, we illustrated the assemblage of dominant drivers diagrammatically within a GIS mapping platform, presented below. The latter is particularly handy for streams as it spatially depicts the relationship between changes in dominant drivers and changes in hydrochemistry from source to confluence or source to river mouth.

Alpine derived rivers and streams

Within this cohort surface waters characterised as 1B0b by Daughney et al. (2015) are a mix of ARR and BRR. These streams and rivers originate within alpine areas, gaining significant BRR below the tree line (ca. 800 m RSL) prior to debouching onto the lowland plains (example in Figures A 18, A19 and A 20). Unlike purely BRR derived surface waters these waters are not characterised by reducing signatures or Na-HCO_3/Cl facies rather they retain a dominant ARR signature from source to mouth. However, an assessment of median hydrochemistry between 1B0a and 1B0b waters reveals a 3.5, 1.2 and 4.5 times increase in dissolved Mn(II), Fe(II) and TOC for 1B0b waters supporting a significant contribution from areas of reducing BRR.

Figure A 18 illustrates the changes in hydrochemistry of the Mataura River from source to mouth. Specifically, redox state changes from oxic to mixed (oxic-anoxic) due to the increasing proportion/influence of reducing soil/geology on the concentration of redox sensitive. Major ions composition changes from Ca dominated to Na dominated with decreasing proximity to the coast due to increasing marine aerosolic Na (and Cl) load in precipitation.

Further, HCA cluster membership of the Mataura River samples changes from a 1B0b to 1C2a and finally to 1C2b as the solute contribution from BRR and LSR increases. Similarly, HCA cluster membership of the Oreti River changes from 1B0a to 1C2a, and that of the Aparima River changes from 1B0a to 1C2a to finally 2A2c from source to mouth (Figures A 19 and A 20, respectively). The change from 1B0b to 1C2a occurs as the proportion of LSR exceeds 10%. 1C2b waters differ from 1C2a waters in that the latter have a higher proportion of BRR. The change from 1C2a/1C2b to 2A2c

facies reflects negligible ARR and greater BRR and most importantly LSR derived solute inputs associated with young mafic soils and geology.

Bedrock Derived Streams and Rivers

All of the Bedrock Derived Streams and Rivers sites have a source associated within currently or historically forested bedrock outcrop (*senso lato* 'Hill Country'), that occurs below 800 m RSL and have no Alpine River Recharge (ARR) within their capture zone.

Changes in stream chemistry as the proportion of LSR (or degree of influence associated with LSR) within a capture zone increases coincides with a shift in hydrochemical facies from BRR dominated (1C2a/1C2b cluster membership) to BRR derived waters that have a significant LSR component (2A2c facies) and to carbonate influenced waters (2A1o facies) (examples illustrated in Figure A 21 and A 22).

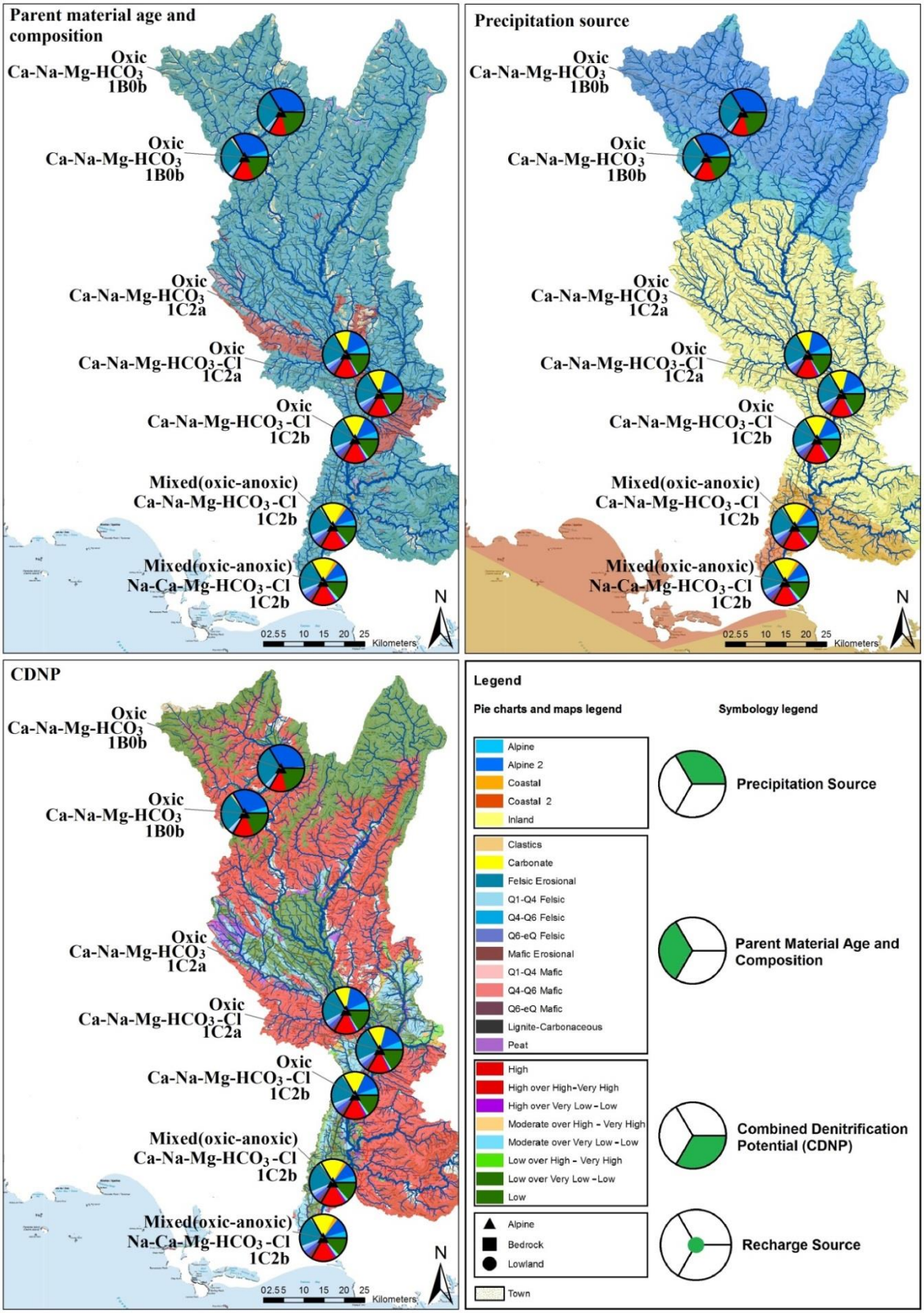


Figure A 18: Diagrammatical representation of the hydrochemical evolution of freshwater within the Mataura River with respect to the three drivers (parent material and age, precipitation source, CRP). Pie charts represent the proportion of area associated with driver categories at each site. Labels detail redox state, long water type and HCA assignment after Daughney et al. (2015).

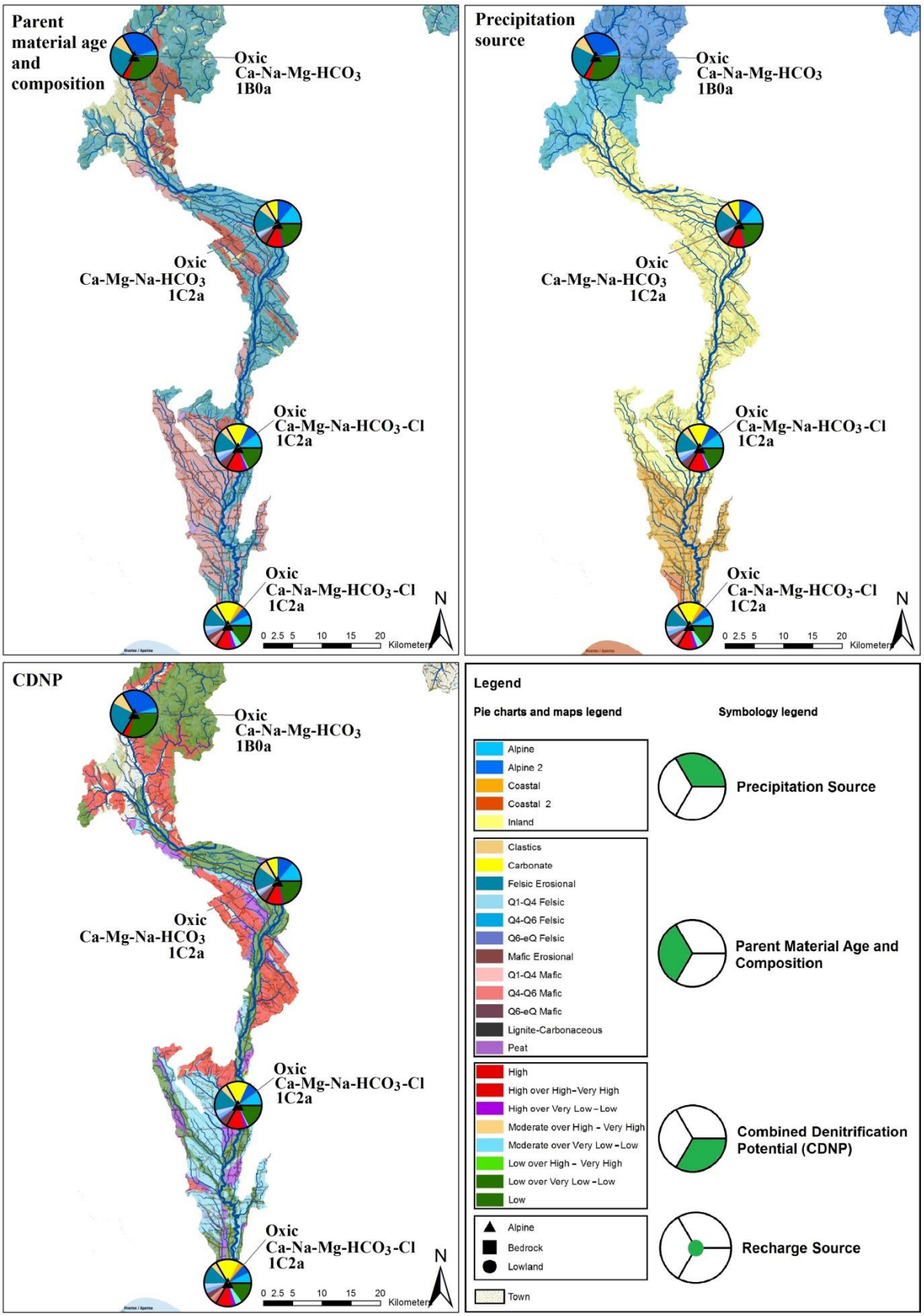


Figure A 19: Diagrammatical representation of the hydrochemical evolution of freshwater within the Oreti River with respect to the three drivers (parent material and age, precipitation source, CRP). Pie charts represent the proportion of area associated with driver categories at each site. Labels detail redox state, long water type and HCA assignment after Daughney et al. (2015).

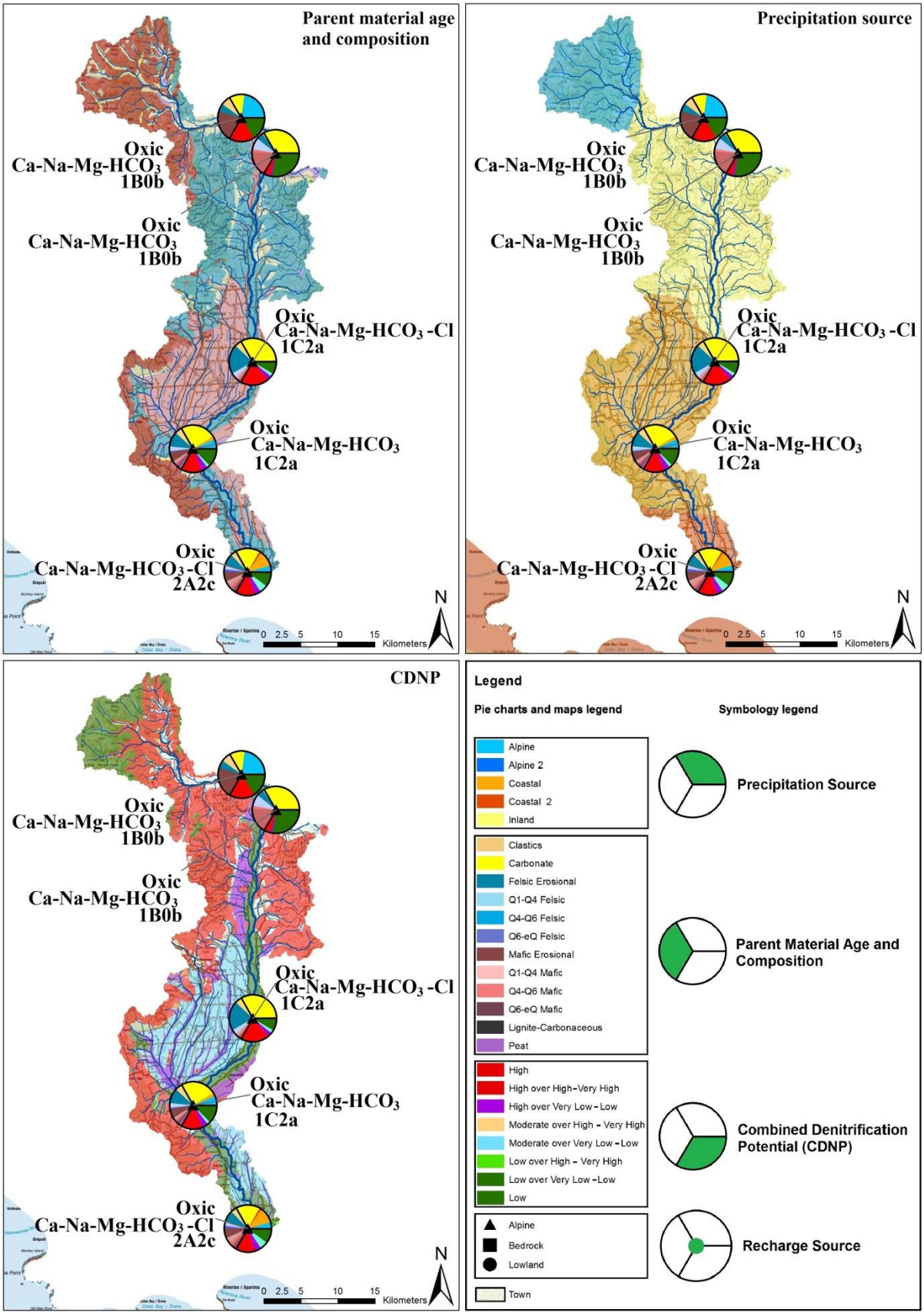


Figure A 20: Diagrammatical representation of the hydrochemical evolution of freshwater within the Aparima River with respect to the three drivers (parent material and age, precipitation source, CRP). Pie charts represent the proportion of area associated with driver categories at each site. Labels detail redox state, long water type and HCA assignment after Daughney et al. (2015).

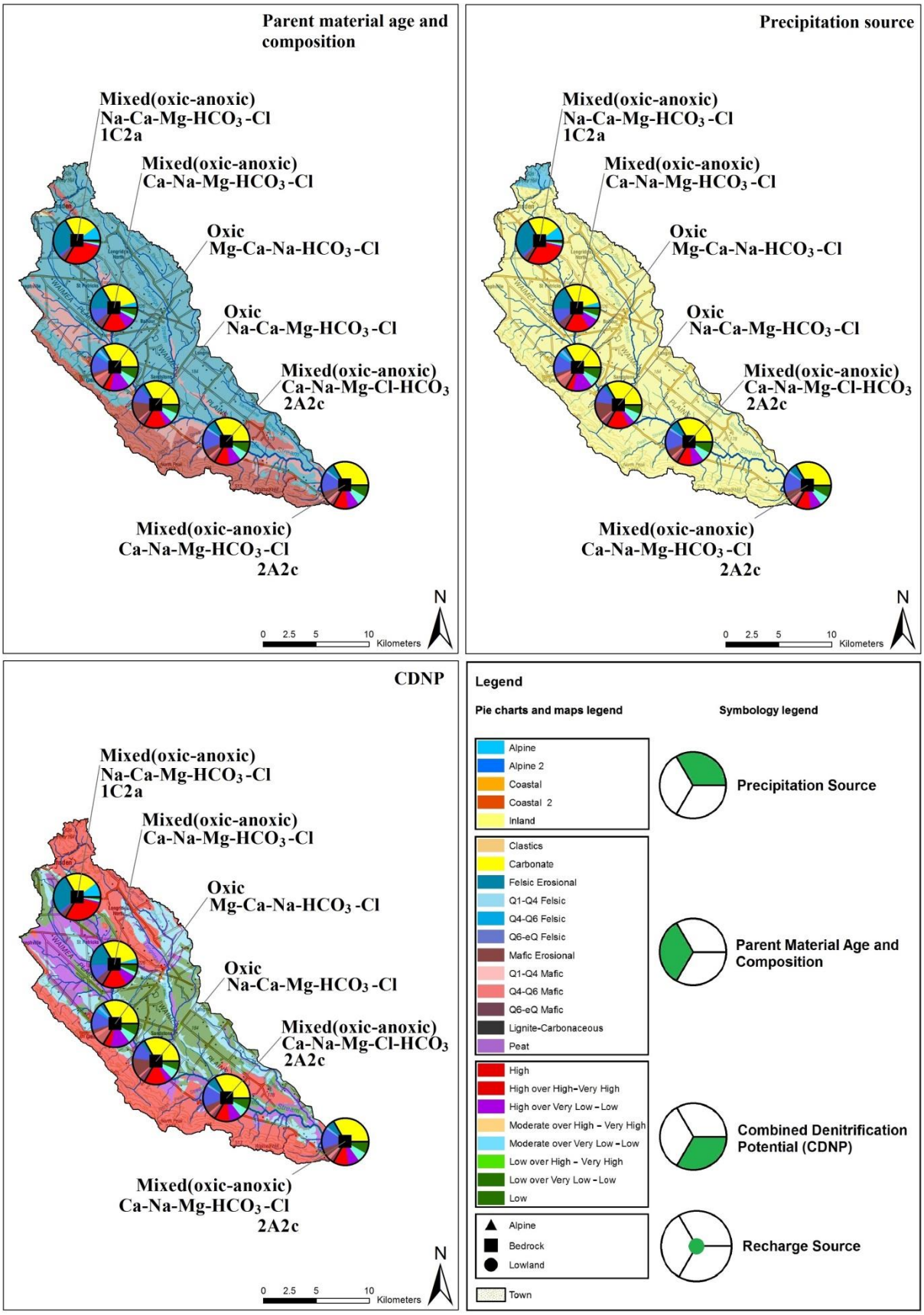


Figure A 21: Diagrammatical representation of the hydrochemical evolution of freshwater within the Waimea Stream with respect to the three drivers (parent material and age, precipitation source, CRP). Pie charts represent the proportion of area associated with driver categories at each site. Labels detail redox state, long water type and HCA assignment after Daughney et al. (2015).

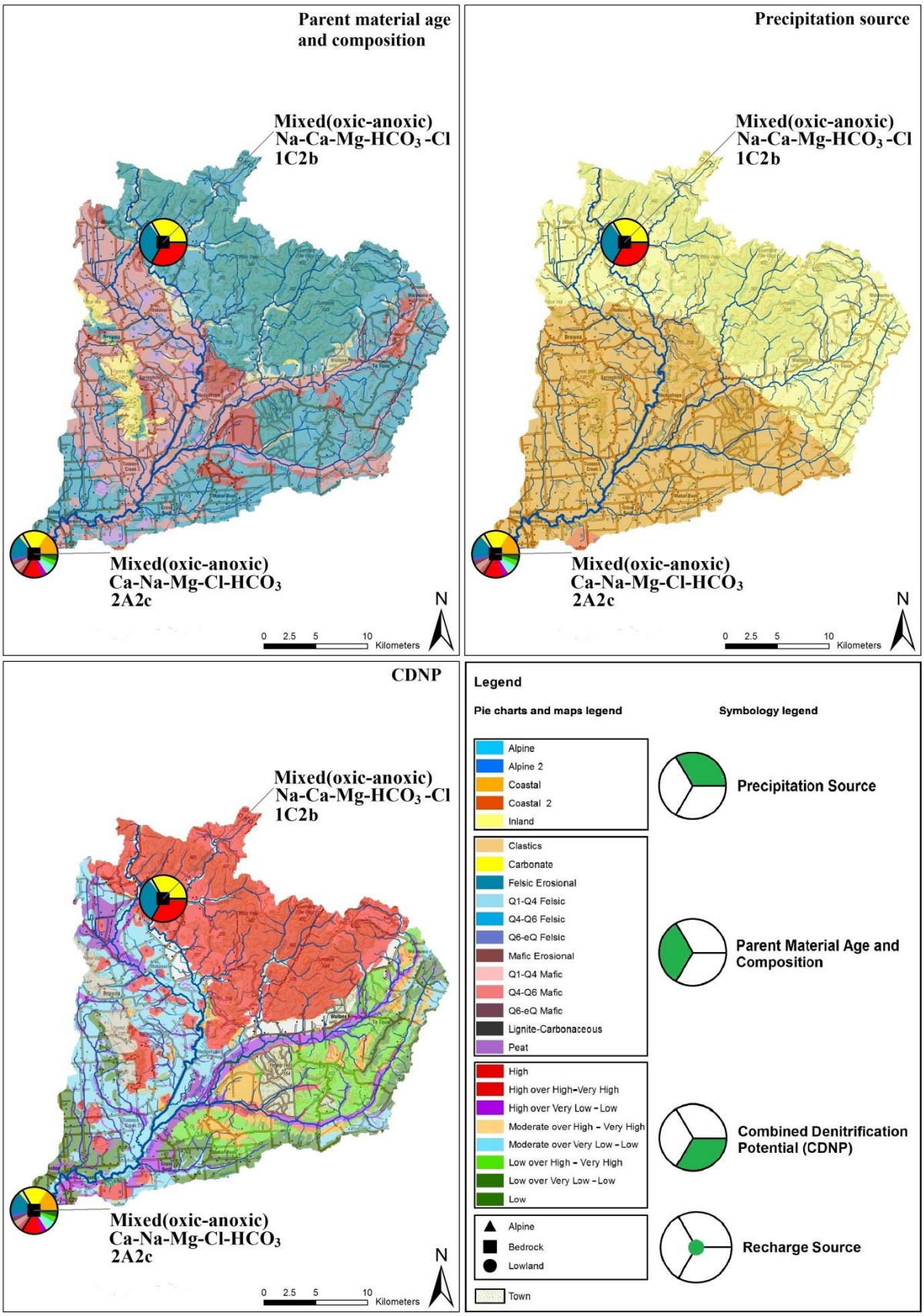


Figure A 22: Diagrammatical representation of the hydrochemical evolution of freshwater within the Makarewa River with respect to the three drivers (parent material and age, precipitation source, CRP). Pie charts represent the proportion of area associated with driver categories at each site. Labels detail redox state, long water type and HCA assignment after Daughney et al. (2015).

Lowland - Land Surface Recharged (LSR), Streams and Rivers

Sites classified as having a lowland-LSR source constitute the largest and most varied category of Southland surface waters (n = 38 sites). These surface waters are generally more mineralised (at their source) than ARR and BRR derived streams. Only one site, Sandstone Stream at Kingston Crossing Rd, also has a minor (1.4% of the catchment area) BRR component within its capture zone. None of the sites have ARR within the capture zone.

Generally, we often see a change in redox state due to changes in proportion/influence of reducing and oxidising soil and geology in the capture zone. Similarly, major ion composition changes as a function of increasing/decreasing influence of marine aerosolic Na and Cl with decreasing/increasing proximity to the coast and/or change in geology (e.g. proportion of felsic and carbonate geology). Figure A 23 depicts the hydrochemical evolution of freshwater within the Waimatuku Stream with respect to the three drivers (parent material and age, precipitation source, CRP). Redox state switches from mixed (oxic, anoxic) to oxic and back to mixed (oxic, anoxic) as the proportion of oxidising to reducing soil and geology switches between high and low. Major ions of the Waimatuku Stream remain Ca and HCO_3, despite a significant proportion of coastal precipitation (from source to mouth) suggesting that the Waimatuku stream is significantly impacted by the presence of carbonate rock within its capture zone. Secondary major cations swap between Na and Mg as the proportion of felsic and mafic geology increases, respectively. HCA cluster membership of the Waimatuku stream at its source is 3B2a (Organic carbon/reducing clusters LSR). Further downstream and at its mouth HCA cluster membership is 2A1o (carbonate influenced waters that exhibit periods of saturation with respect to calcite), again reflecting the significant influence of carbonate on the composition of the Waimatuku stream.

Only sites from with the Central Plains area (Waimatuku, Middle Creek and Ayr Creek) deviate from the general trend. Specifically, despite strongly reducing soils these sites all show oxidising characteristics such as elevated TON and low Mn(II) and Fe(II) concentrations. For all of the sites within the Central Plains area there is strong evidence for macropore cracking during the summer months and the bypass of the reducing soil zone by autumn recharge (TC 6) suggesting reduction of redox sensitive species (including TON, Mn(IV) and Fe(III)) does not occur despite reducing soil zone.

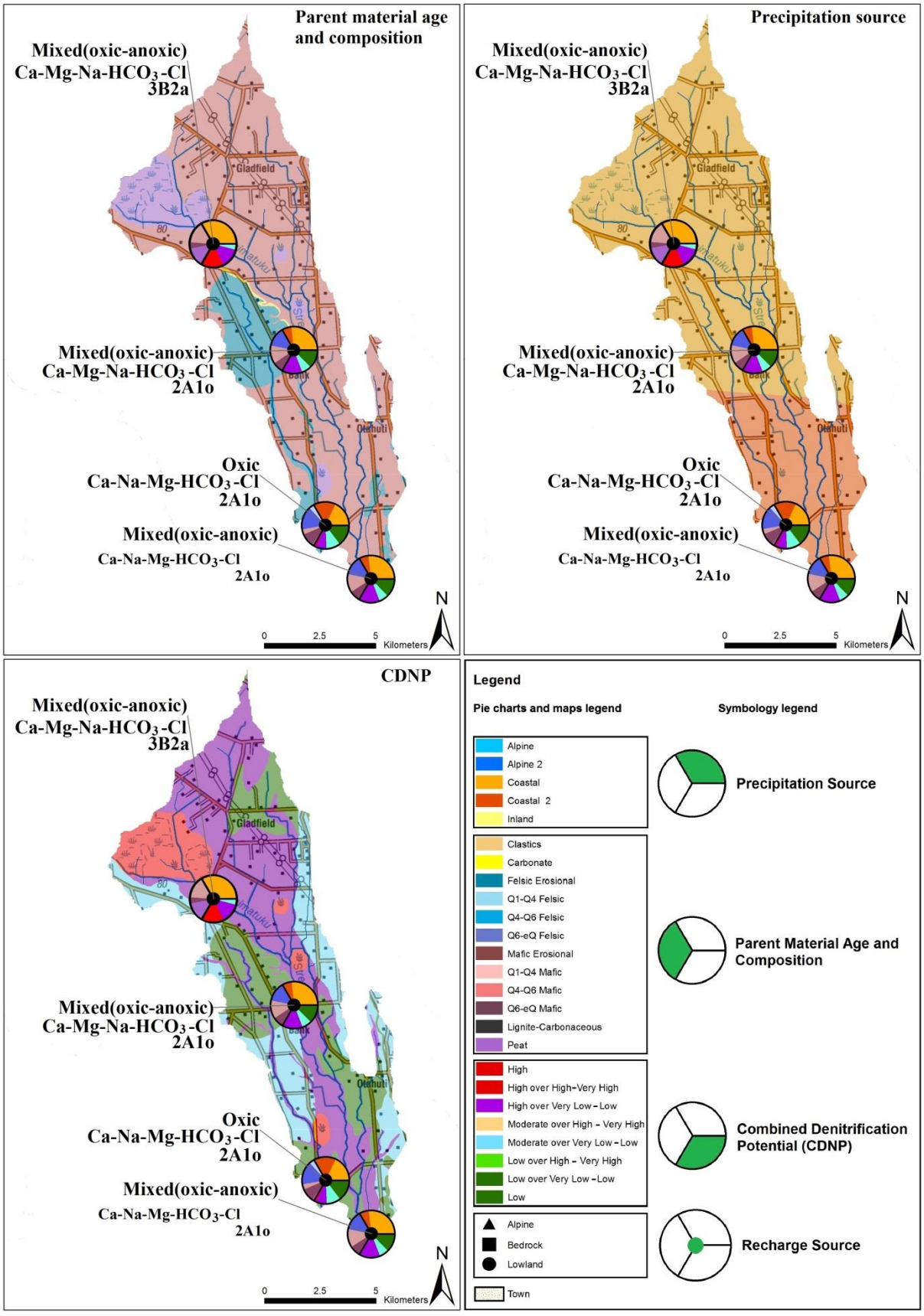
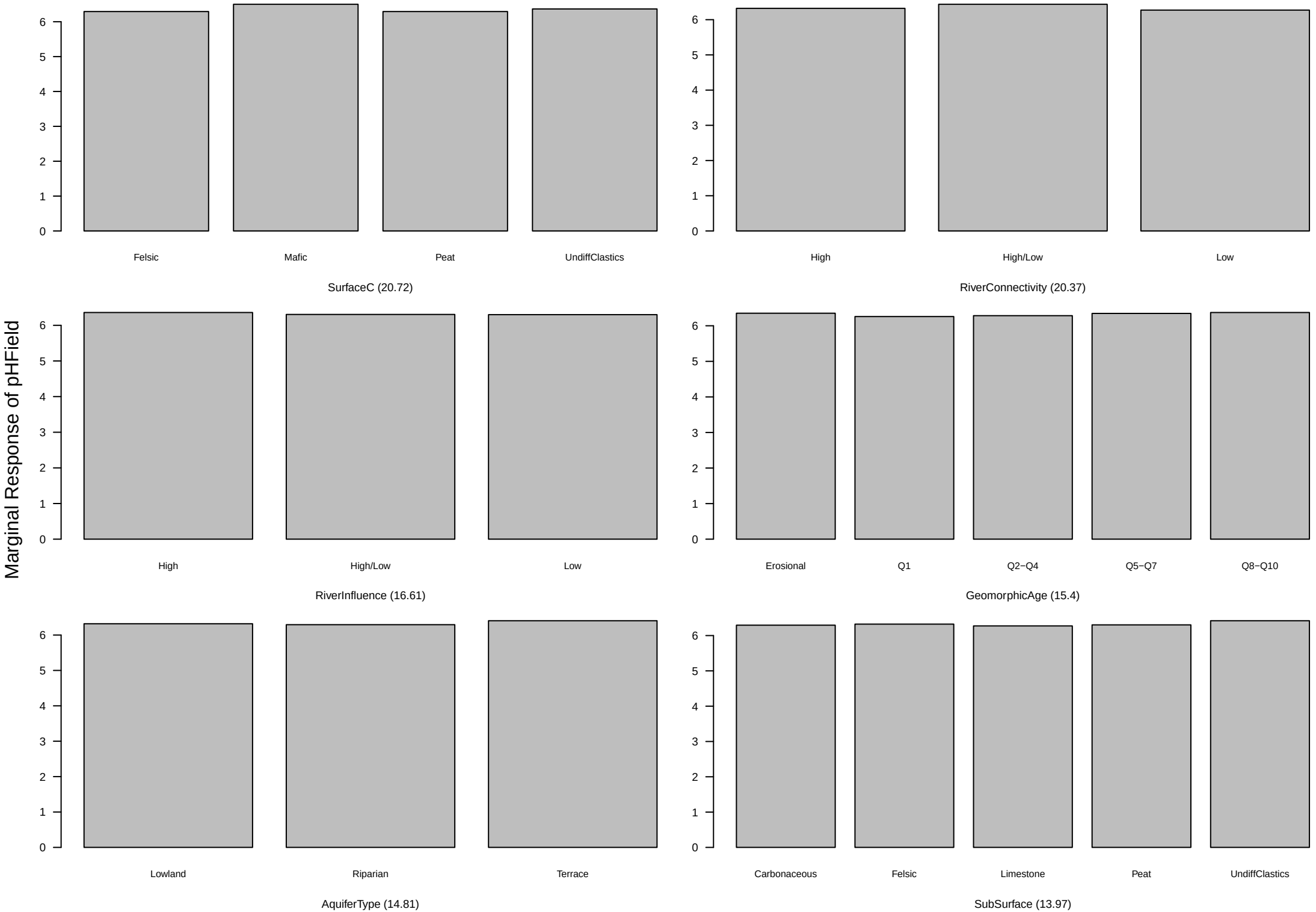


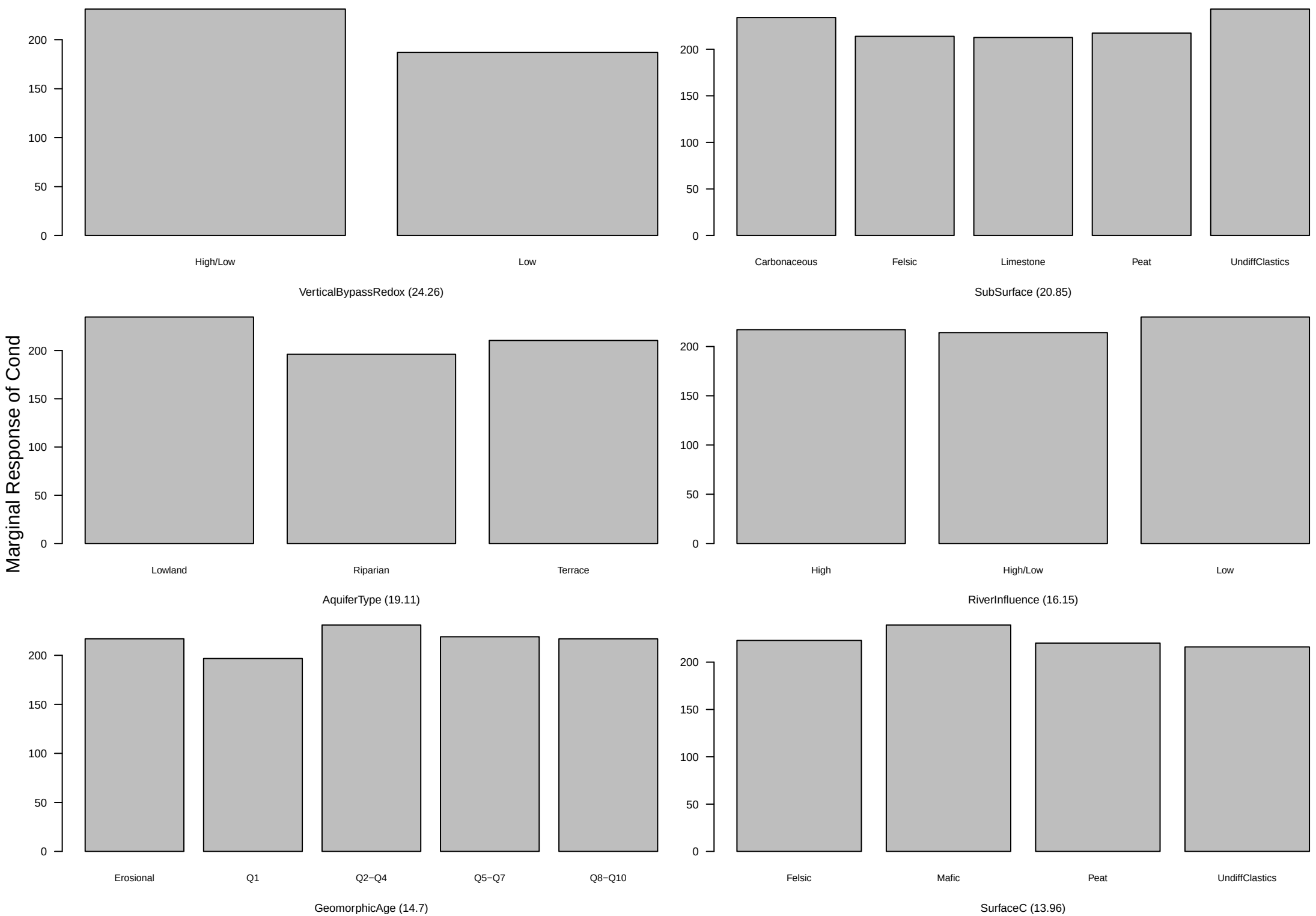
Figure A 23: Diagrammatical representation of the hydrochemical evolution of freshwater within the Waimatuku Stream with respect to the three drivers (parent material and age, precipitation source, CRP). Pie charts represent the proportion of area associated with driver categories at each site. Labels detail redox state, long water type and HCA assignment after Daughney et al. (2015).

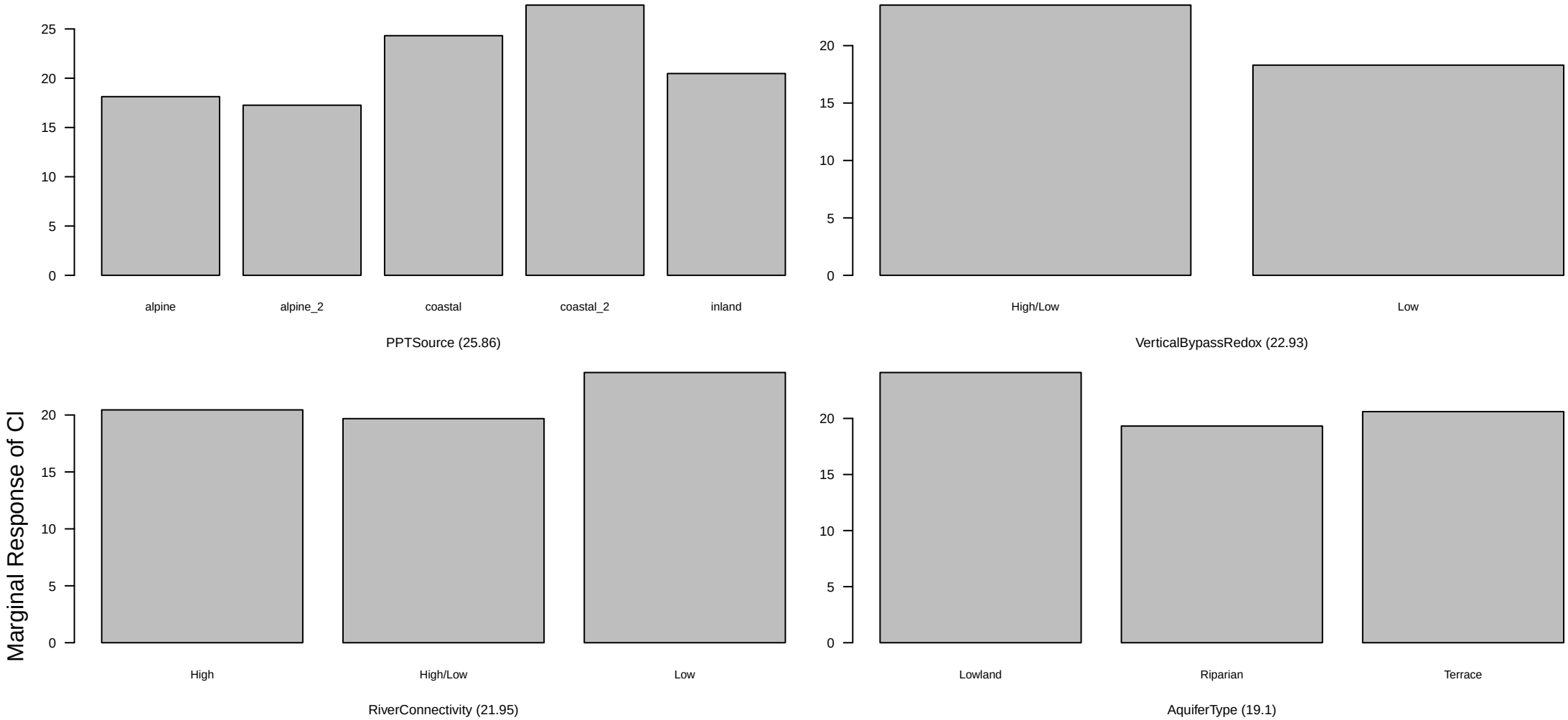
Appendix G – Data

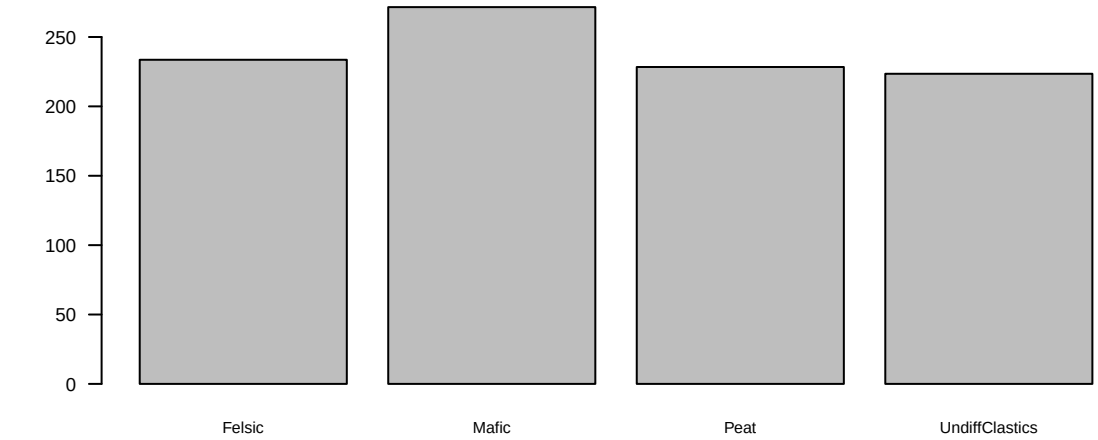
This appendix details data used for validation of the conceptual model, i.e. stratification and empirical modelling.

Links to CSV files compiling data in addition to descriptions of the datasets will be provided at a later stage. In the interim, the reader is welcome to contact the authors for provision of the data.

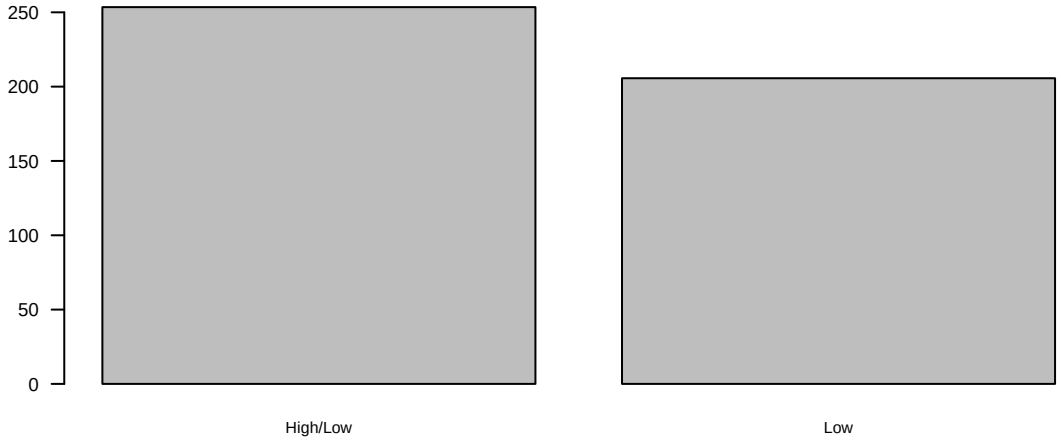

Marginal Response of pHField
0
1
2
3
4
5
6
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (20.72)
High
High/Low
Low
RiverConnectivity (20.37)
High
High/Low
Low
RiverInfluence (16.61)
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge (15.4)
Lowland
Riparian
Terrace
AquiferType (14.81)
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (13.97)


Marginal Response of Cond
0
50
100
150
200
High/Low
Low
VerticalBypassRedox (24.26)
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (20.85)
Lowland
Riparian
Terrace
AquiferType (19.11)
High
High/Low
Low
RiverInfluence (16.15)
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge (14.7)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (13.96)

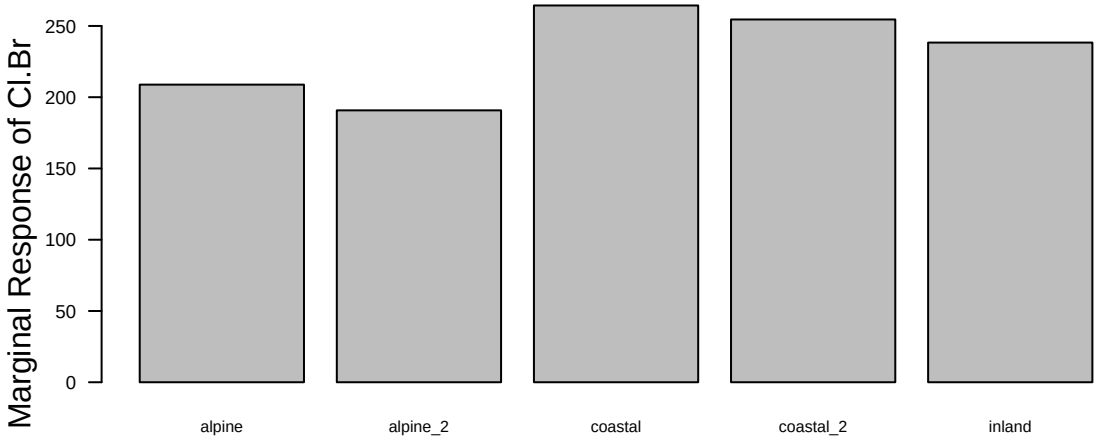

Marginal Response of Cl
0
5
10
15
20
25
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (25.86)
High/Low
Low
VerticalBypassRedox (22.93)
High
High/Low
Low
RiverConnectivity (21.95)
Lowland
Riparian
Terrace
AquiferType (19.1)



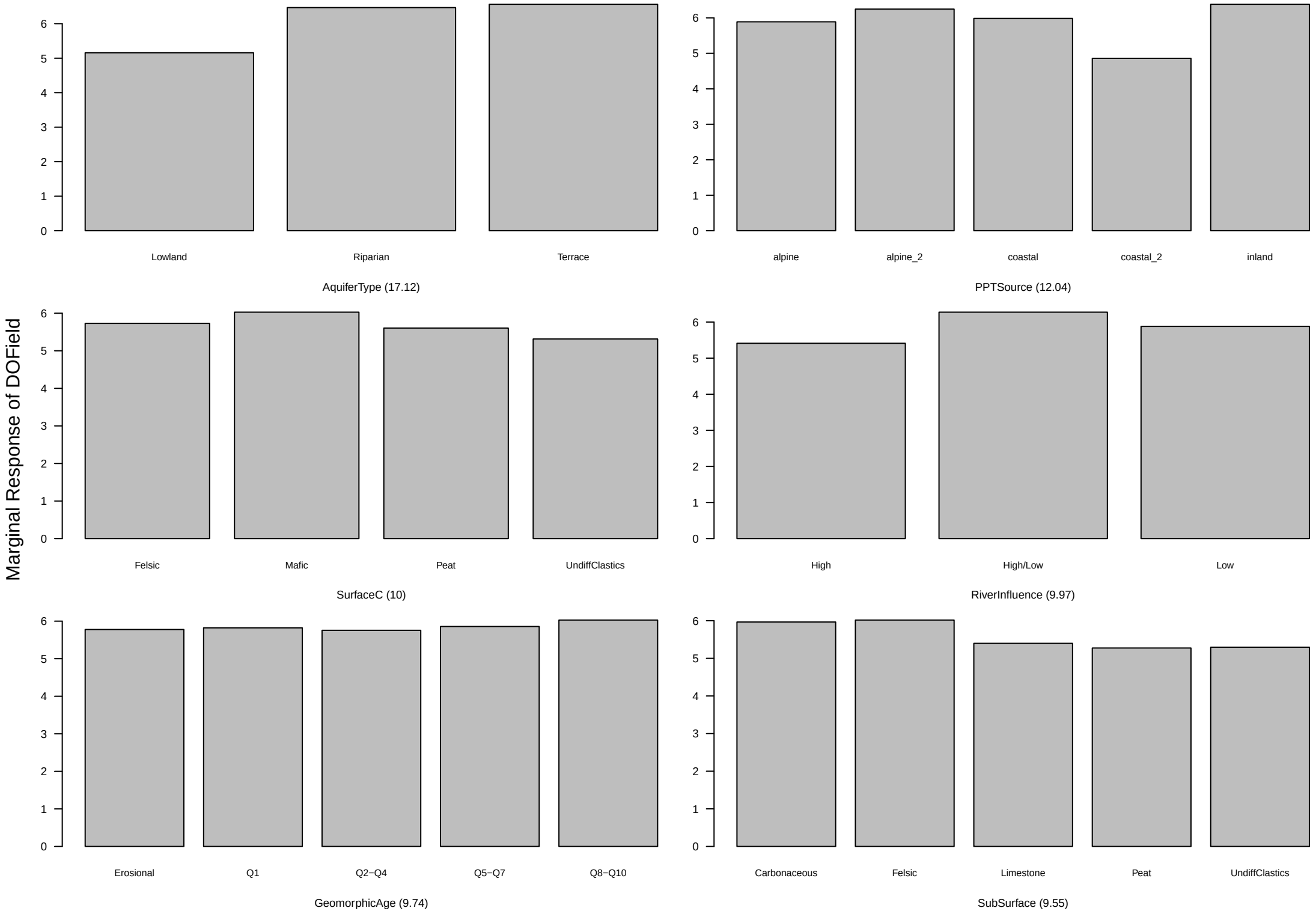
SurfaceC (26.03)

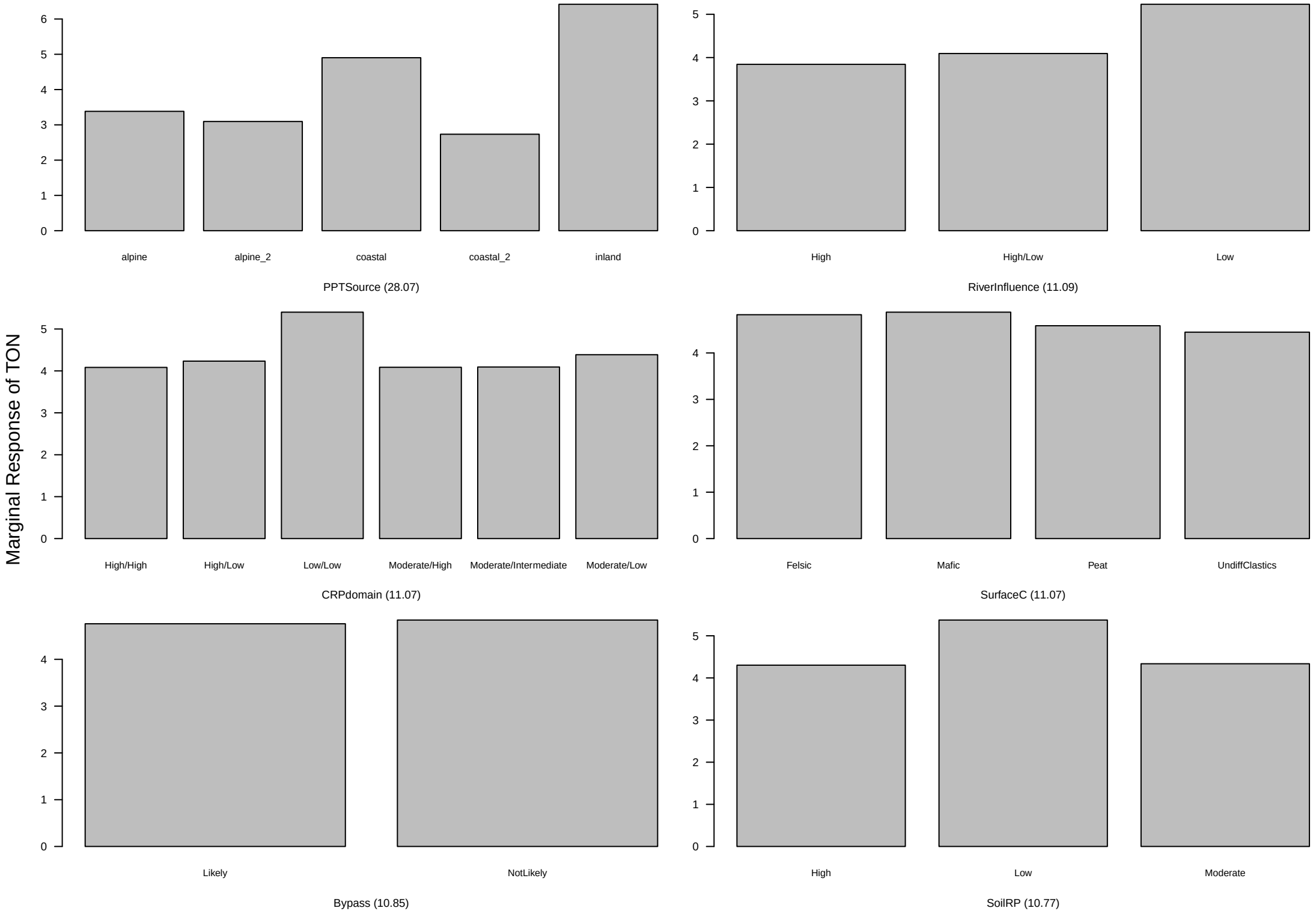


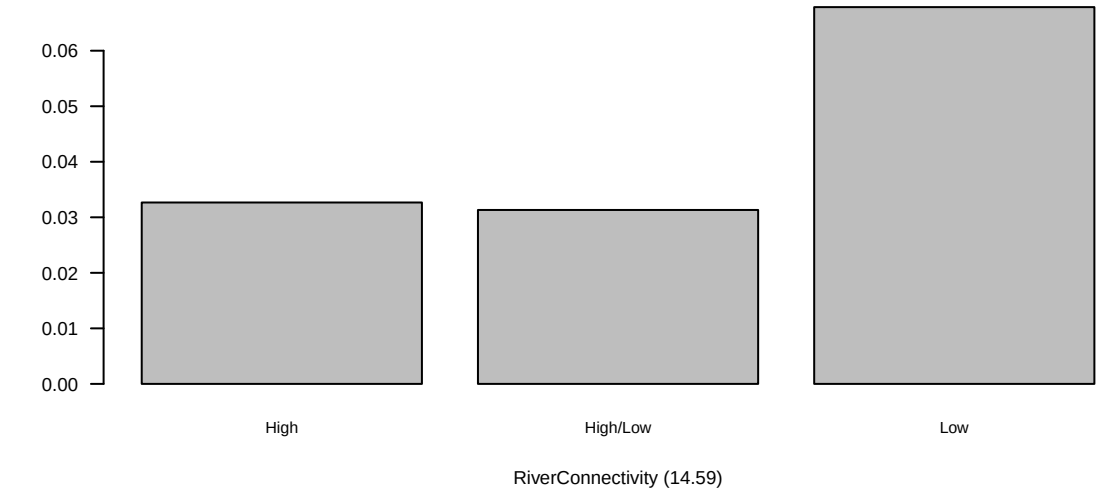
VerticalBypassRedox (24.94)

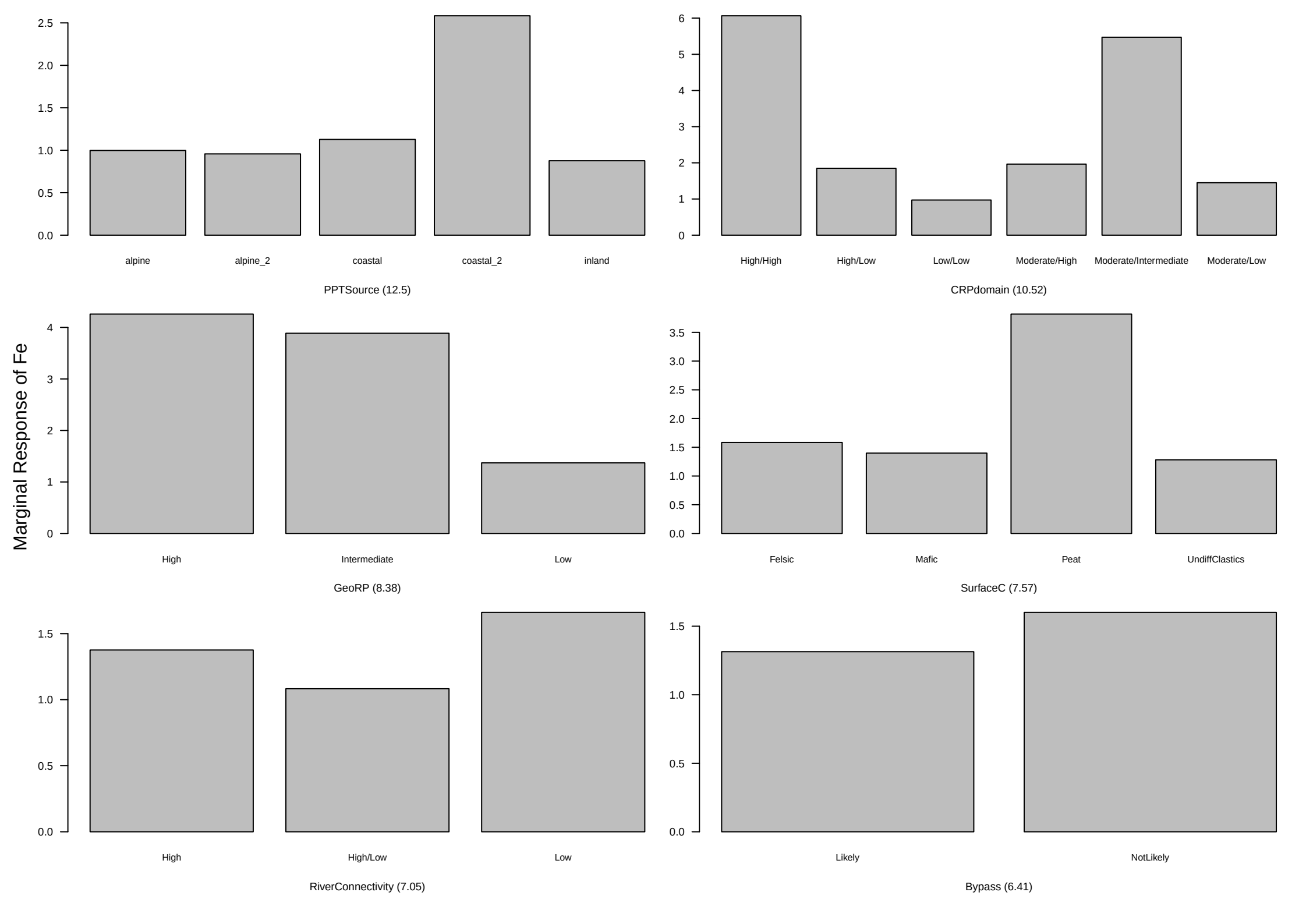


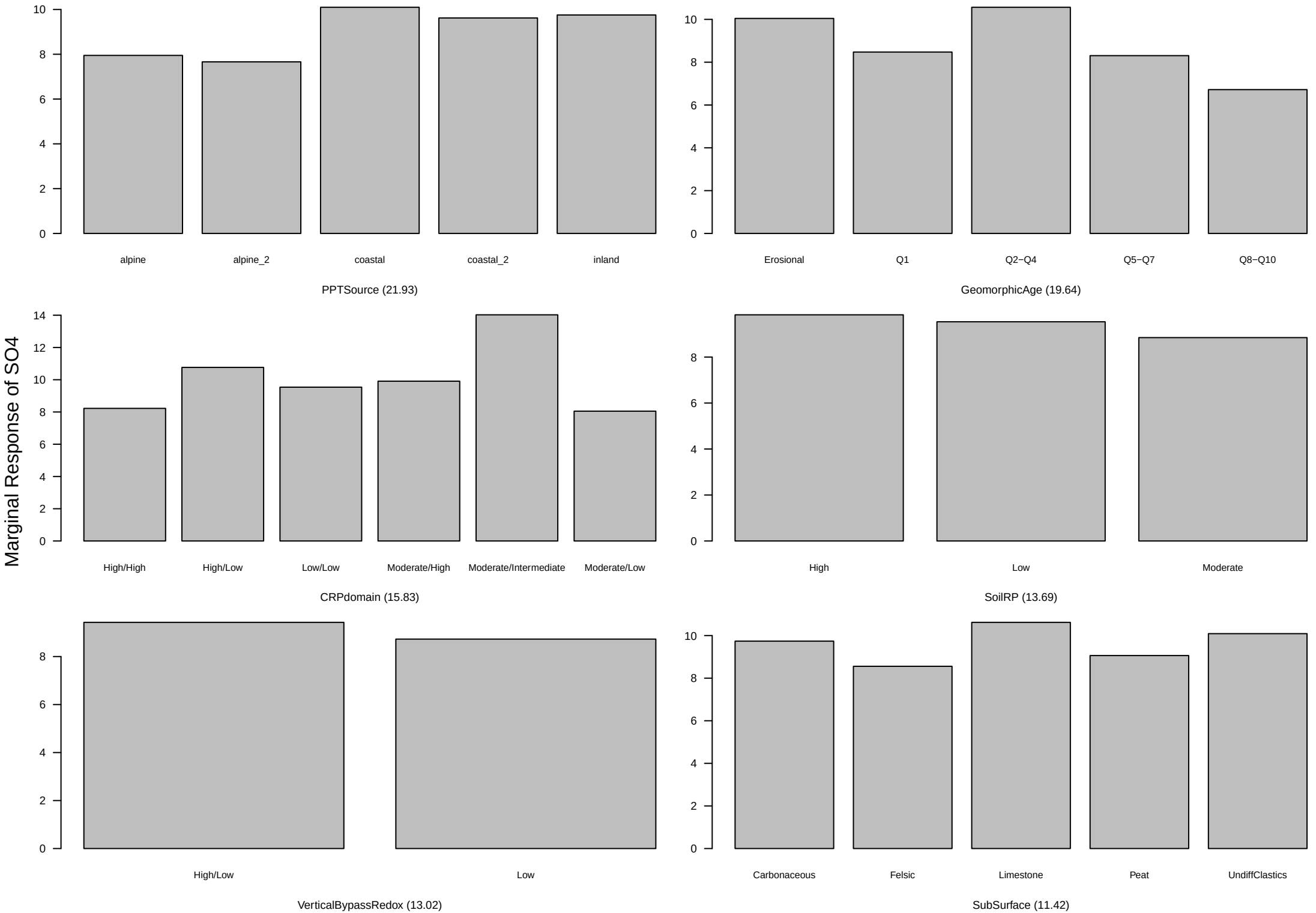
PPTSource (21.53)

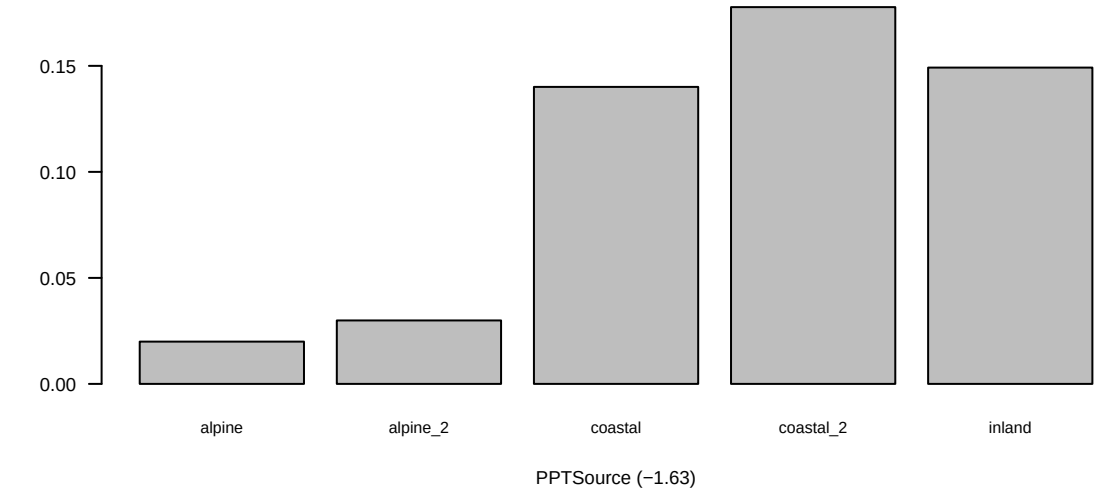

Marginal Response of DOField
0
1
2
3
4
5
6
Lowland
Riparian
Terrace
AquiferType (17.12)
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (12.04)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (10)
High
High/Low
Low
RiverInfluence (9.97)
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge (9.74)
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (9.55)


Marginal Response of TON
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (28.07)
High
High/Low
Low
RiverInfluence (11.09)
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (11.07)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (11.07)
Likely
NotLikely
Bypass (10.85)
High
Low
Moderate
SoilRP (10.77)

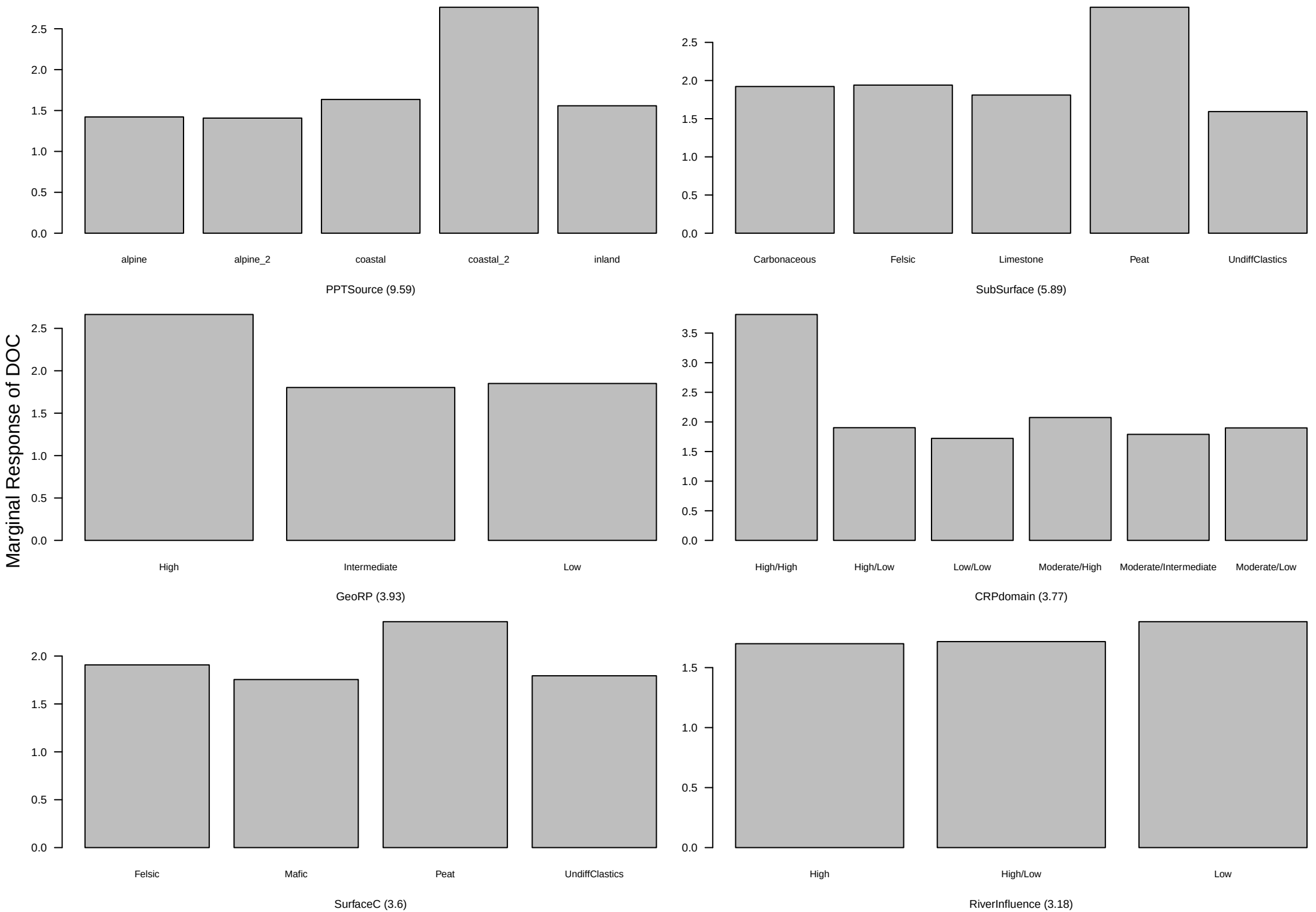

0.06
0.05
0.04
0.03
0.02
0.01
0.00
High
High/Low
Low
RiverConnectivity (14.59)
Marginal Response of Mn

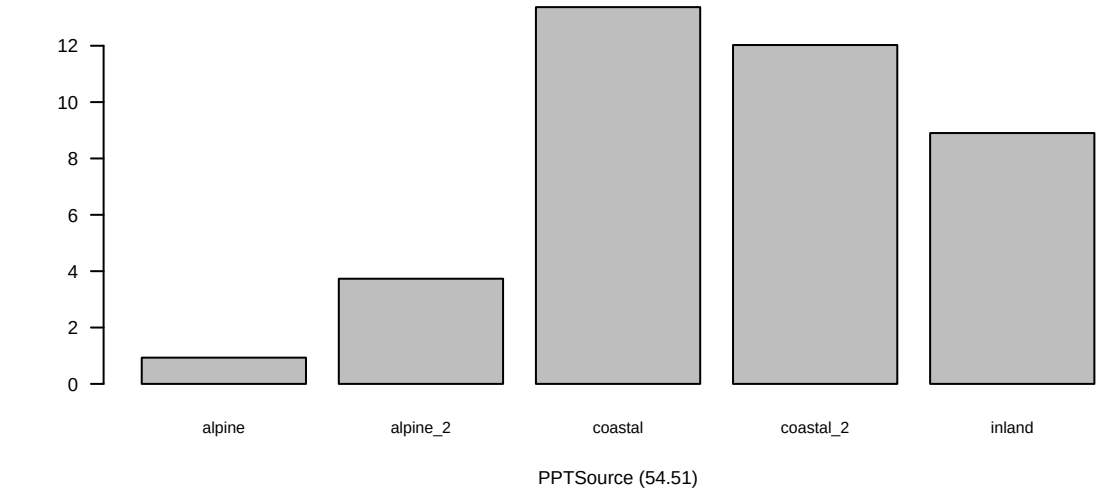

Marginal Response of Fe
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (12.5)
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (10.52)
High
Intermediate
Low
GeoRP (8.38)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (7.57)
High
High/Low
Low
RiverConnectivity (7.05)
Likely
NotLikely
Bypass (6.41)

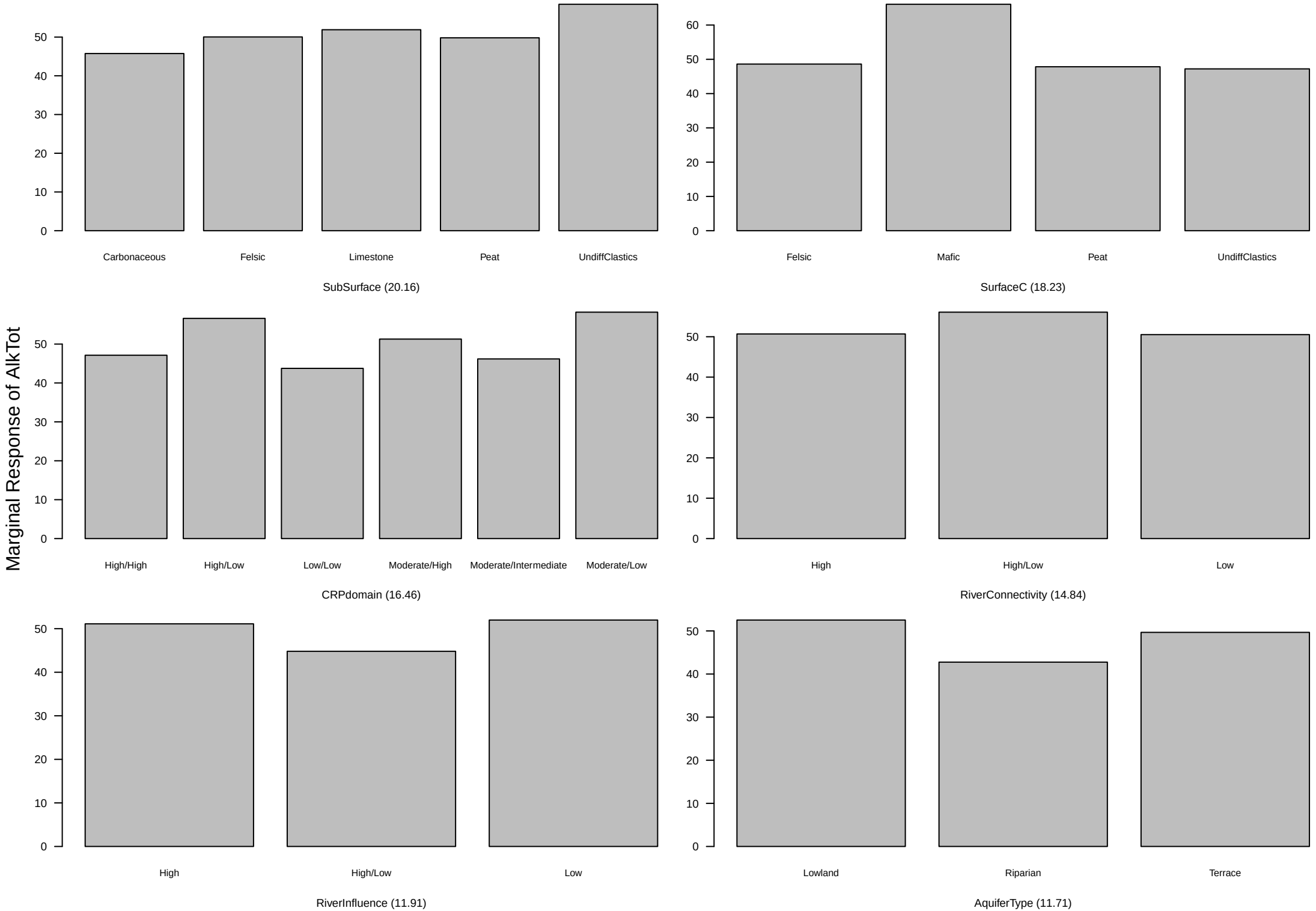

Marginal Response of SO4
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (21.93)
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge (19.64)
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (15.83)
High
Low
Moderate
SoilRP (13.69)
High/Low
Low
VerticalBypassRedox (13.02)
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (11.42)

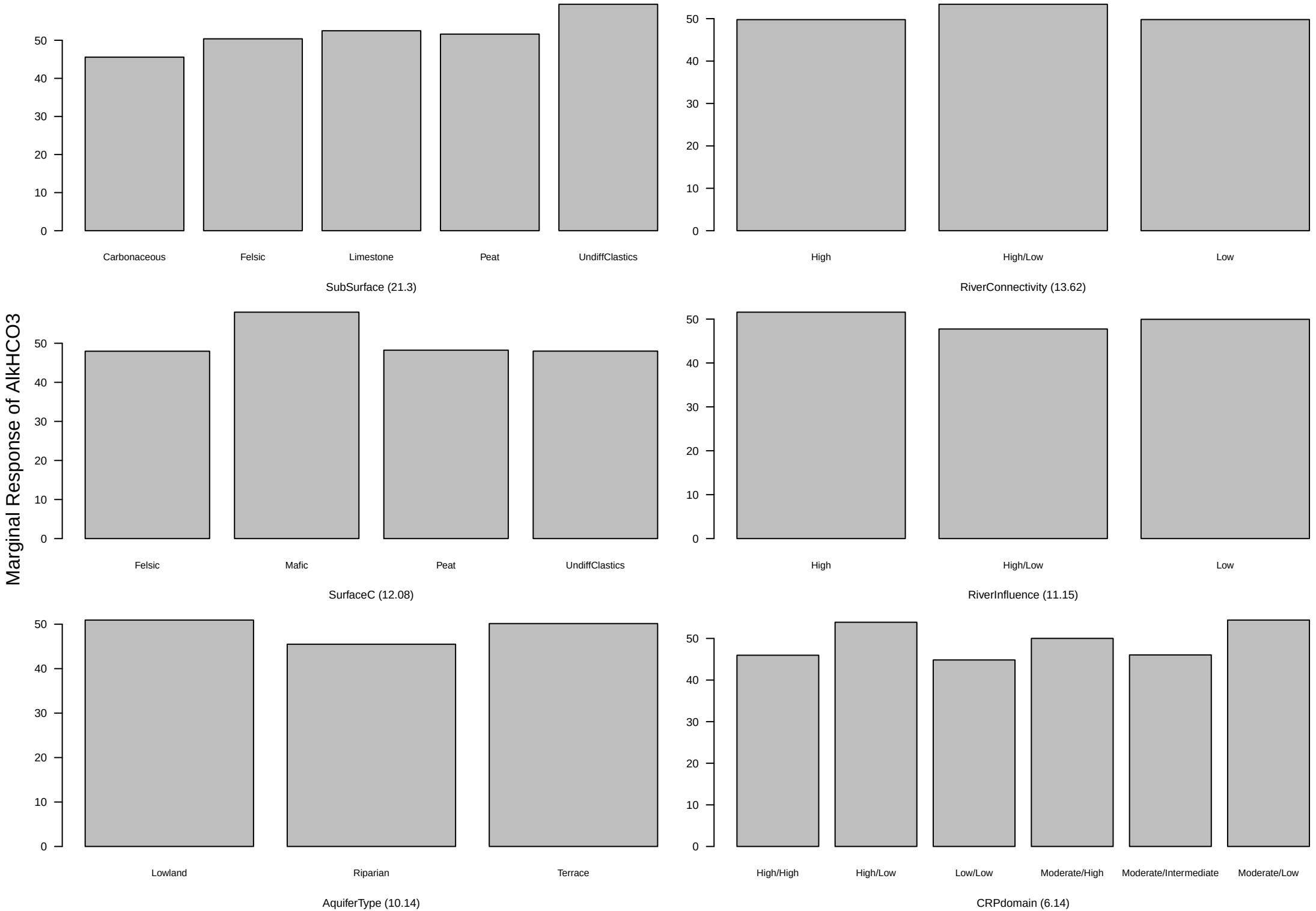


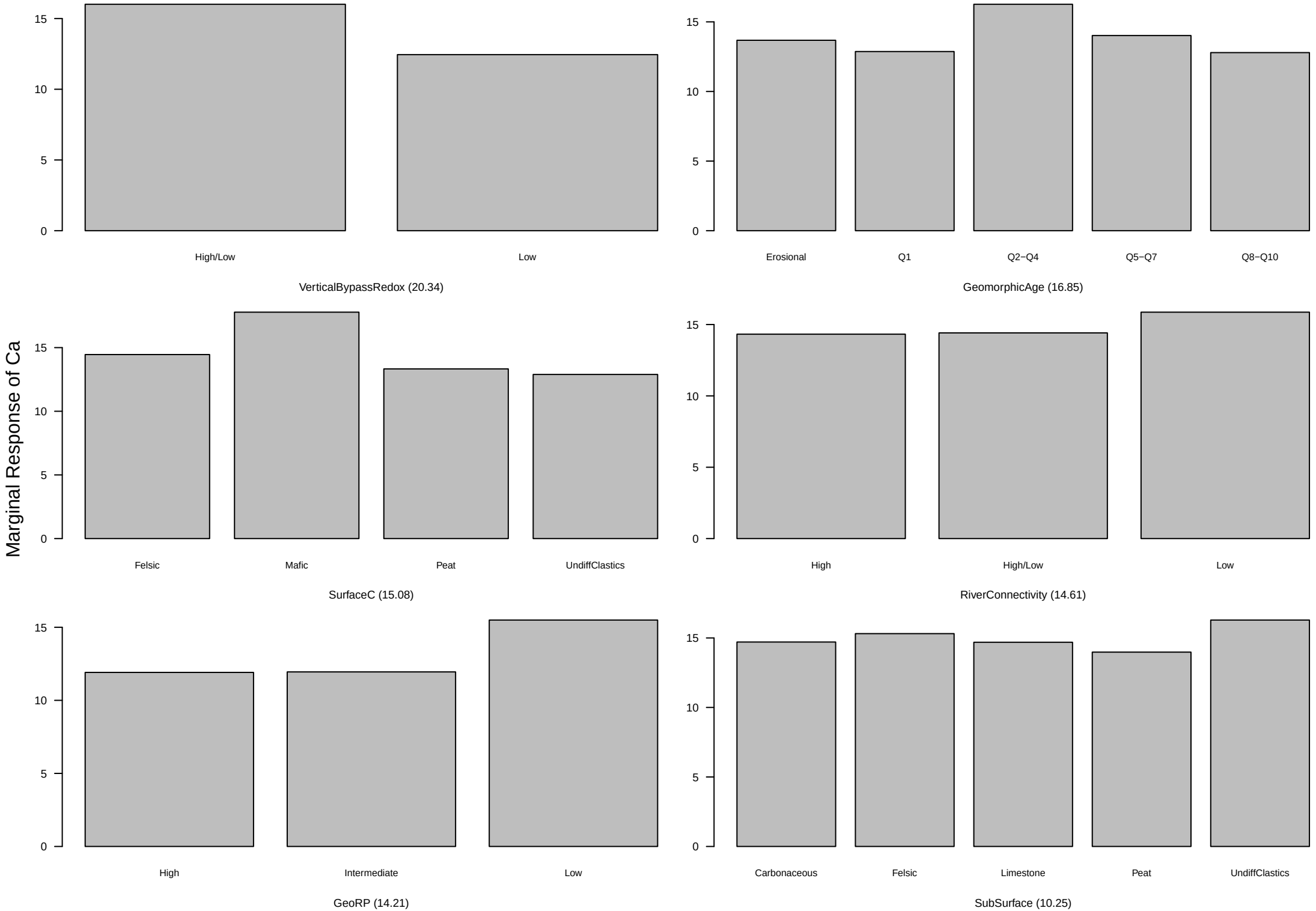
Marginal Response of NH4

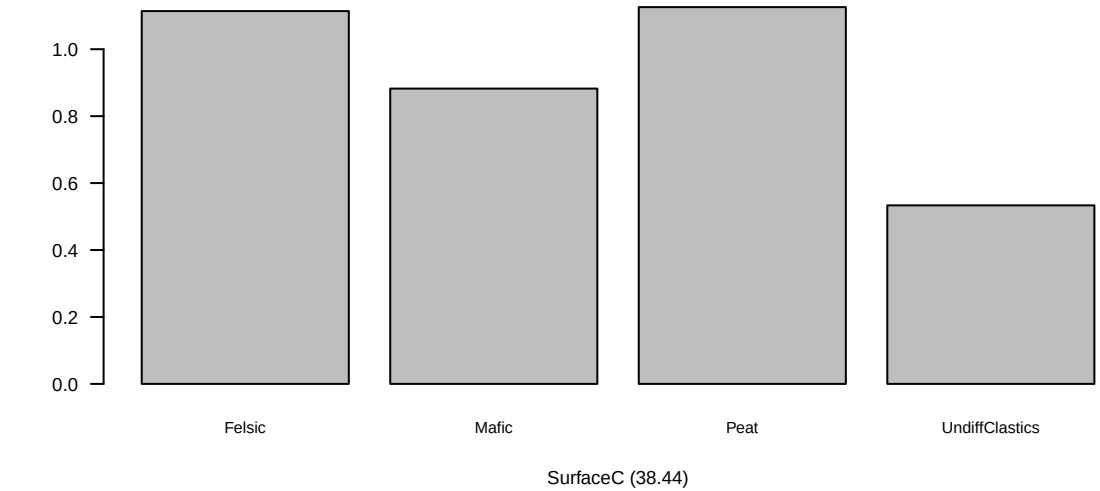

Marginal Response of DOC
PPTSource (9.59)
alpine
alpine_2
coastal
coastal_2
inland
SubSurface (5.89)
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
GeoRP (3.93)
High
Intermediate
Low
CRPdomain (3.77)
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
SurfaceC (3.6)
Felsic
Mafic
Peat
UndiffClastics
RiverInfluence (3.18)
High
High/Low
Low


12
10
8
6
4
2
0
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (54.51)
Marginal Response of TOC

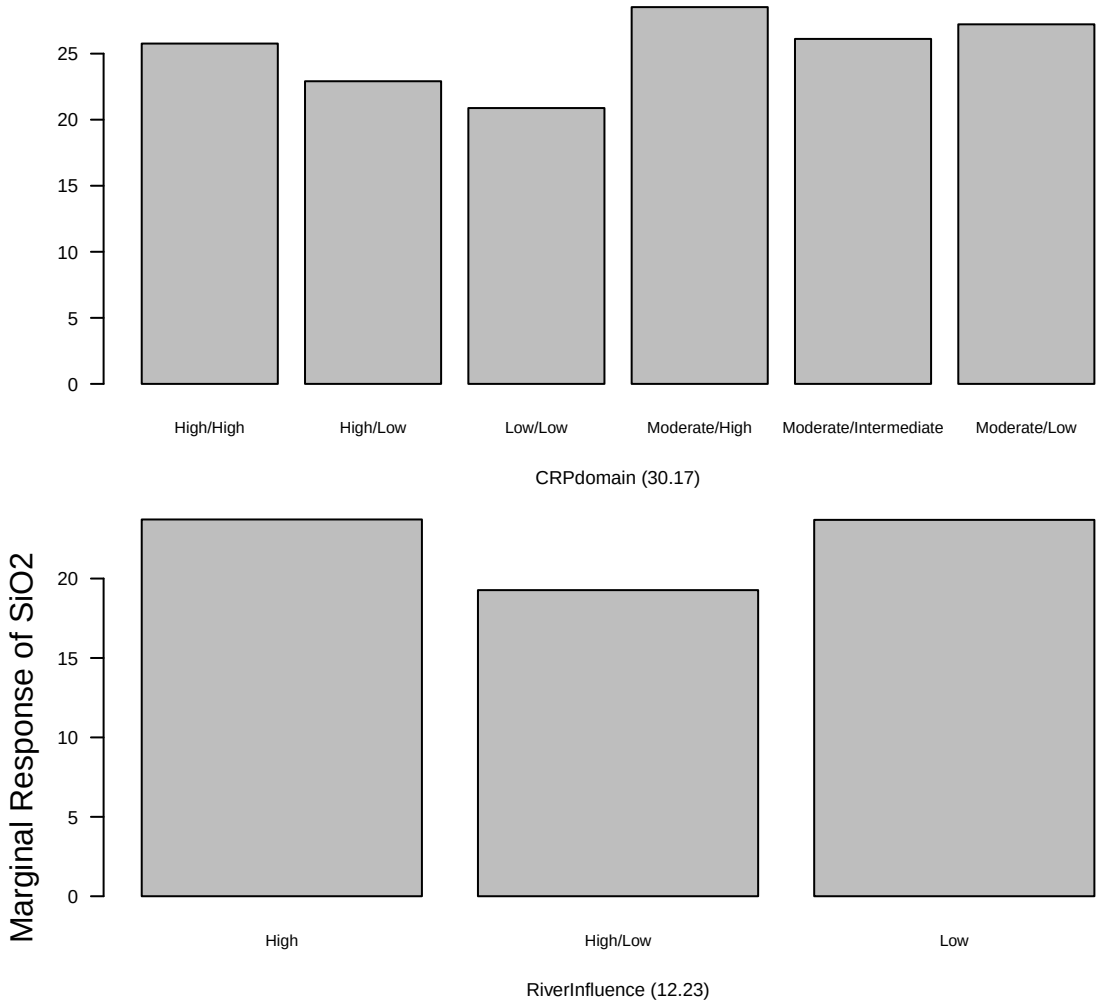

Marginal Response of AlkTot
0
10
20
30
40
50
60
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (20.16)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (18.23)
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (16.46)
High
High/Low
Low
RiverConnectivity (14.84)
High
High/Low
Low
RiverInfluence (11.91)
Lowland
Riparian
Terrace
AquiferType (11.71)

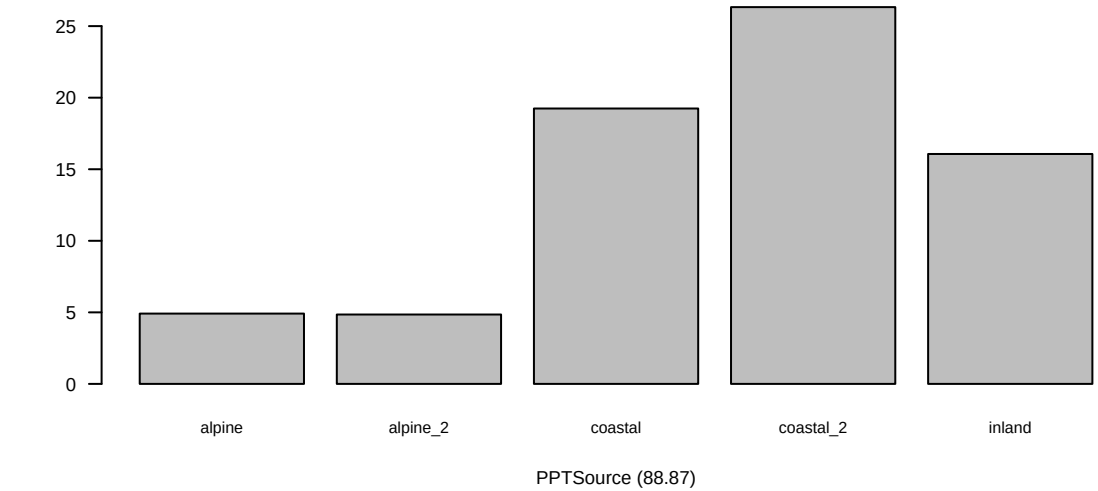

Marginal Response of AlkHCO3
0
10
20
30
40
50
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (21.3)
High
High/Low
Low
RiverConnectivity (13.62)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (12.08)
High
High/Low
Low
RiverInfluence (11.15)
Lowland
Riparian
Terrace
AquiferType (10.14)
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (6.14)


Marginal Response of Ca
15
10
5
0
High/Low
Low
VerticalBypassRedox (20.34)
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge (16.85)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (15.08)
High
High/Low
Low
RiverConnectivity (14.61)
High
Intermediate
Low
GeoRP (14.21)
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (10.25)

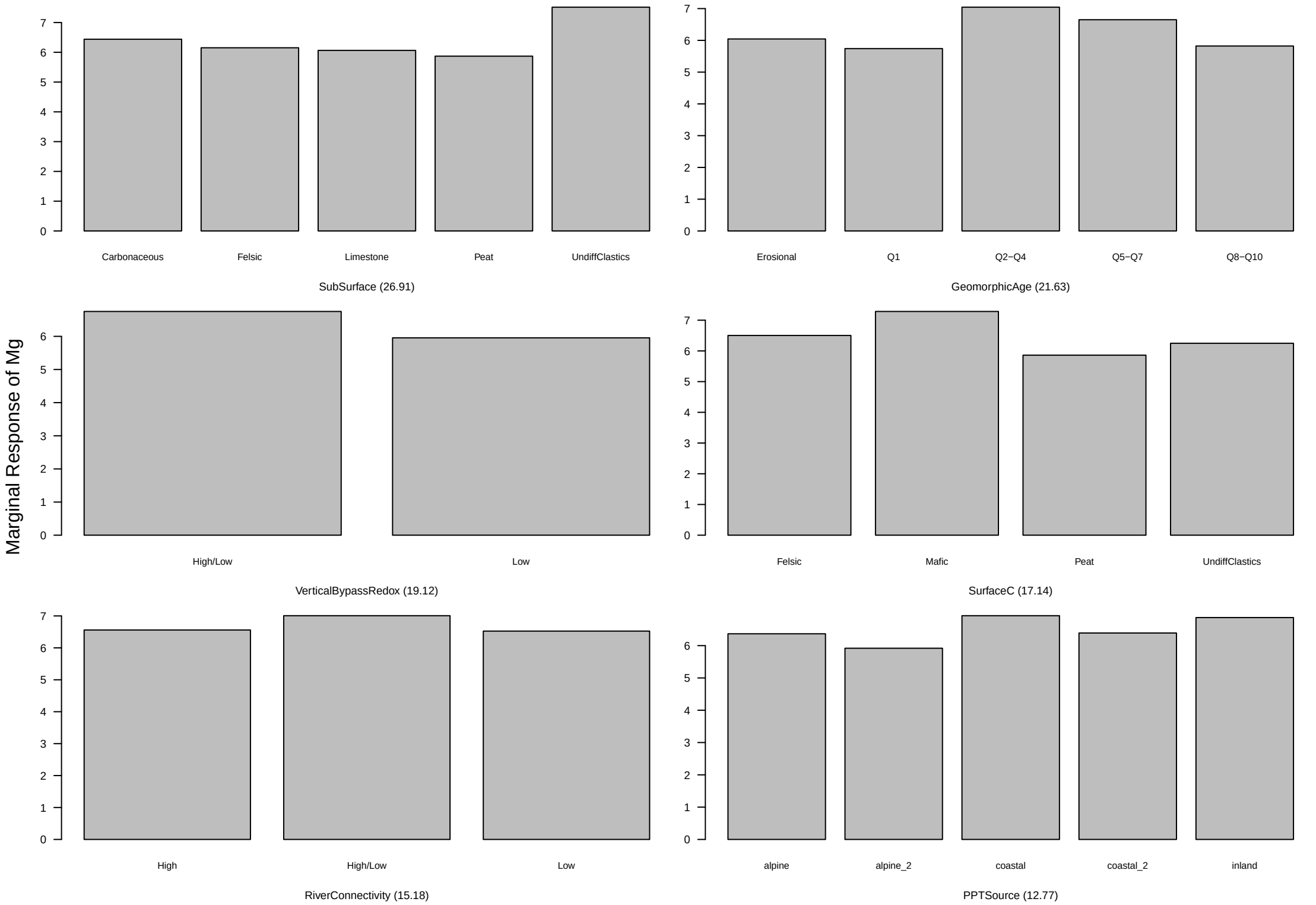


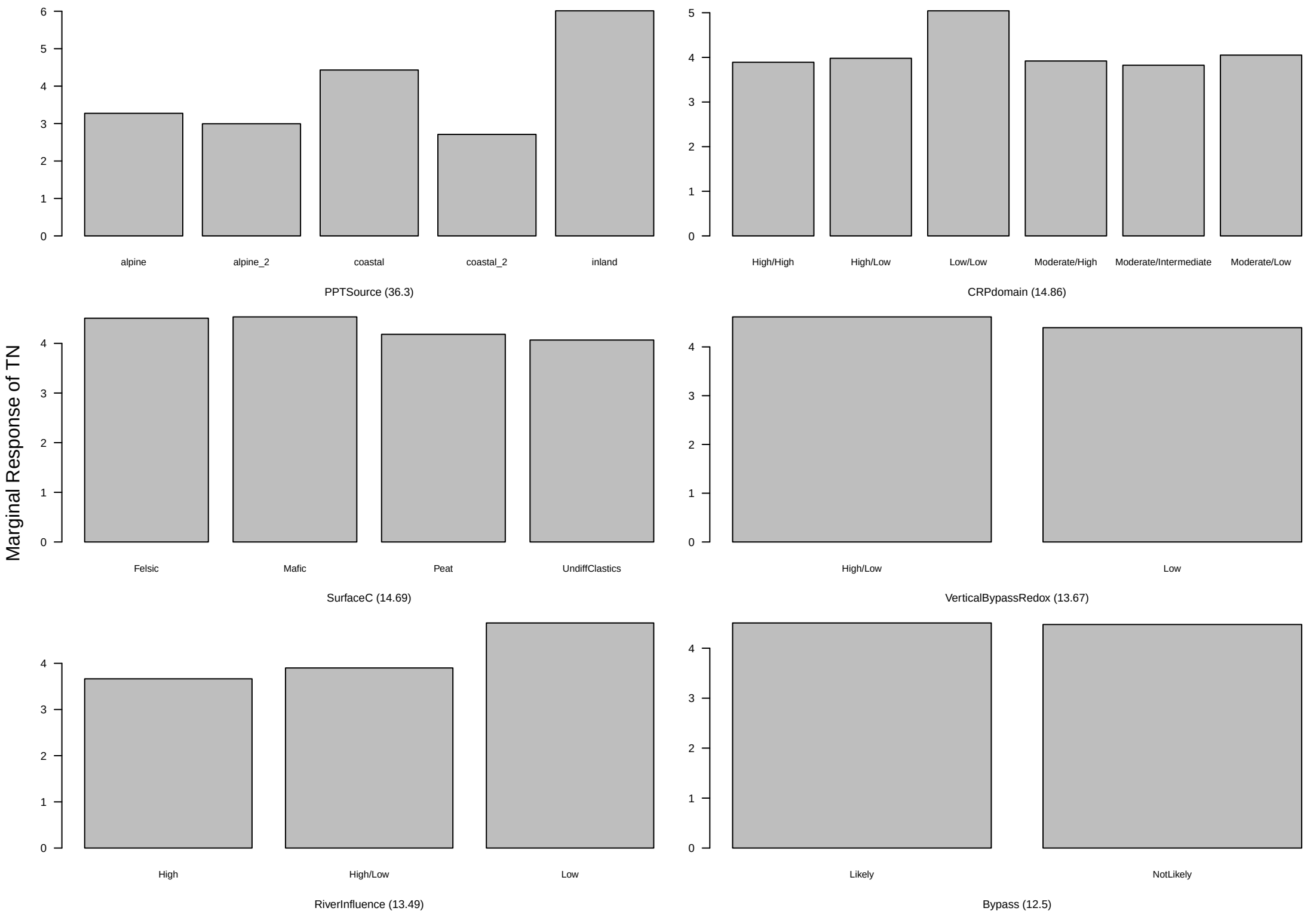
Marginal Response of K

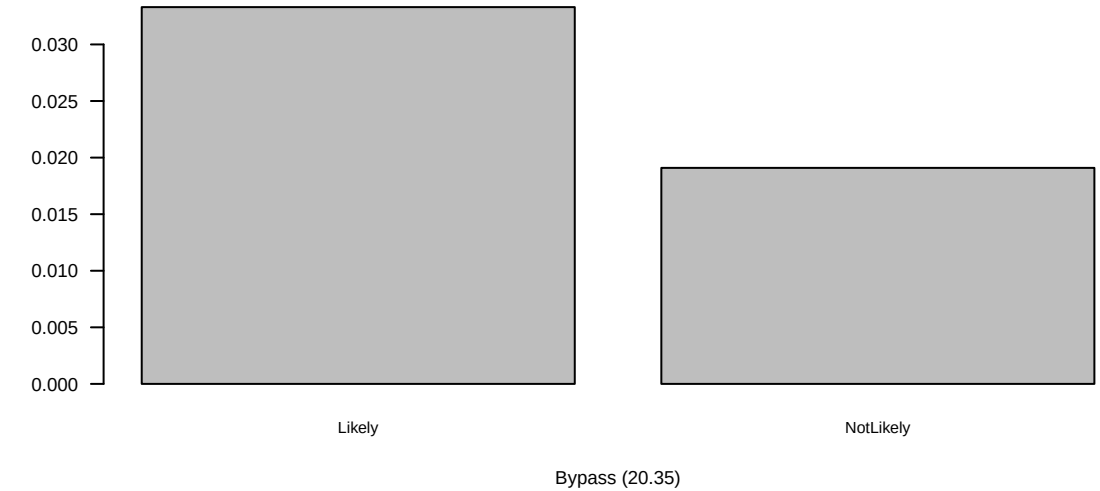

Marginal Response of SiO2
0
5
10
15
20
25
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (30.17)
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (24.44)
High
High/Low
Low
RiverInfluence (12.23)

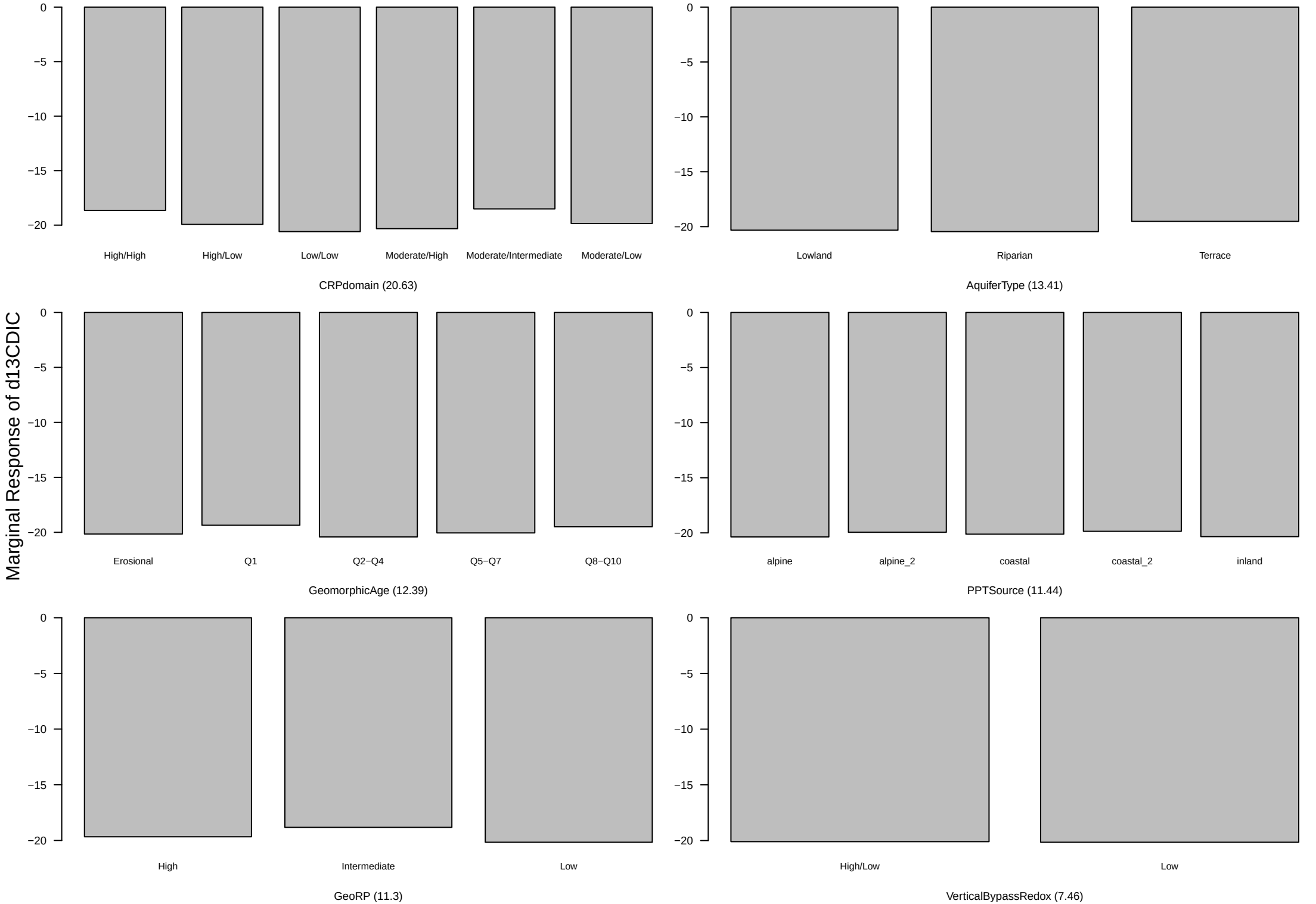


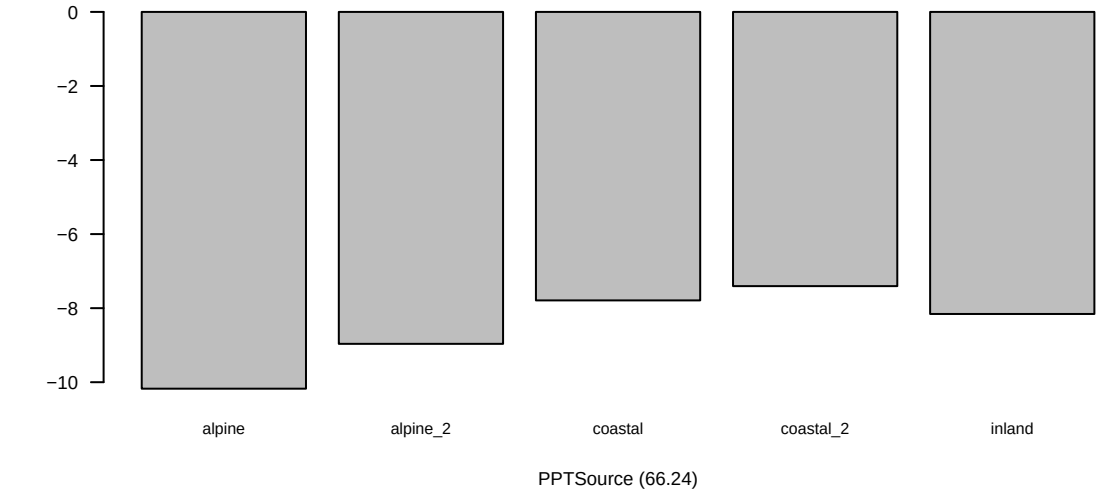
Marginal Response of Na


Marginal Response of Mg
0
1
2
3
4
5
6
7
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface (26.91)
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge (21.63)
High/Low
Low
VerticalBypassRedox (19.12)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (17.14)
High
High/Low
Low
RiverConnectivity (15.18)
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (12.77)

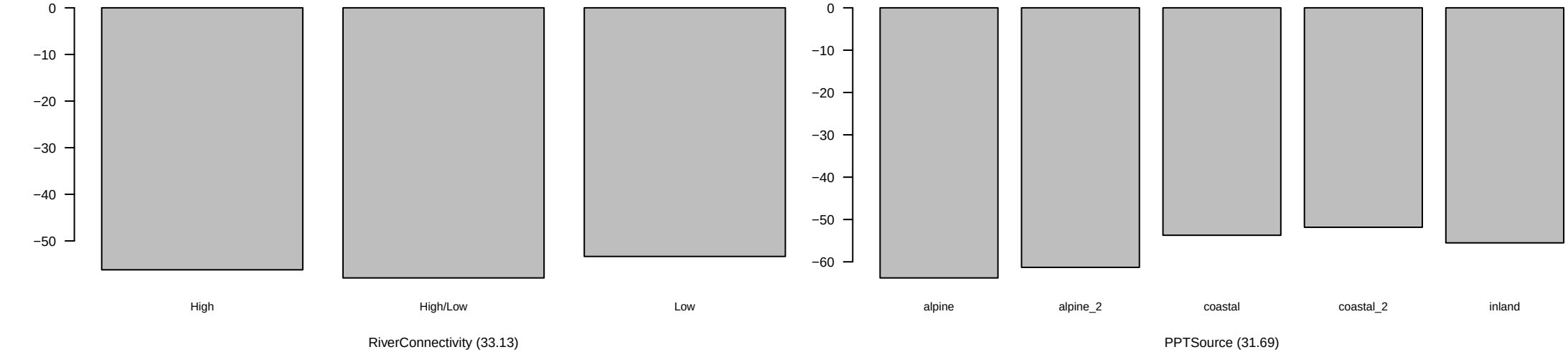

Marginal Response of TN
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (36.3)
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (14.86)
Felsic
Mafic
Peat
UndiffClastics
SurfaceC (14.69)
High/Low
Low
VerticalBypassRedox (13.67)
High
High/Low
Low
RiverInfluence (13.49)
Likely
NotLikely
Bypass (12.5)

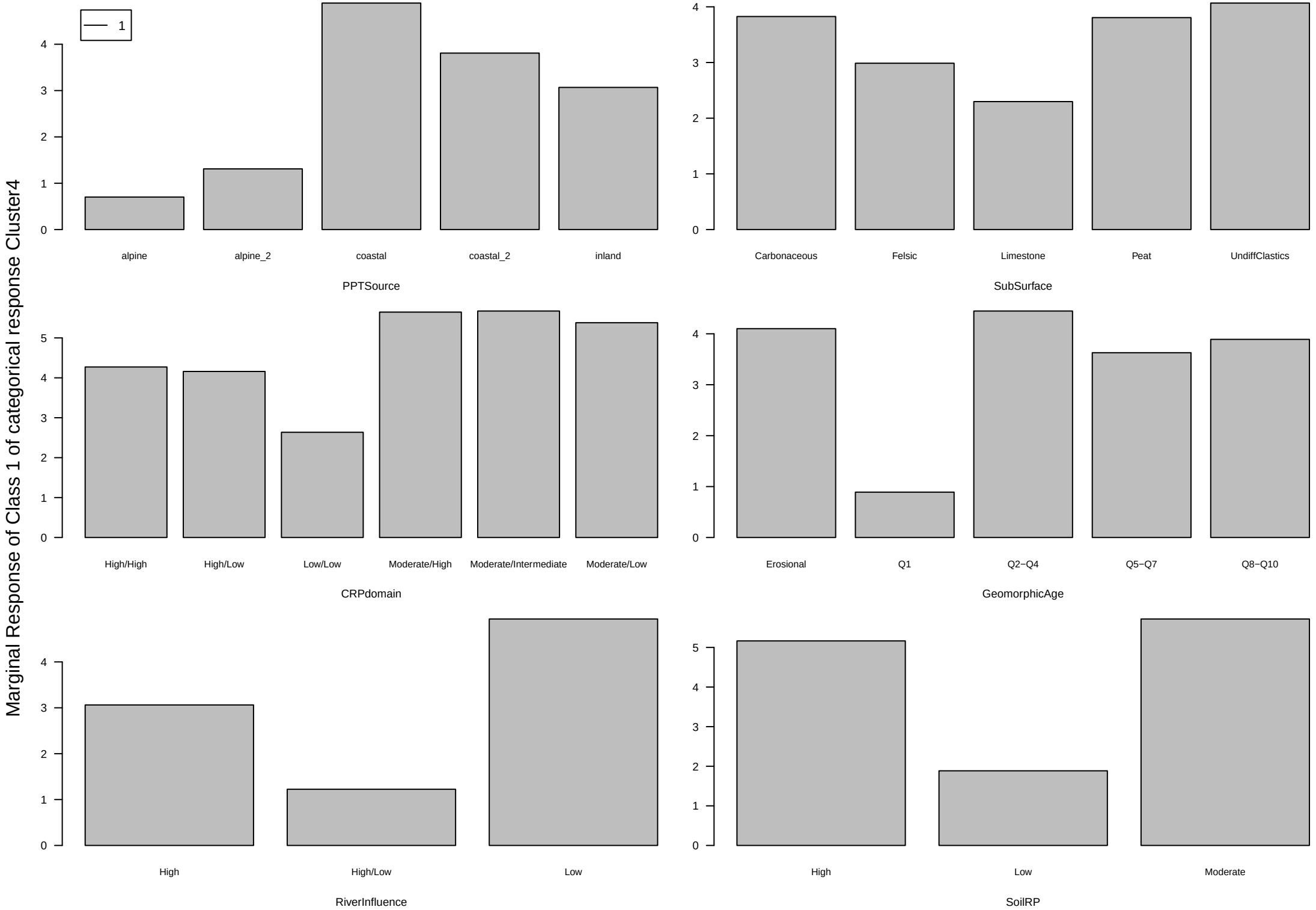

0.030
0.025
0.020
0.015
0.010
0.005
0.000
Likely
NotLikely
Bypass (20.35)
Marginal Response of DRP

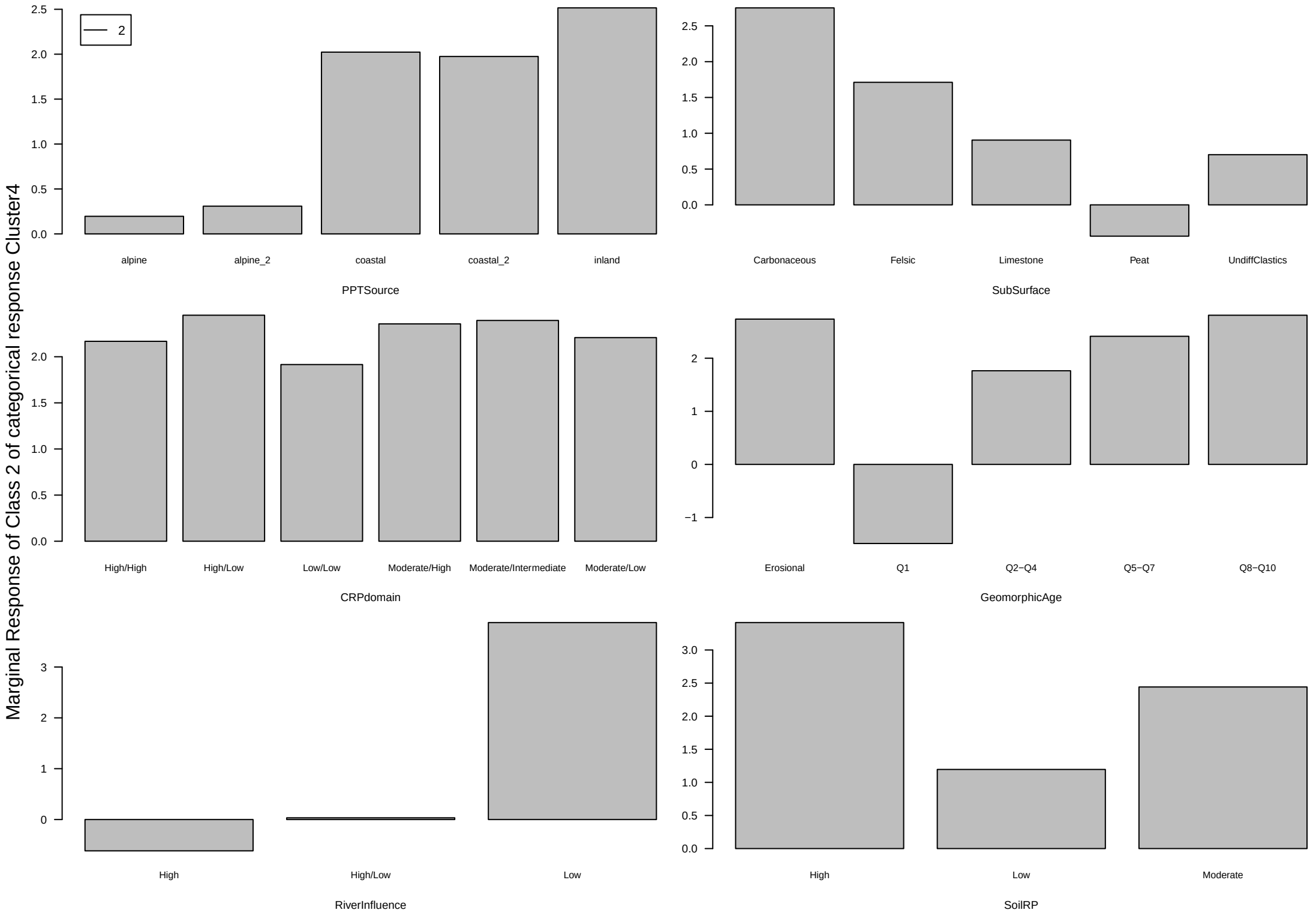

Marginal Response of d13CDIC
0
−5
−10
−15
−20
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain (20.63)
Lowland
Riparian
Terrace
AquiferType (13.41)
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge (12.39)
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (11.44)
High
Intermediate
Low
GeoRP (11.3)
High/Low
Low
VerticalBypassRedox (7.46)

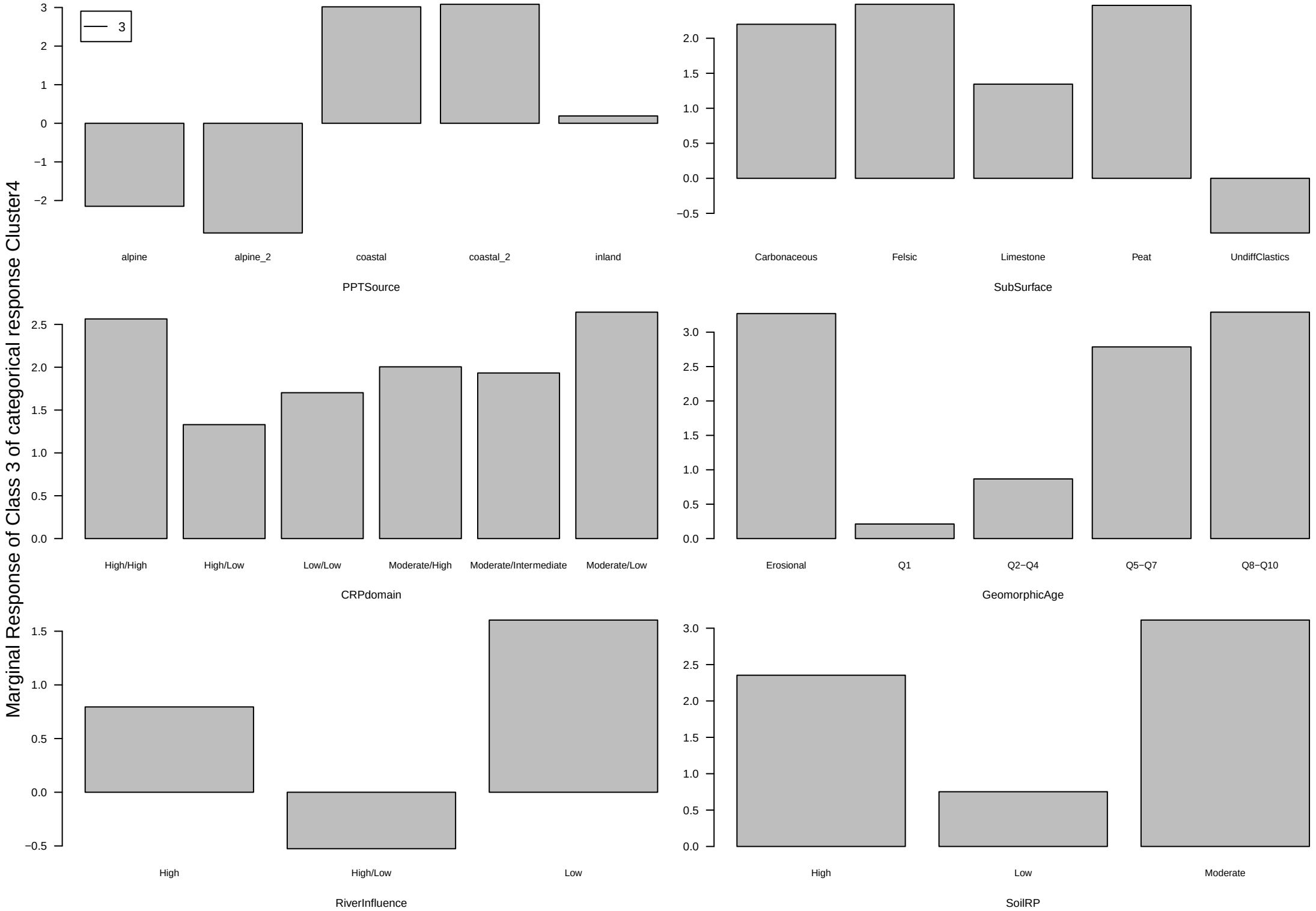


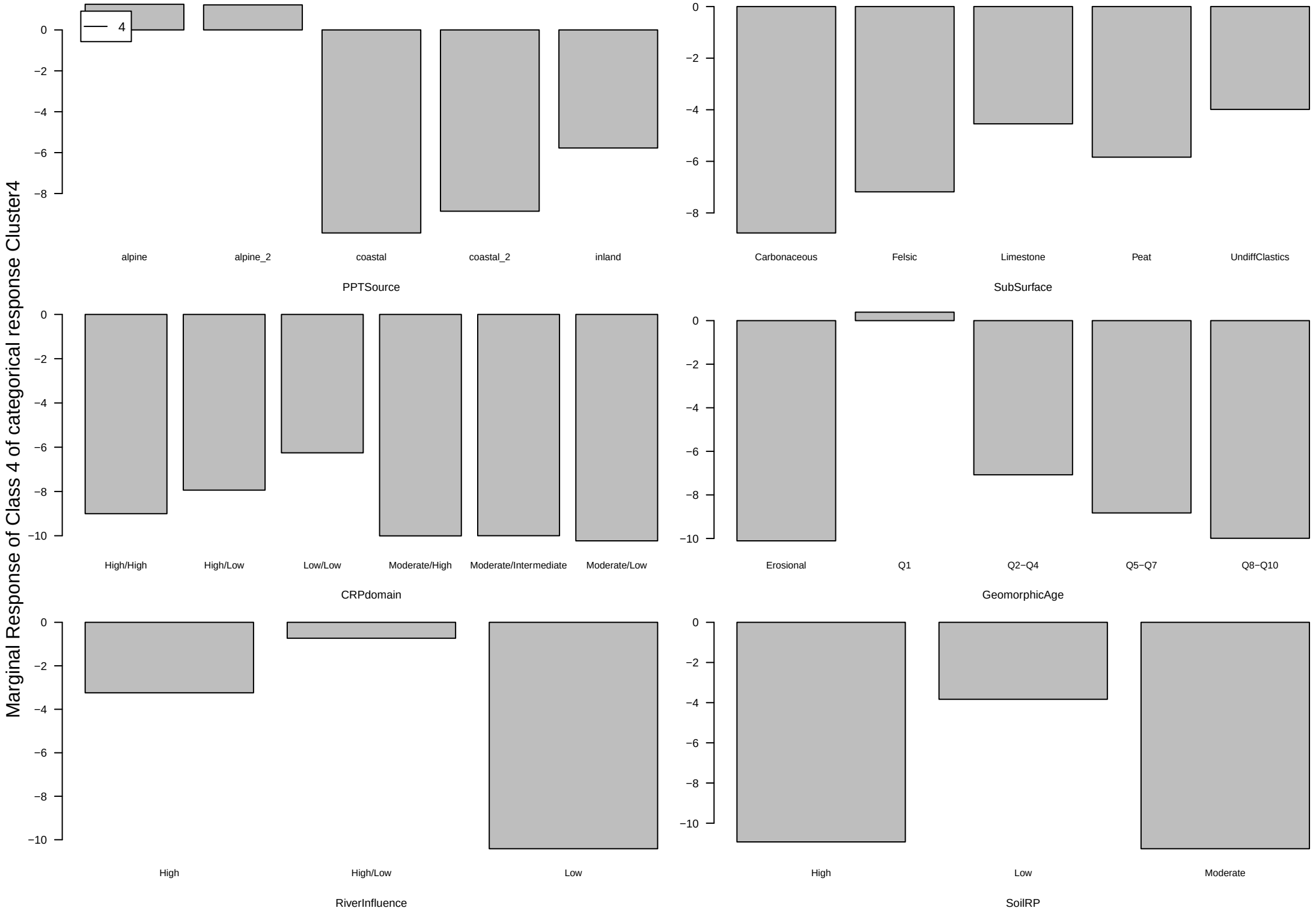
Marginal Response of d18OH2O

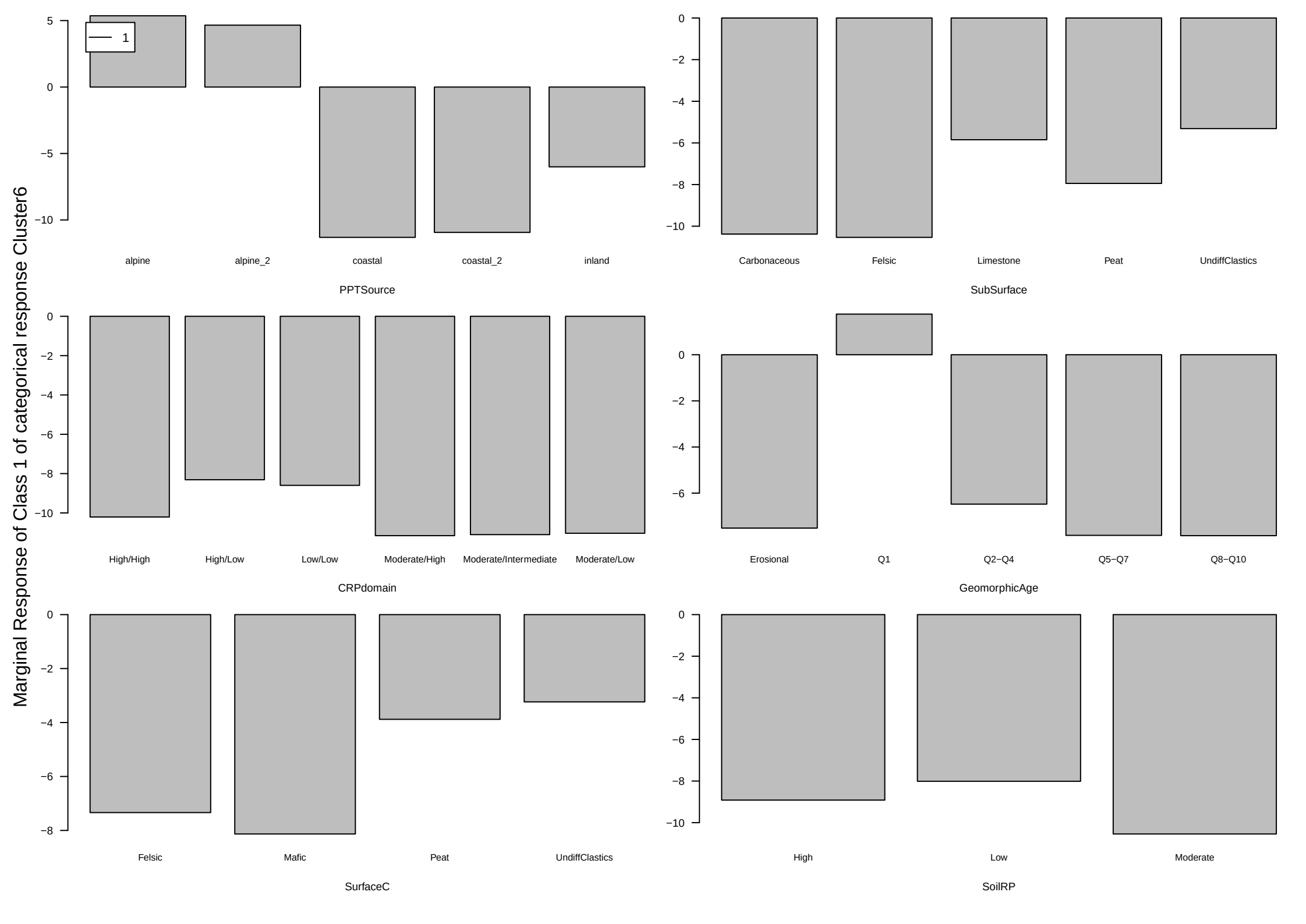

0
−10
−20
−30
−40
−50
High
High/Low
Low
RiverConnectivity (33.13)
0
−10
−20
−30
−40
−50
−60
alpine
alpine_2
coastal
coastal_2
inland
PPTSource (31.69)
Marginal Response of d2HH2O

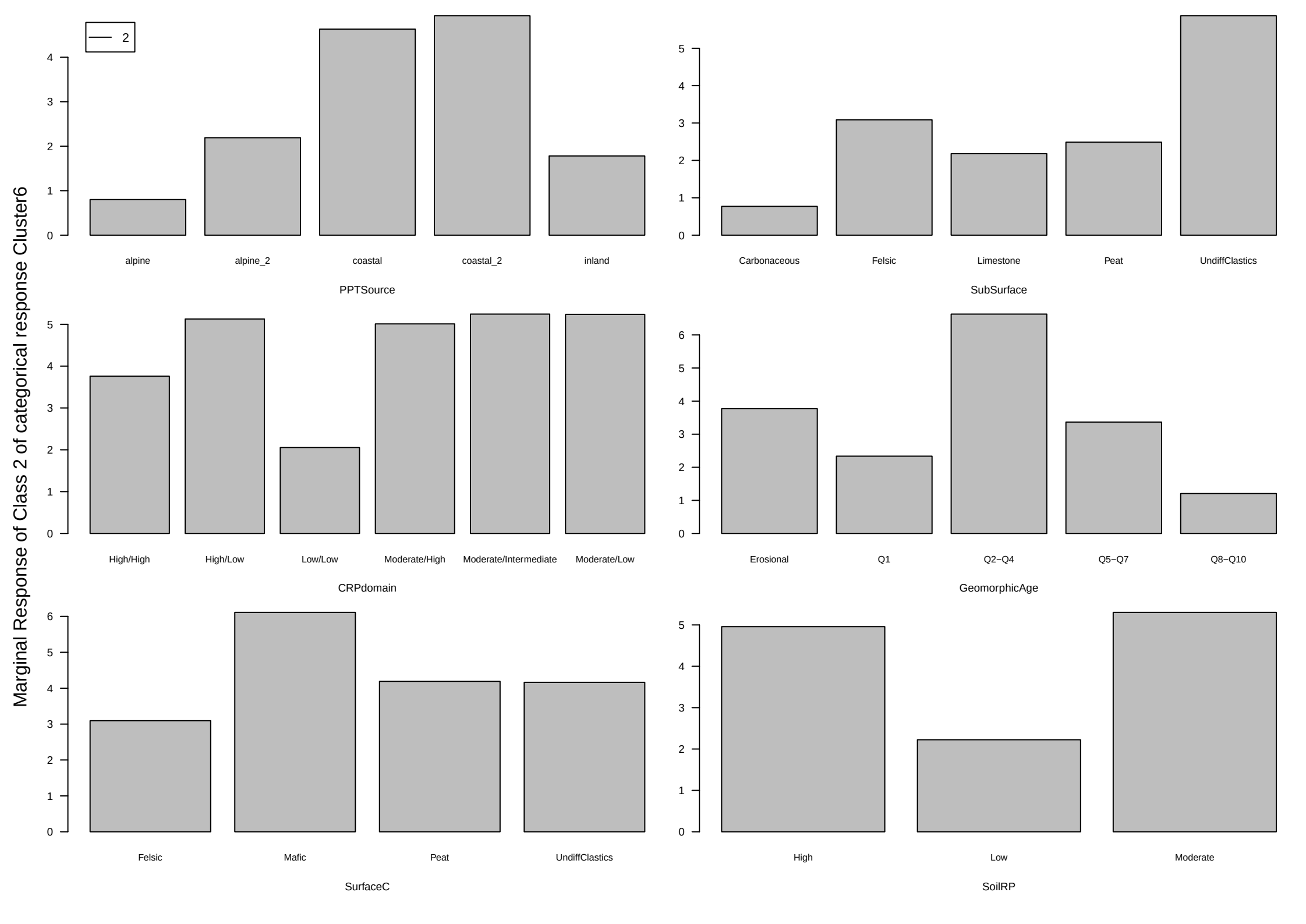

Marginal Response of Class 1 of categorical response Cluster4
1
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
High
High/Low
Low
RiverInfluence
High
Low
Moderate
SoilRP

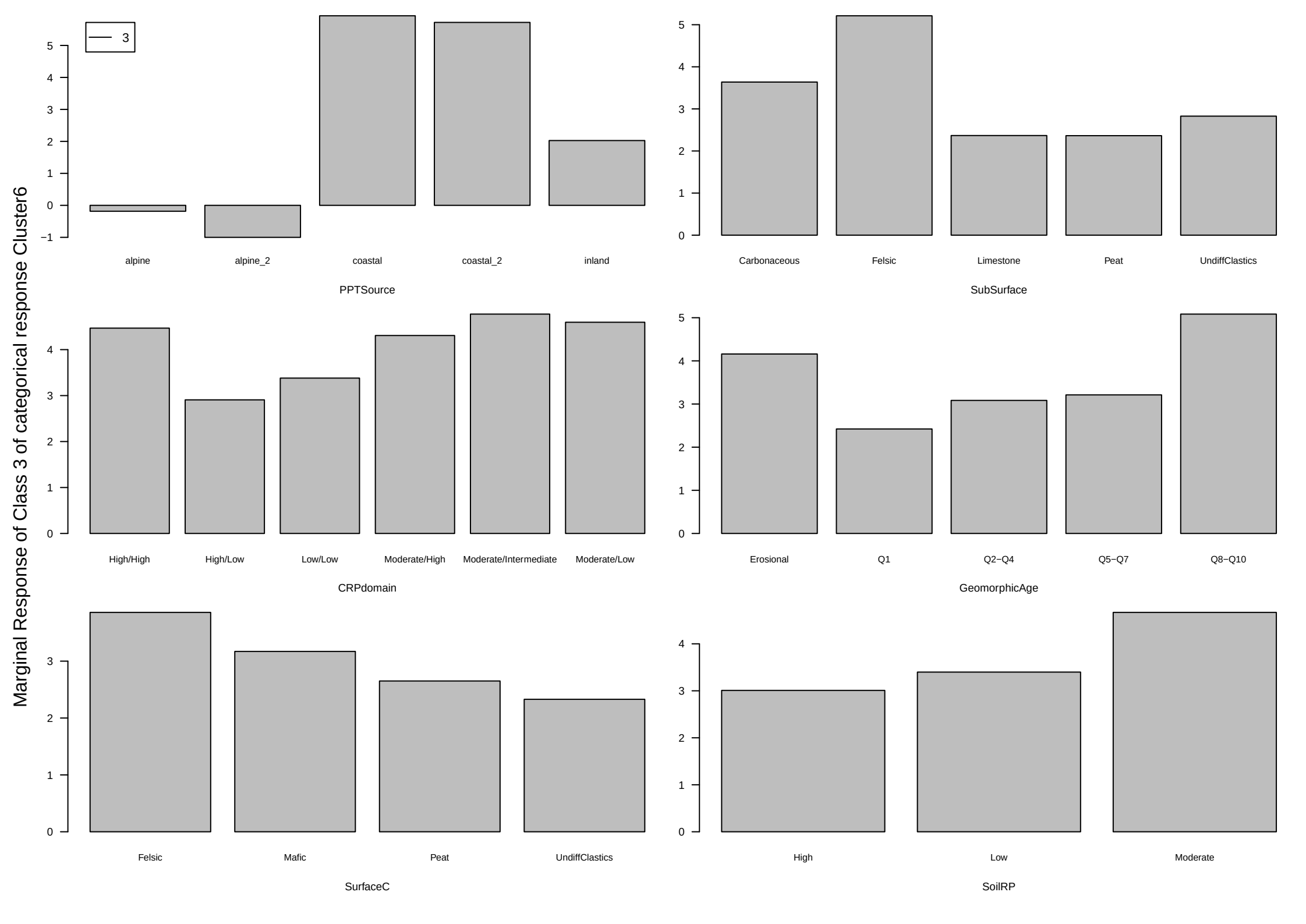

Marginal Response of Class 2 of categorical response Cluster4
2
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
RiverInfluence
High
High/Low
Low
SoilRP
High
Low
Moderate

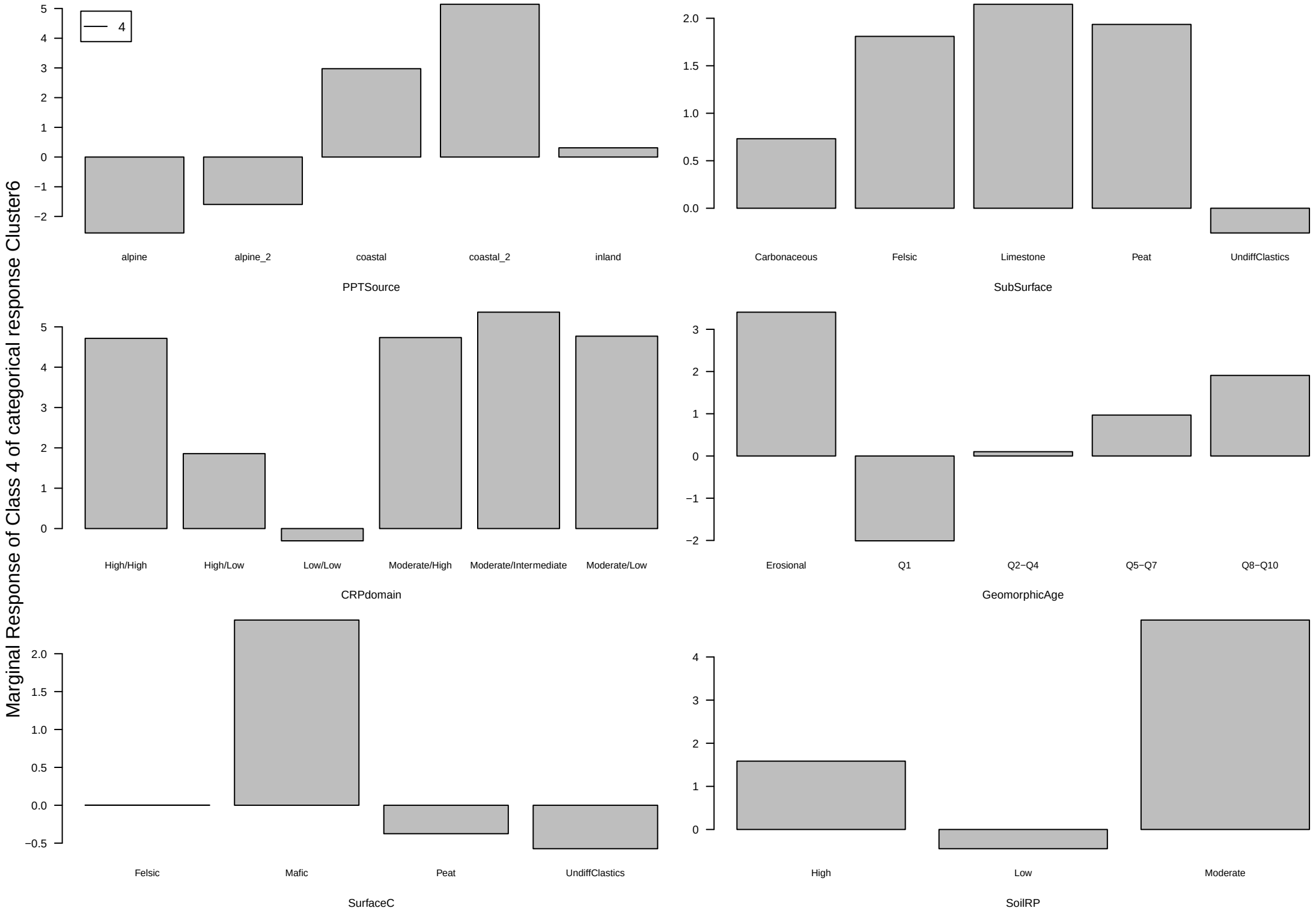

Marginal Response of Class 3 of categorical response Cluster4
3
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
High
High/Low
Low
RiverInfluence
High
Low
Moderate
SoilRP

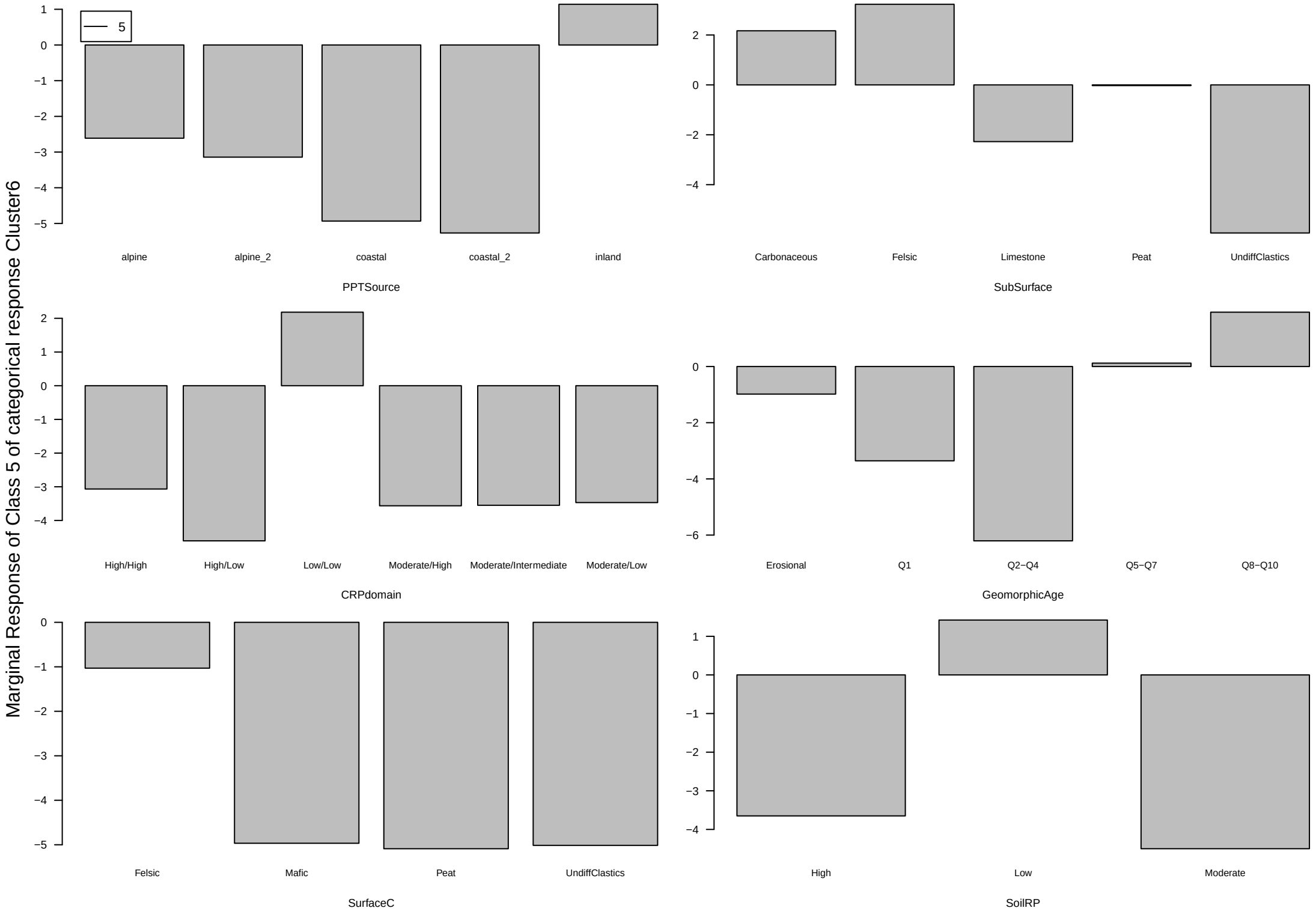

Marginal Response of Class 4 of categorical response Cluster4
4
0
−2
−4
−6
−8
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
−10
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
High
High/Low
Low
RiverInfluence
Moderate
SoilRP

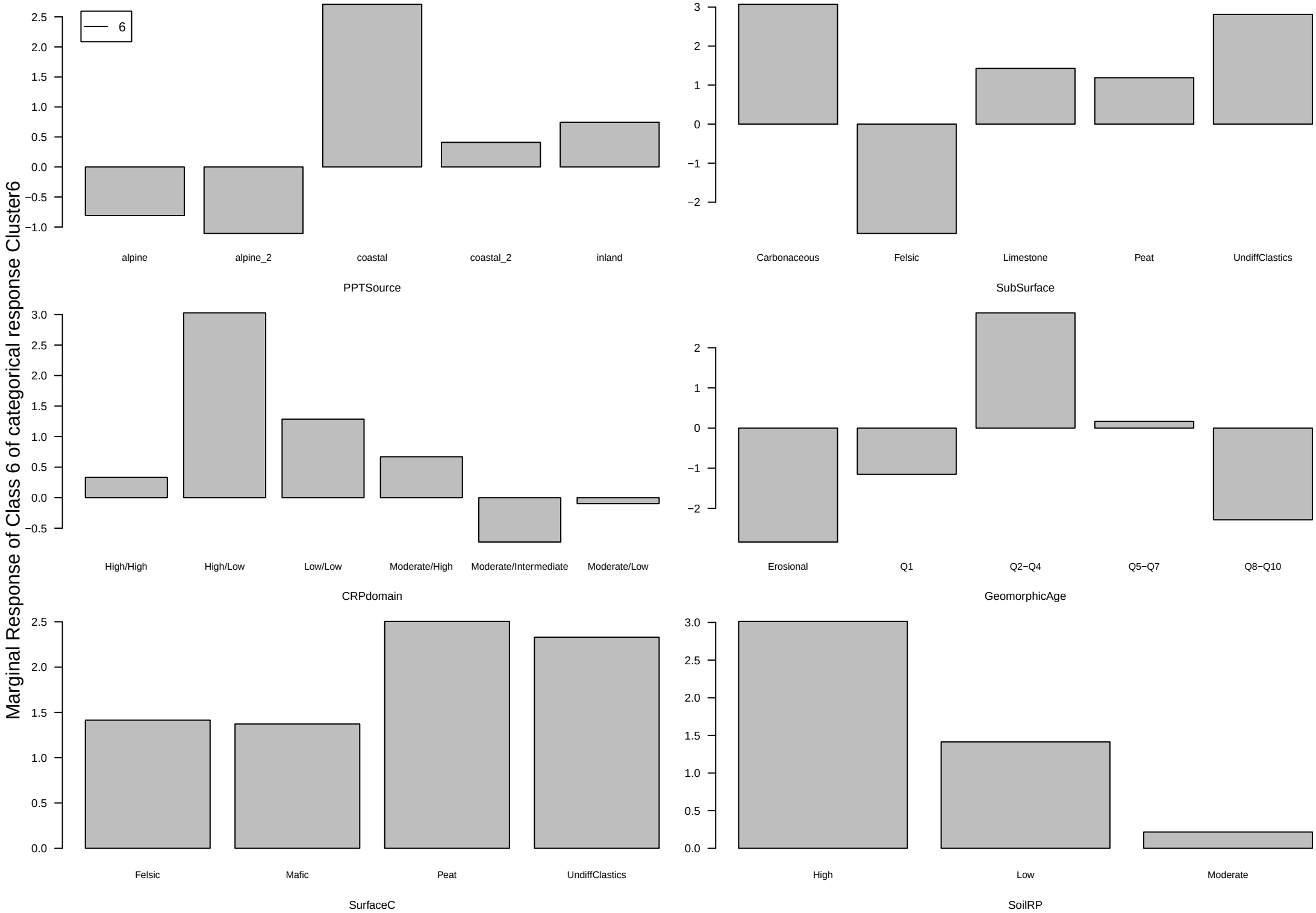

Marginal Response of Class 1 of categorical response Cluster6
1
5
0
−5
−10
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
0
−2
−4
−6
−8
−10
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
0
−2
−4
−6
−8
−10
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
0
−2
−4
−6
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
0
−2
−4
−6
−8
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
0
−2
−4
−6
−8
−10
High
Low
Moderate
SoilRP

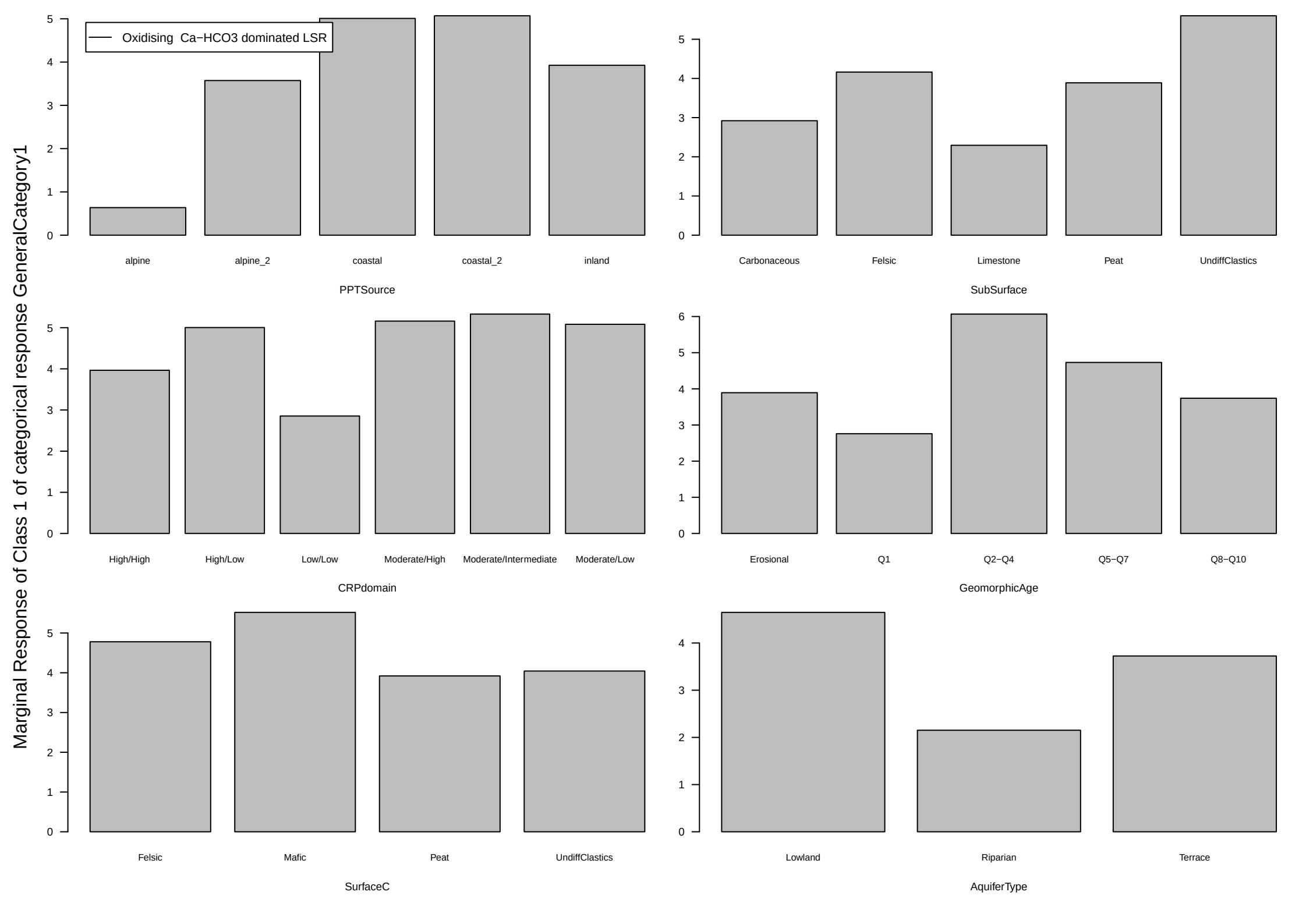

Marginal Response of Class 2 of categorical response Cluster6
2
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
SoilRP
High
Low
Moderate

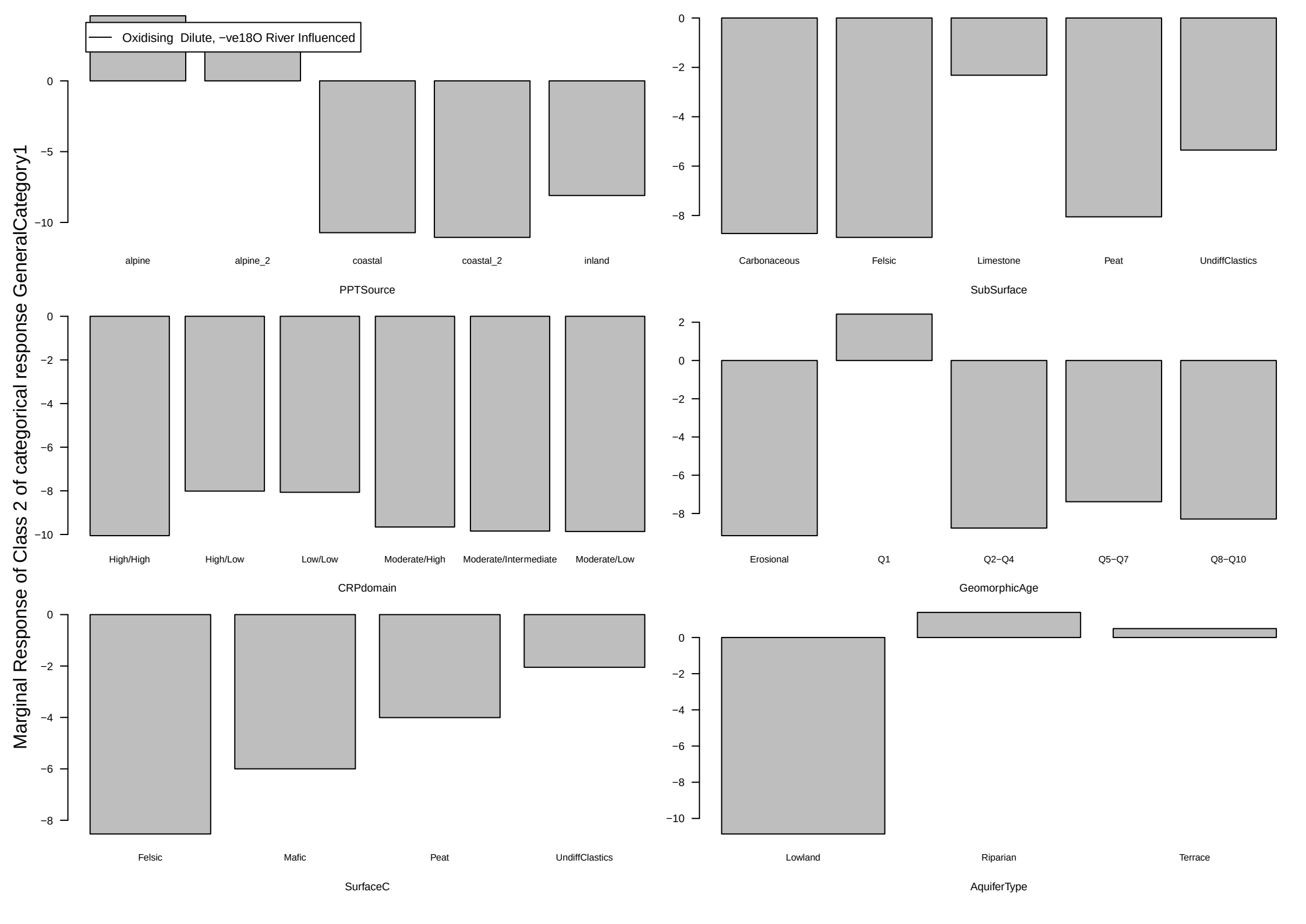

Marginal Response of Class 3 of categorical response Cluster6
3
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
SoilRP
High
Low
Moderate

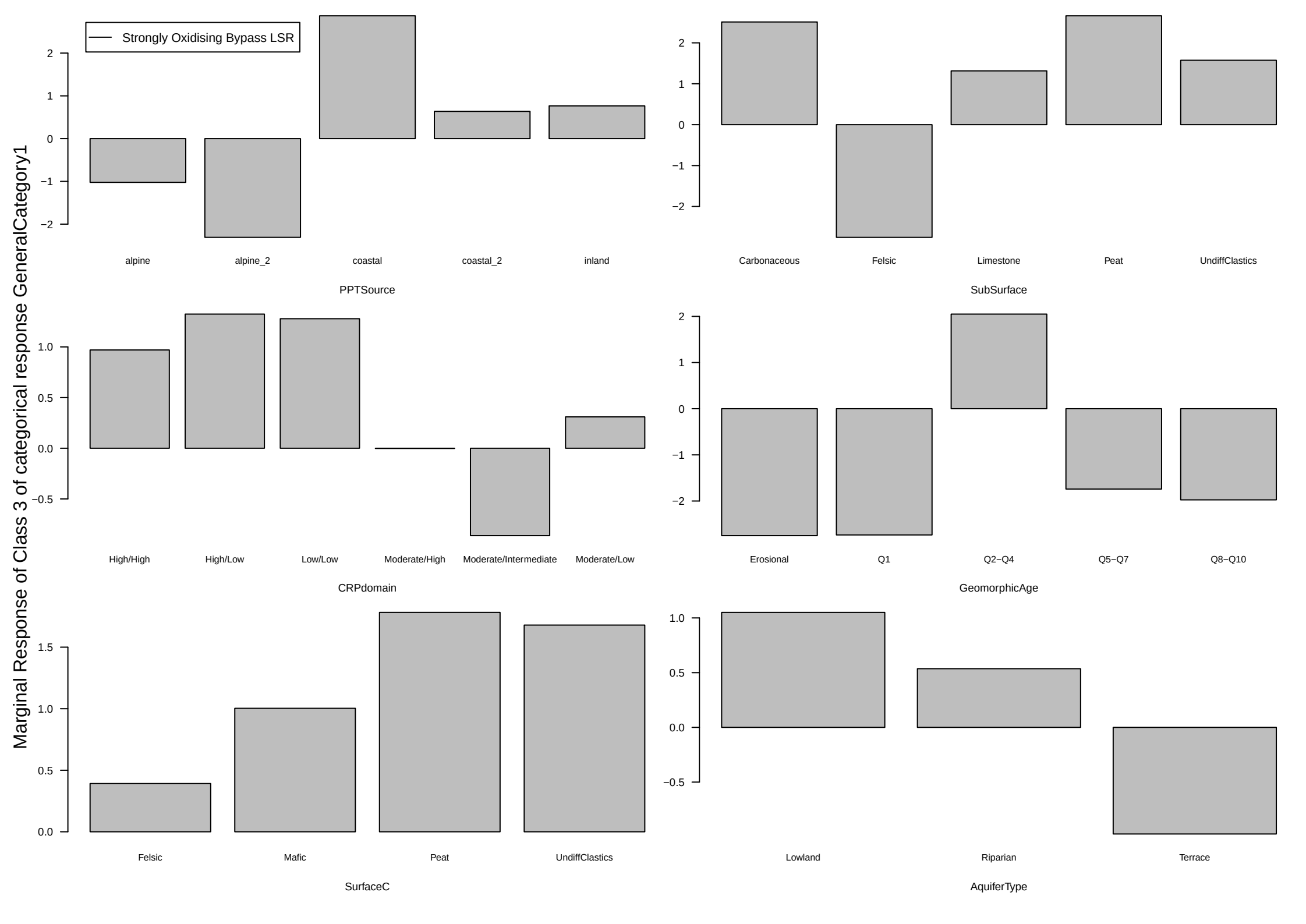

Marginal Response of Class 4 of categorical response Cluster6
4
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
High
Low
Moderate
SoilRP

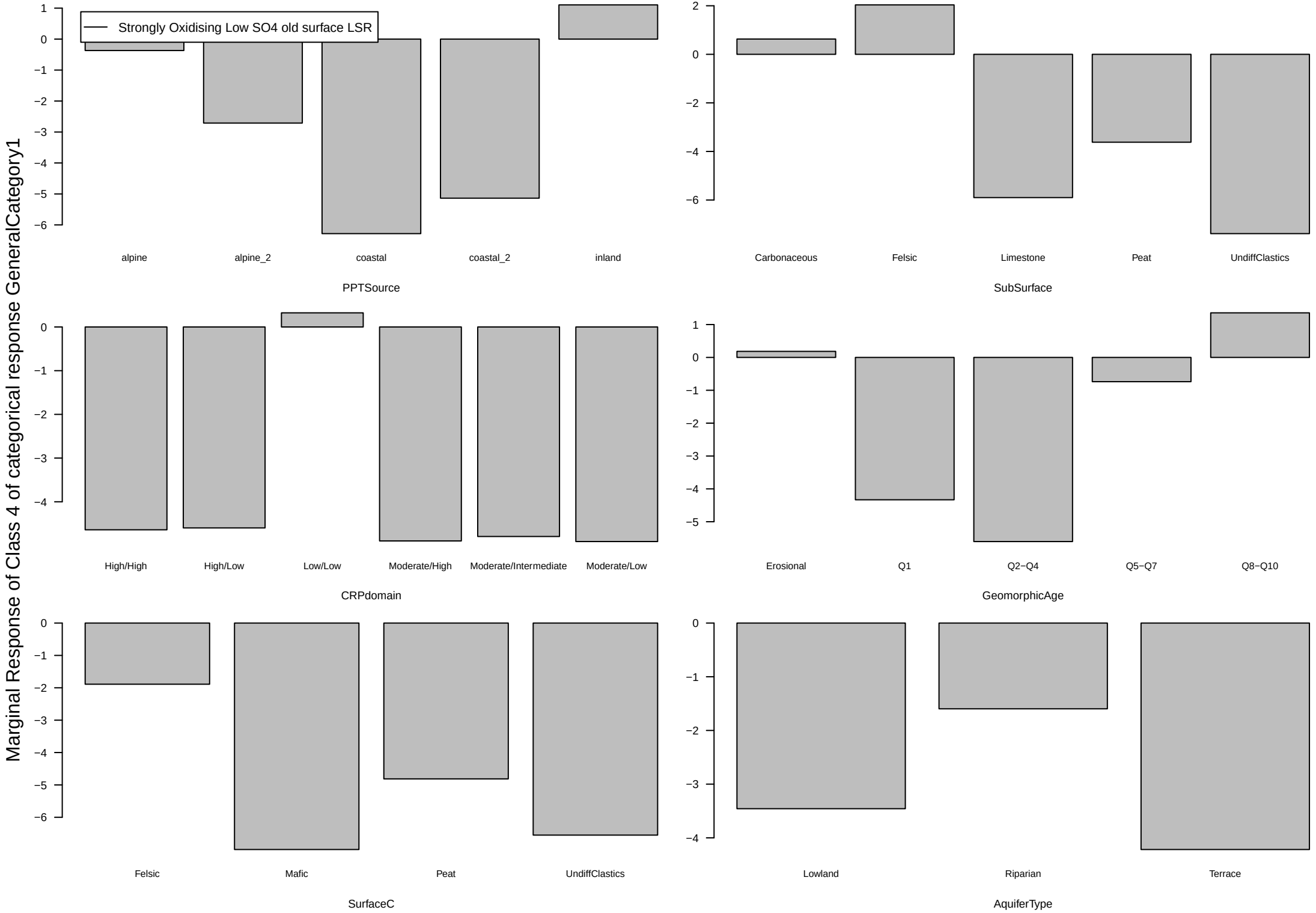

Marginal Response of Class 5 of categorical response Cluster6
5
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
SoilRP
High
Low
Moderate

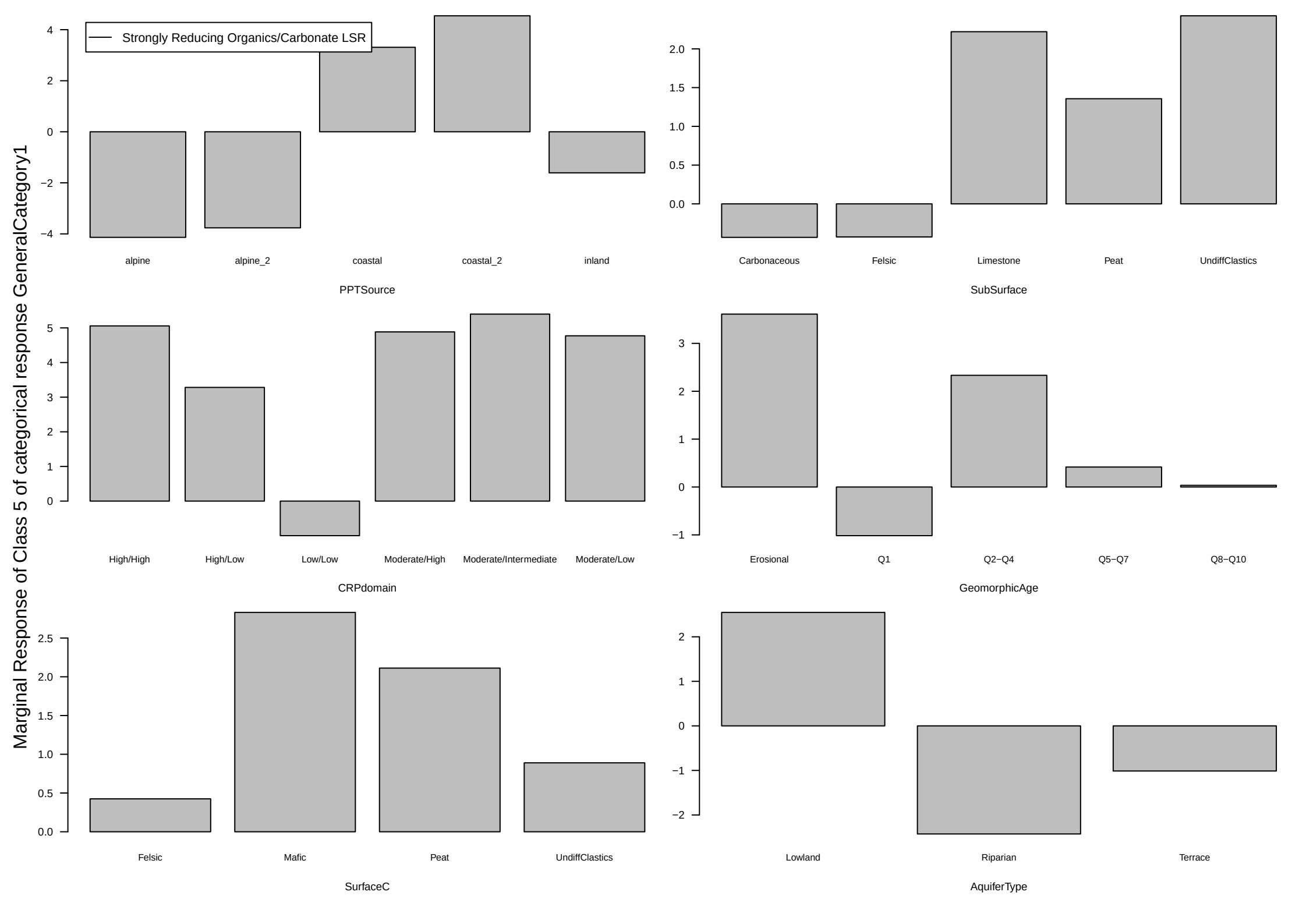

Marginal Response of Class 6 of categorical response Cluster6
6
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
SoilRP
High
Low
Moderate

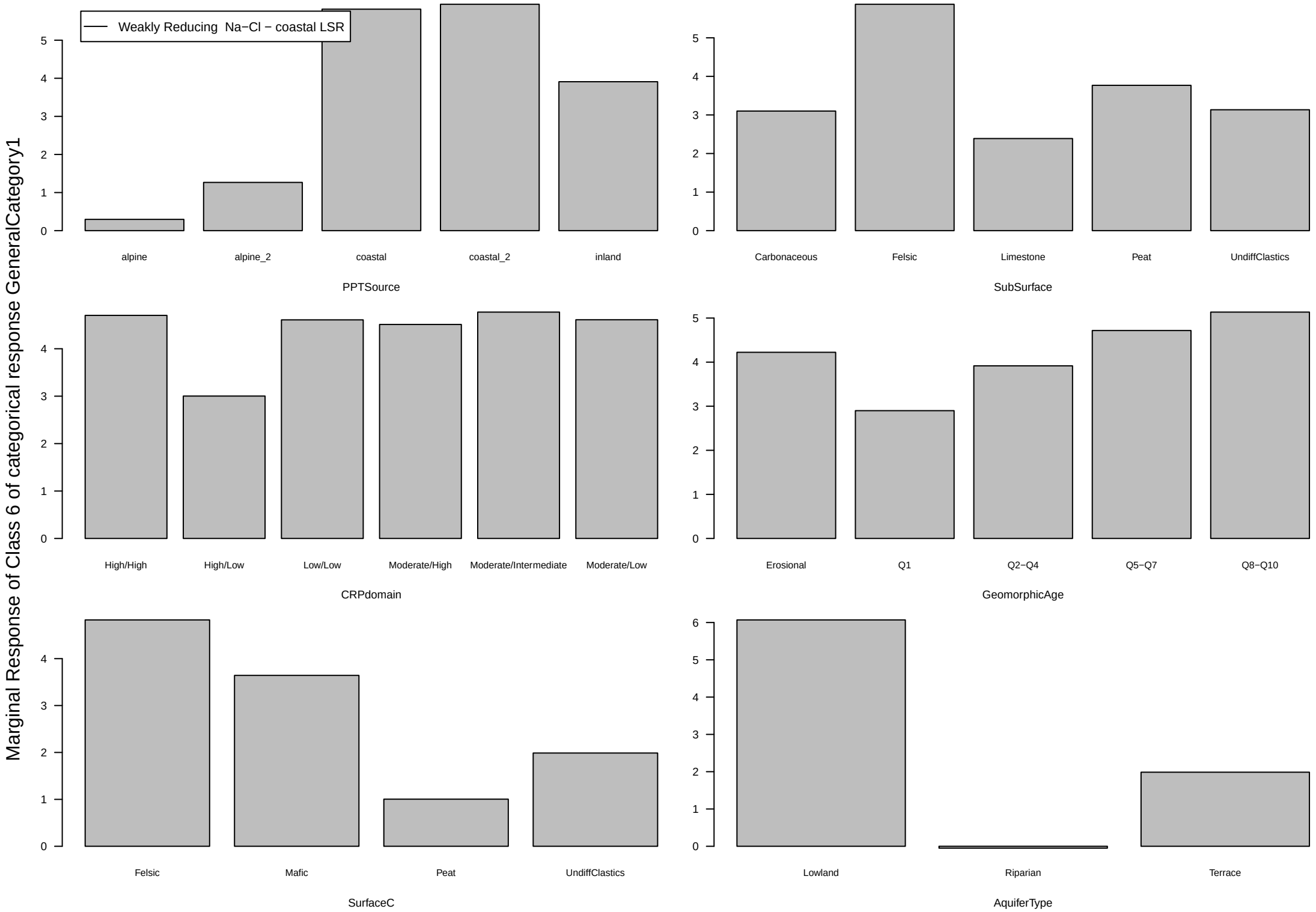

Oxidising Ca−HCO3 dominated LSR
Marginal Response of Class 1 of categorical response GeneralCategory1
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
AquiferType
Lowland
Riparian
Terrace

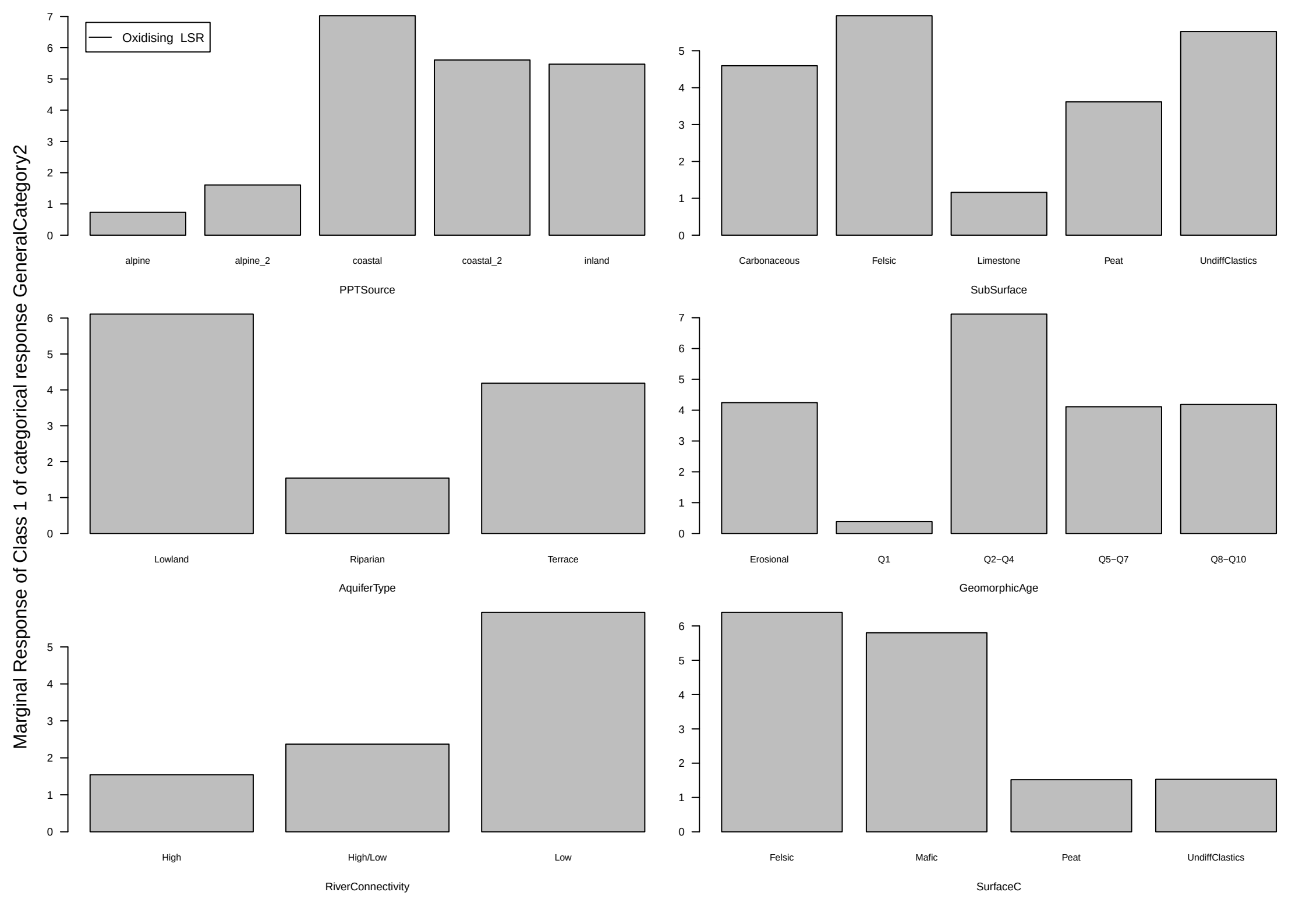

Oxidising Dilute, −ve18O River Influenced
Marginal Response of Class 2 of categorical response GeneralCategory1
0
−5
−10
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
0
−2
−4
−6
−8
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
0
−2
−4
−6
−8
−10
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
2
0
−2
−4
−6
−8
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
0
−2
−4
−6
−8
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
0
−2
−4
−6
−8
−10
Lowland
Riparian
Terrace
AquiferType

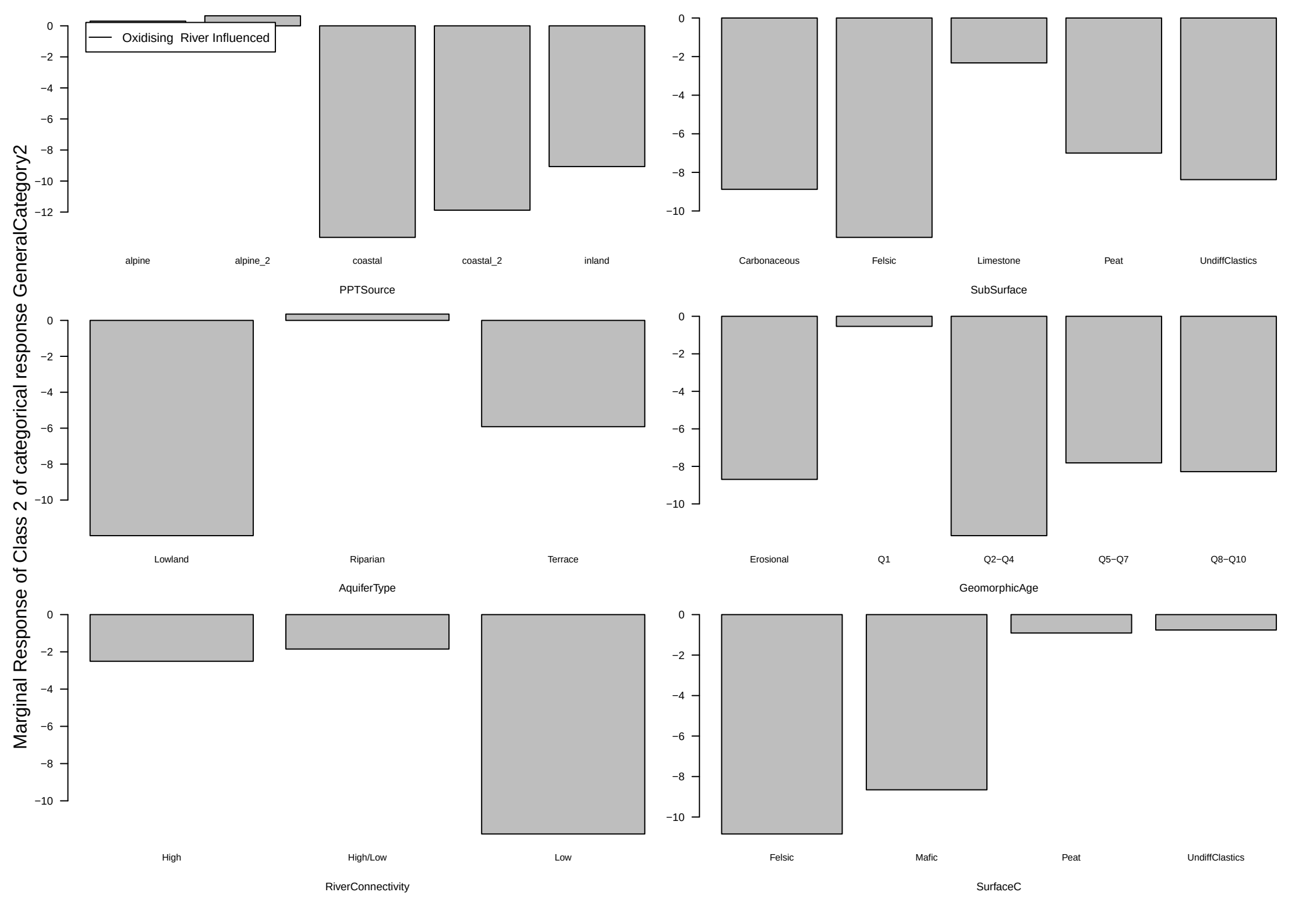

Strongly Oxidising Bypass LSR
Marginal Response of Class 3 of categorical response GeneralCategory1
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
Lowland
Riparian
Terrace
AquiferType

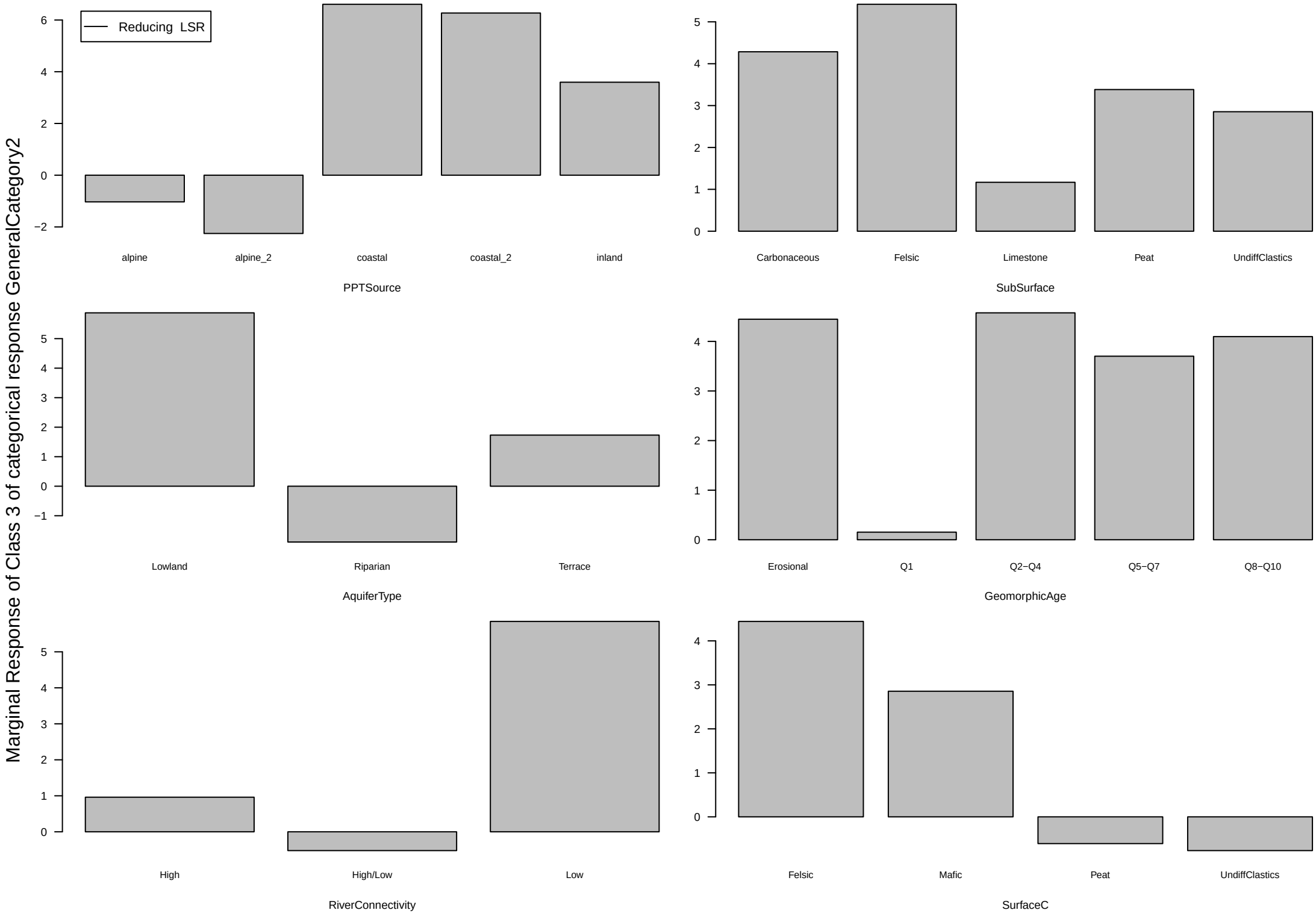

Strongly Oxidising Low SO4 old surface LSR
Marginal Response of Class 4 of categorical response GeneralCategory1
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
AquiferType
Lowland
Riparian
Terrace

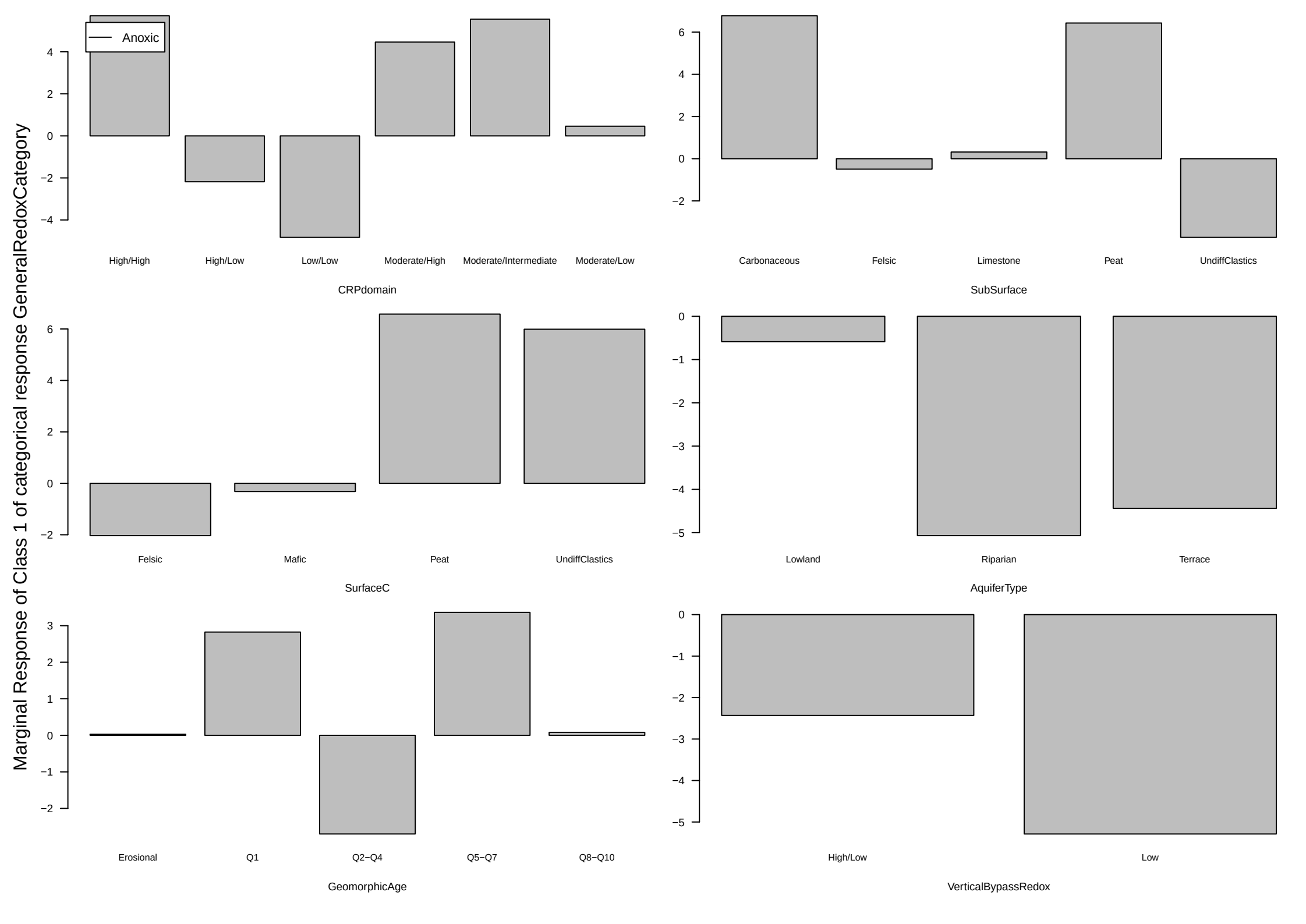

Strongly Reducing Organics/Carbonate LSR
Marginal Response of Class 5 of categorical response GeneralCategory1
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
Lowland
Riparian
Terrace
AquiferType

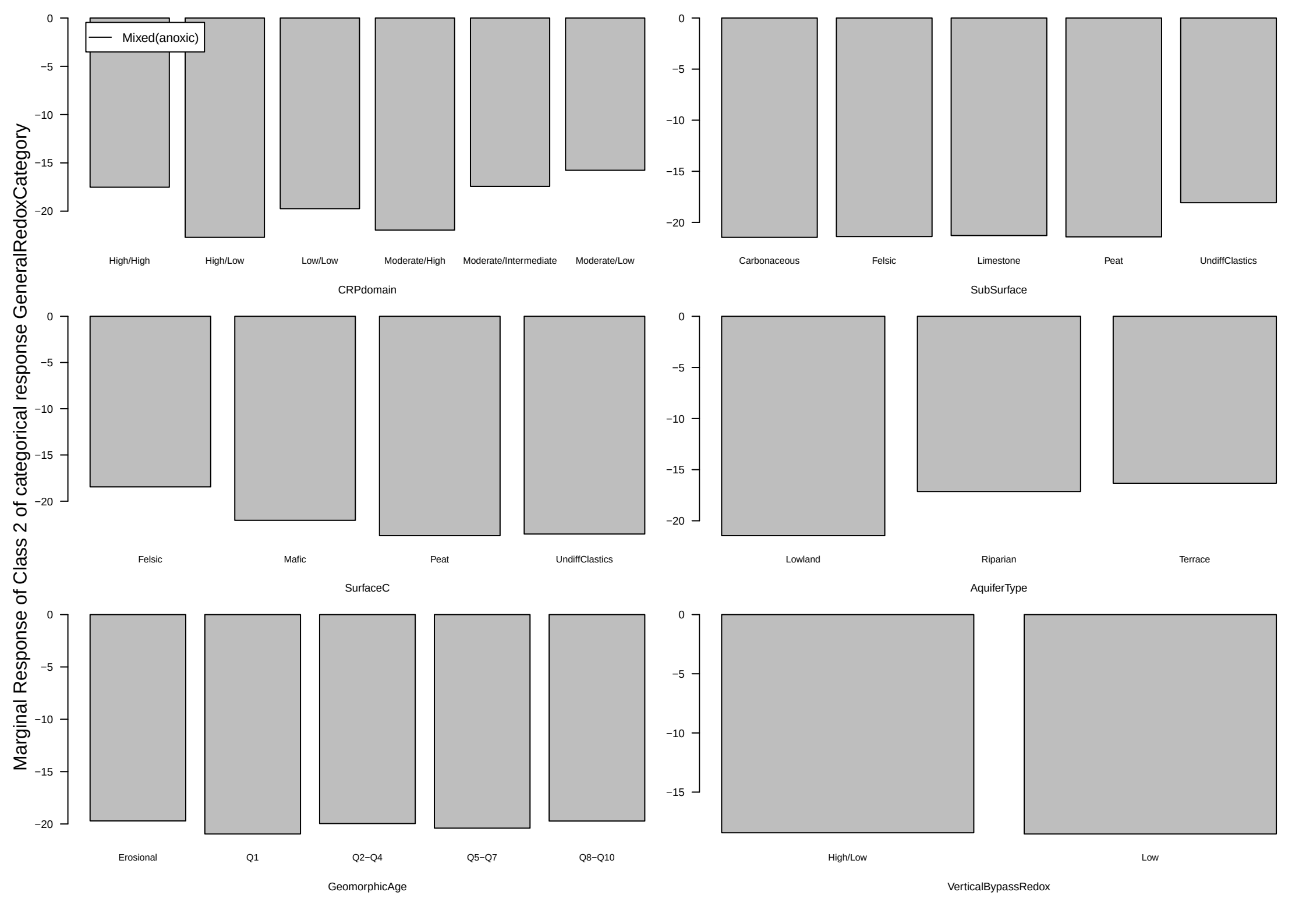

Weakly Reducing Na−Cl − coastal LSR
Marginal Response of Class 6 of categorical response GeneralCategory1
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
AquiferType
Lowland
Riparian
Terrace

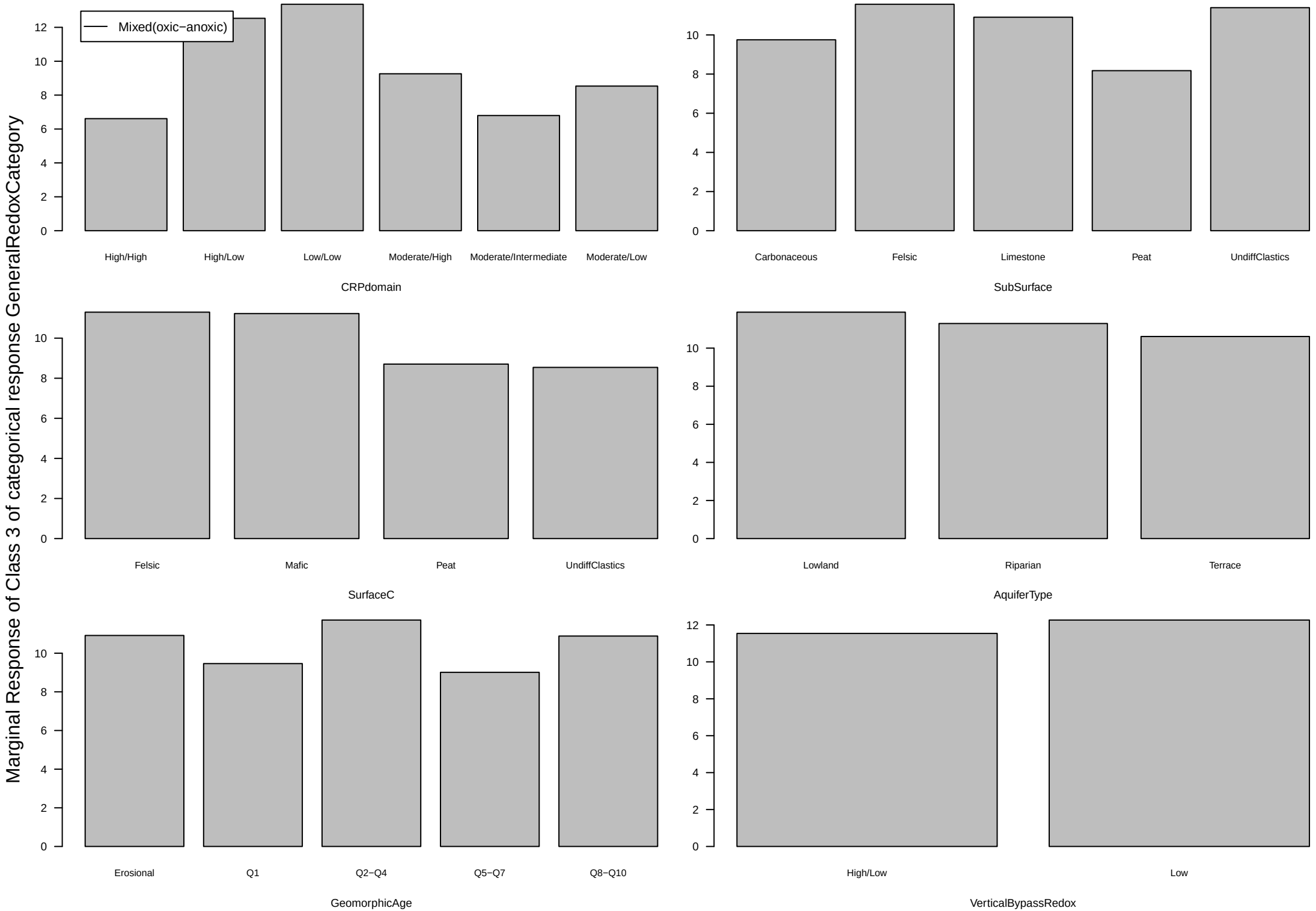

Marginal Response of Class 1 of categorical response GeneralCategory2
Oxidising LSR
PPTSource
alpine
alpine_2
coastal
coastal_2
inland
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
AquiferType
Lowland
Riparian
Terrace
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
RiverConnectivity
High
High/Low
Low
SurfaceC
Felsic
Mafic
Peat
UndiffClastics

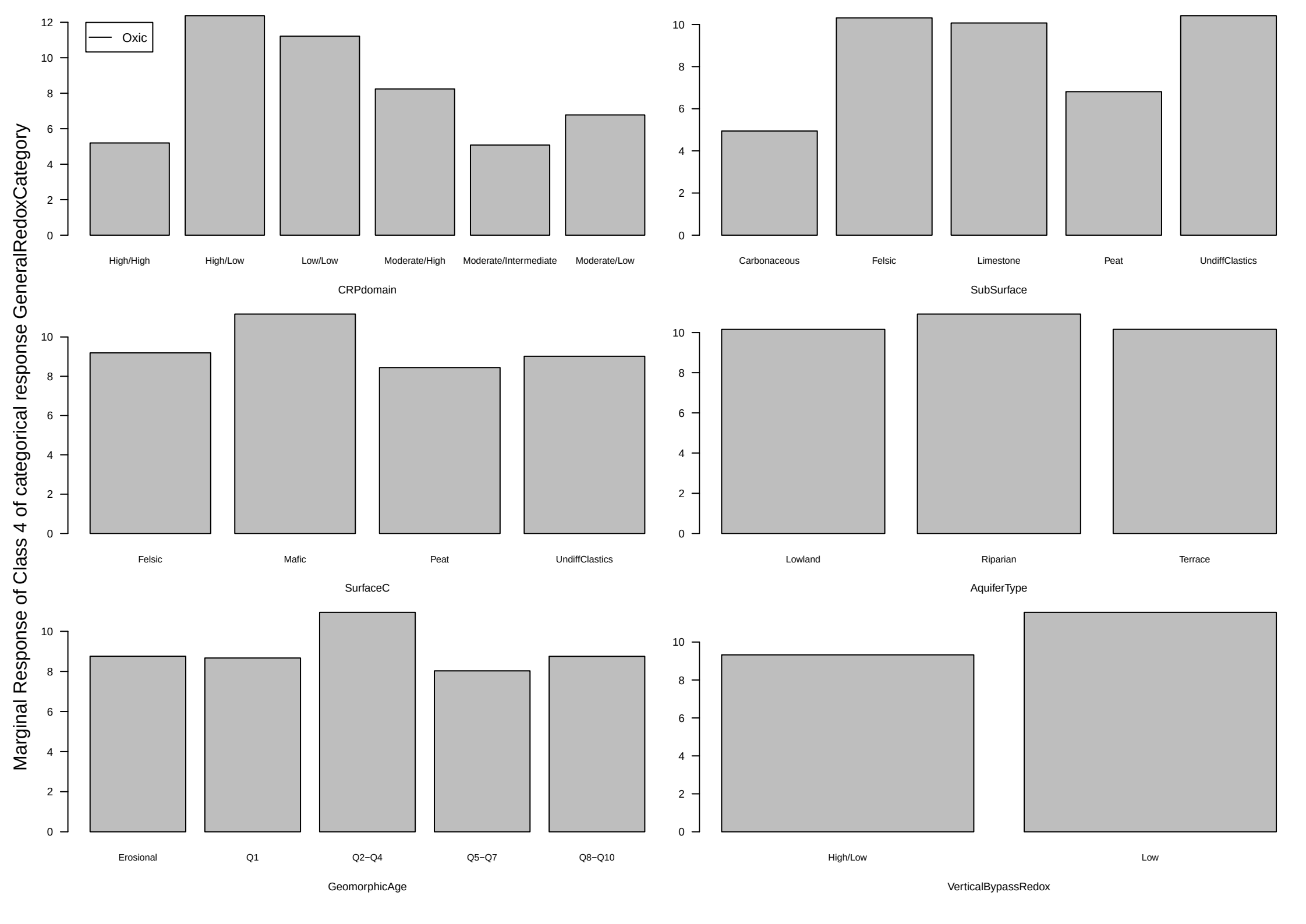

Oxidising River Influenced
Marginal Response of Class 2 of categorical response GeneralCategory2
0
−2
−4
−6
−8
−10
−12
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
Lowland
Riparian
Terrace
AquiferType
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
High
High/Low
Low
RiverConnectivity
Felsic
Mafic
Peat
UndiffClastics
SurfaceC

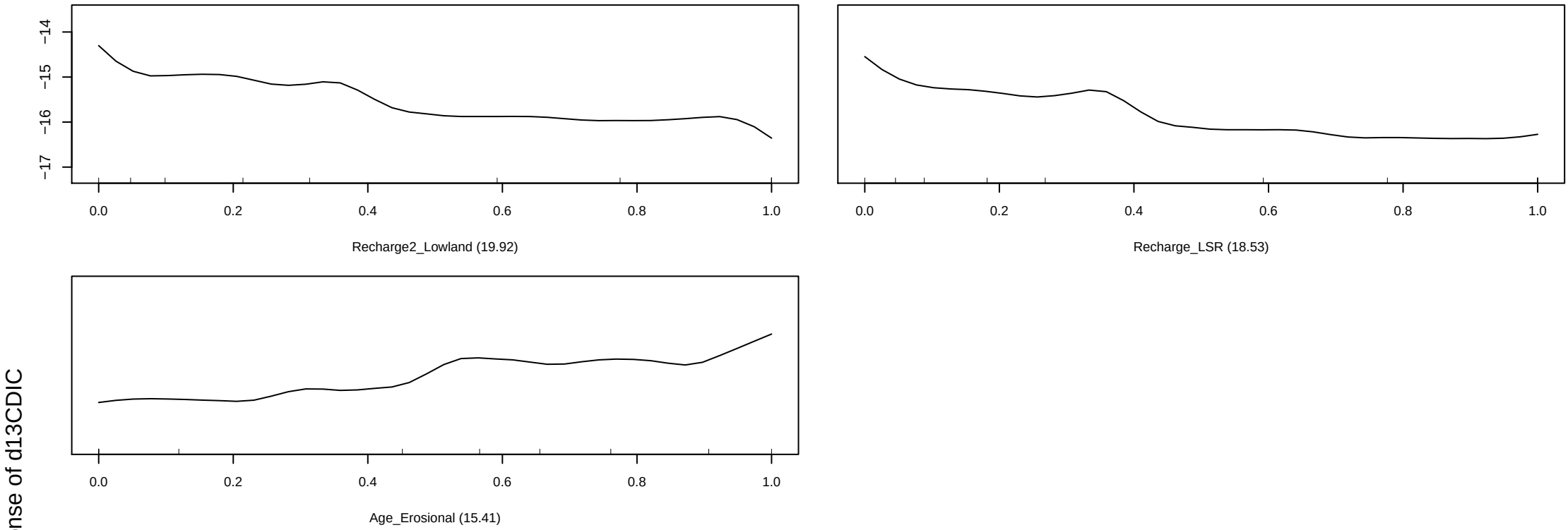

Marginal Response of Class 3 of categorical response GeneralCategory2
Reducing LSR
6
4
2
0
−2
alpine
alpine_2
coastal
coastal_2
inland
PPTSource
5
4
3
2
1
0
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
5
4
3
2
1
0
−1
Lowland
Riparian
Terrace
AquiferType
4
3
2
1
0
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
5
4
3
2
1
0
High
High/Low
Low
RiverConnectivity
4
3
2
1
0
Felsic
Mafic
Peat
UndiffClastics
SurfaceC

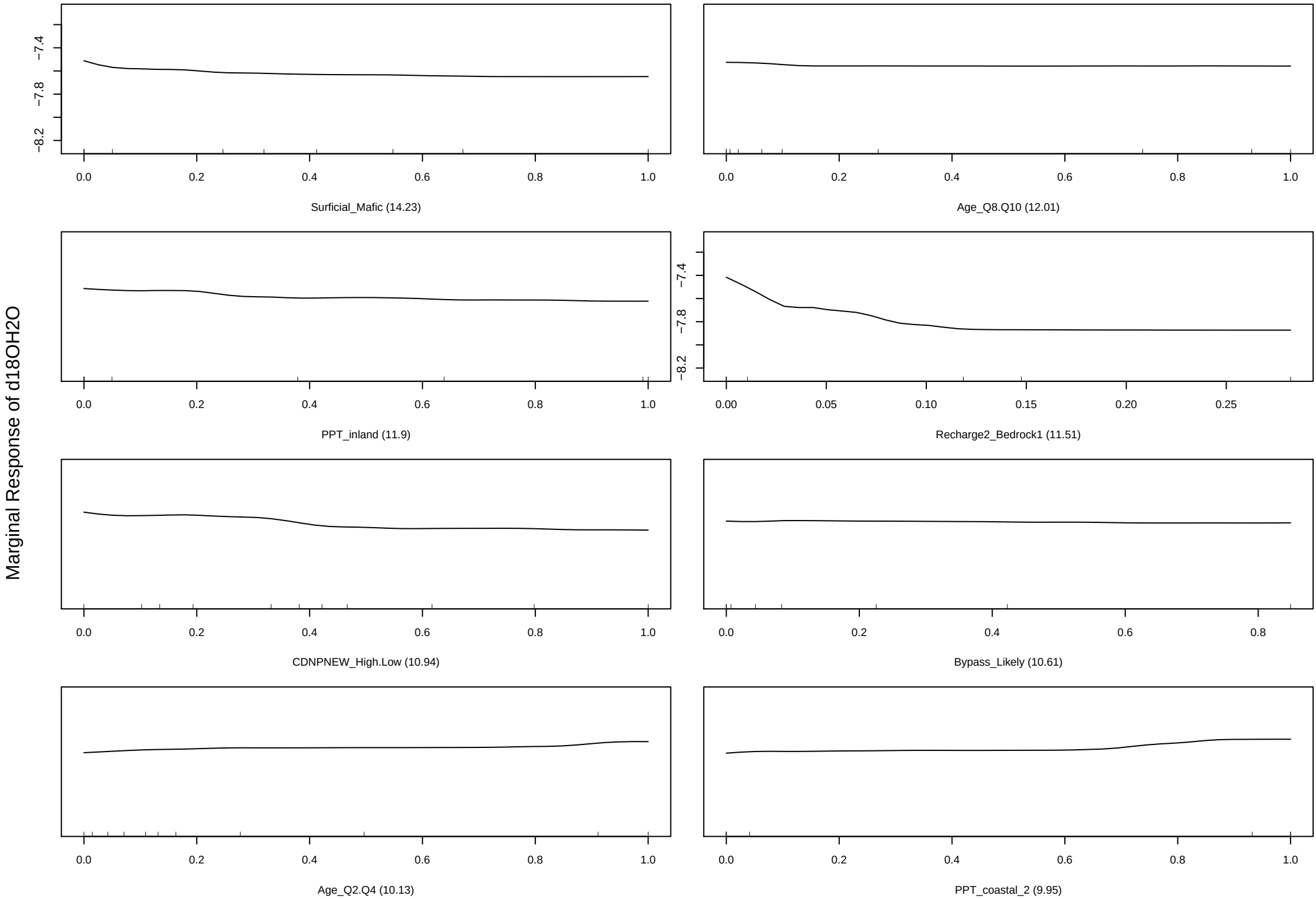

Anoxic
Marginal Response of Class 1 of categorical response GeneralRedoxCategory
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
Lowland
Riparian
Terrace
AquiferType
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
High/Low
Low
VerticalBypassRedox

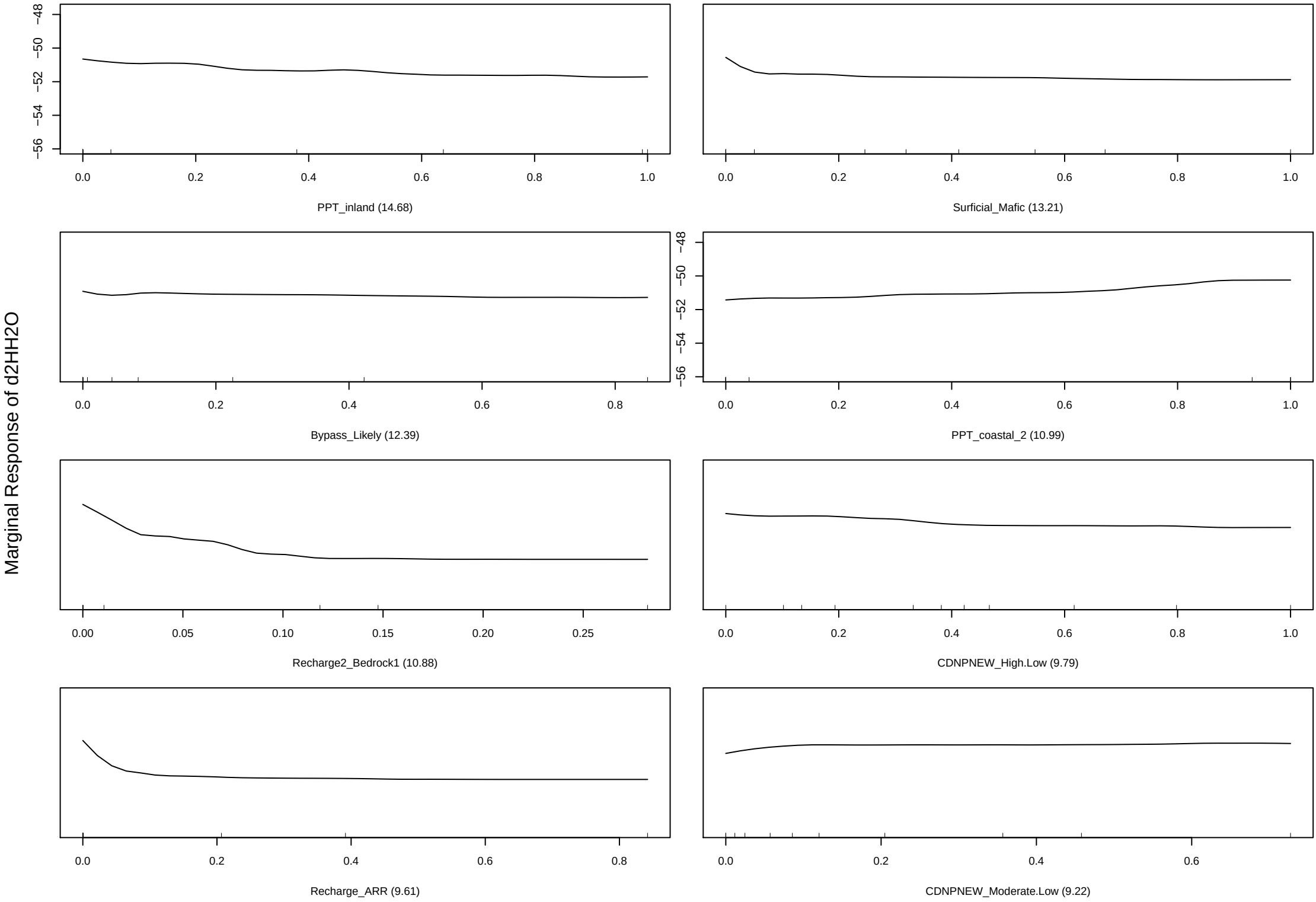

Marginal Response of Class 2 of categorical response GeneralRedoxCategory
Mixed(anoxic)
0
−5
−10
−15
−20
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
Lowland
Riparian
Terrace
AquiferType
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
High/Low
Low
VerticalBypassRedox

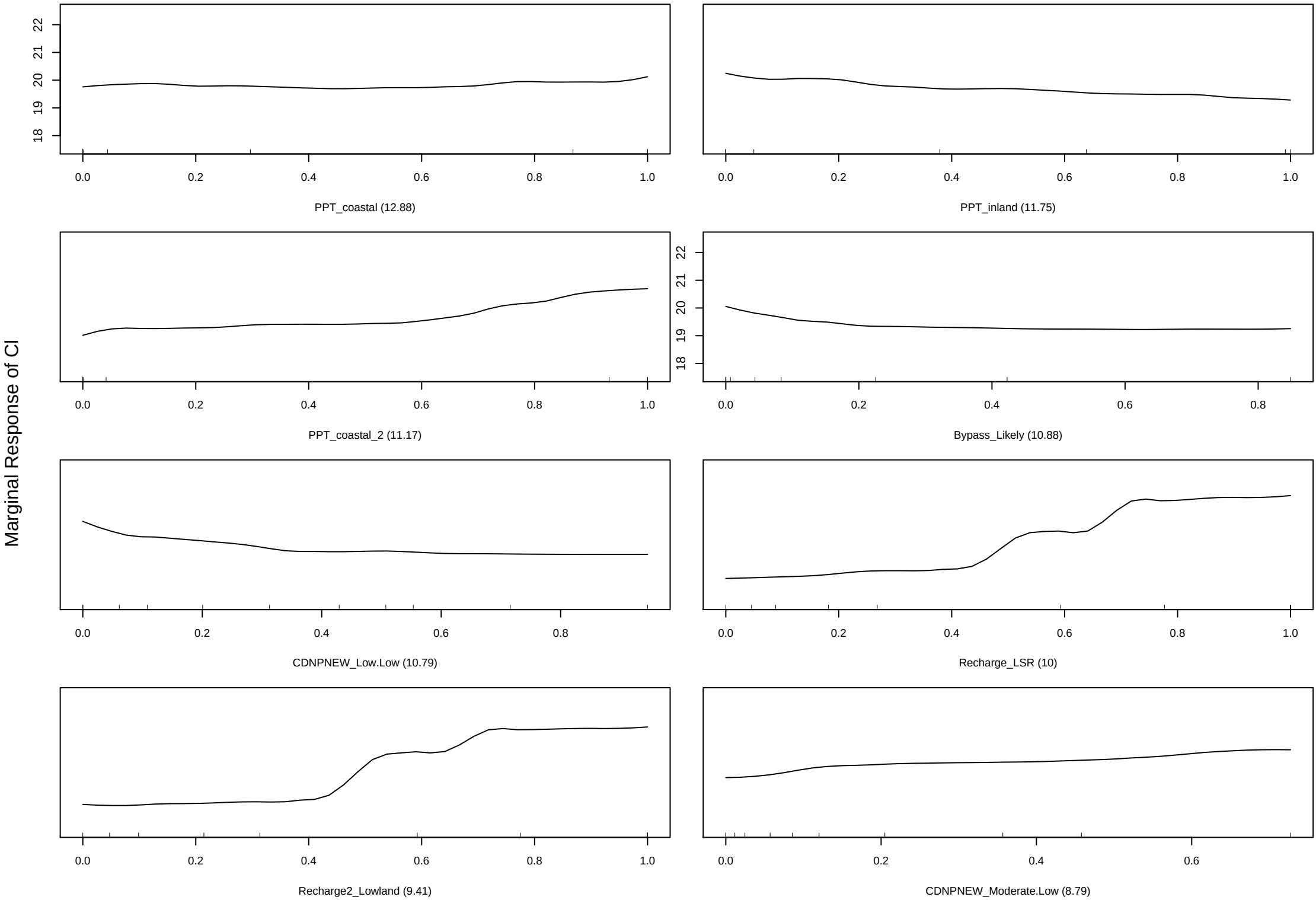

Mixed(oxic−anoxic)
Marginal Response of Class 3 of categorical response GeneralRedoxCategory
CRPdomain
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
SubSurface
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SurfaceC
Felsic
Mafic
Peat
UndiffClastics
AquiferType
Lowland
Riparian
Terrace
GeomorphicAge
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
VerticalBypassRedox
High/Low
Low

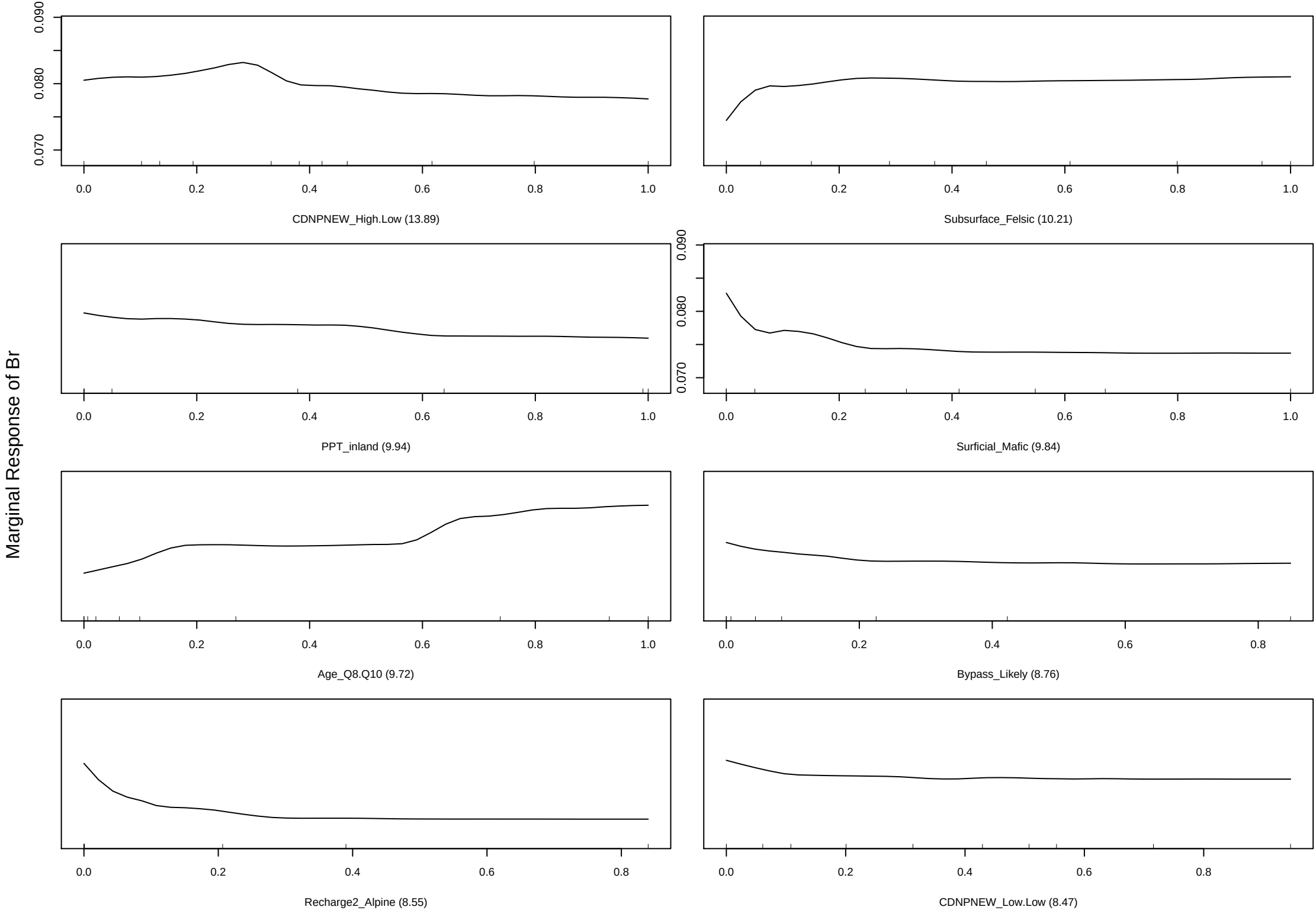

Marginal Response of Class 4 of categorical response GeneralRedoxCategory
Oxic
0
2
4
6
8
10
12
High/High
High/Low
Low/Low
Moderate/High
Moderate/Intermediate
Moderate/Low
CRPdomain
Carbonaceous
Felsic
Limestone
Peat
UndiffClastics
SubSurface
Felsic
Mafic
Peat
UndiffClastics
SurfaceC
Lowland
Riparian
Terrace
AquiferType
Erosional
Q1
Q2−Q4
Q5−Q7
Q8−Q10
GeomorphicAge
High/Low
Low
VerticalBypassRedox

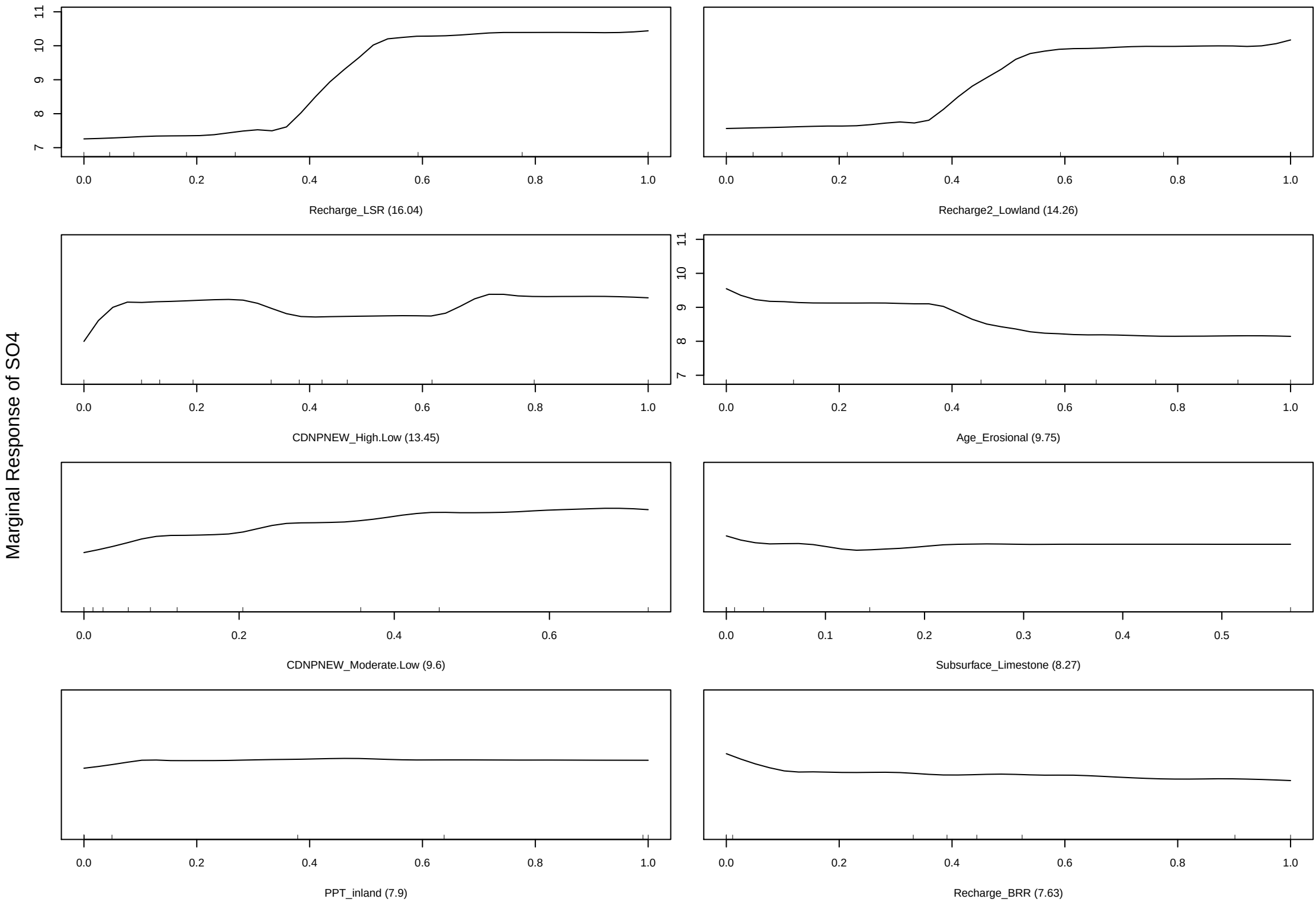

Marginal Response of d13CDIC
−14
−15
−16
−17
0.0
0.2
0.4
0.6
0.8
1.0
Recharge2_Lowland (19.92)
Recharge_LSR (18.53)
Age_Erosional (15.41)

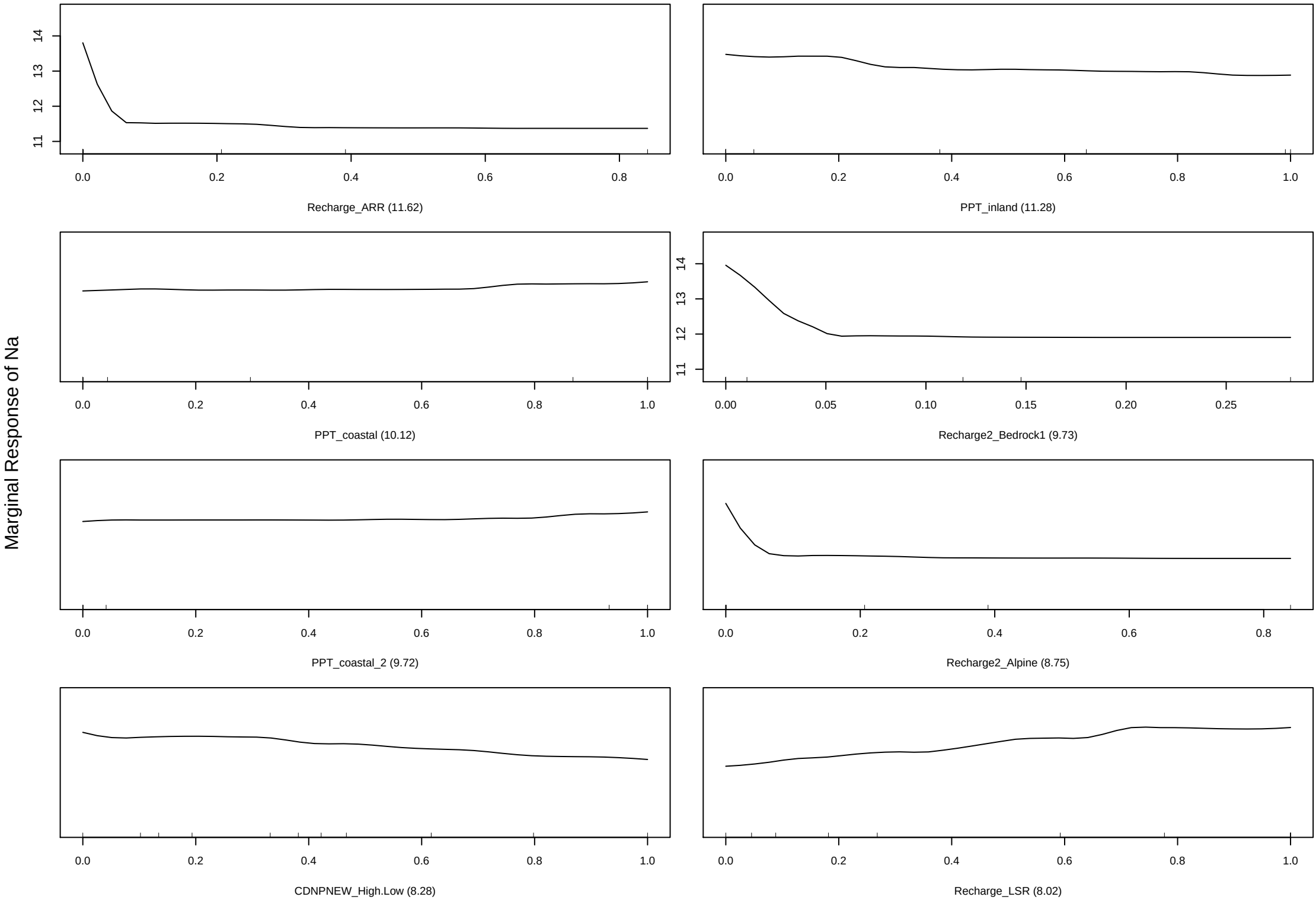

Marginal Response of d18OH2O
−7.4
−7.8
−8.2
0.0
0.2
0.4
0.6
0.8
1.0
Surficial_Mafic (14.23)
Age_Q8.Q10 (12.01)
PPT_inland (11.9)
0.00
0.05
0.10
0.15
0.20
0.25
Recharge2_Bedrock1 (11.51)
CDNPNEW_High.Low (10.94)
Bypass_Likely (10.61)
Age_Q2.Q4 (10.13)
PPT_coastal_2 (9.95)

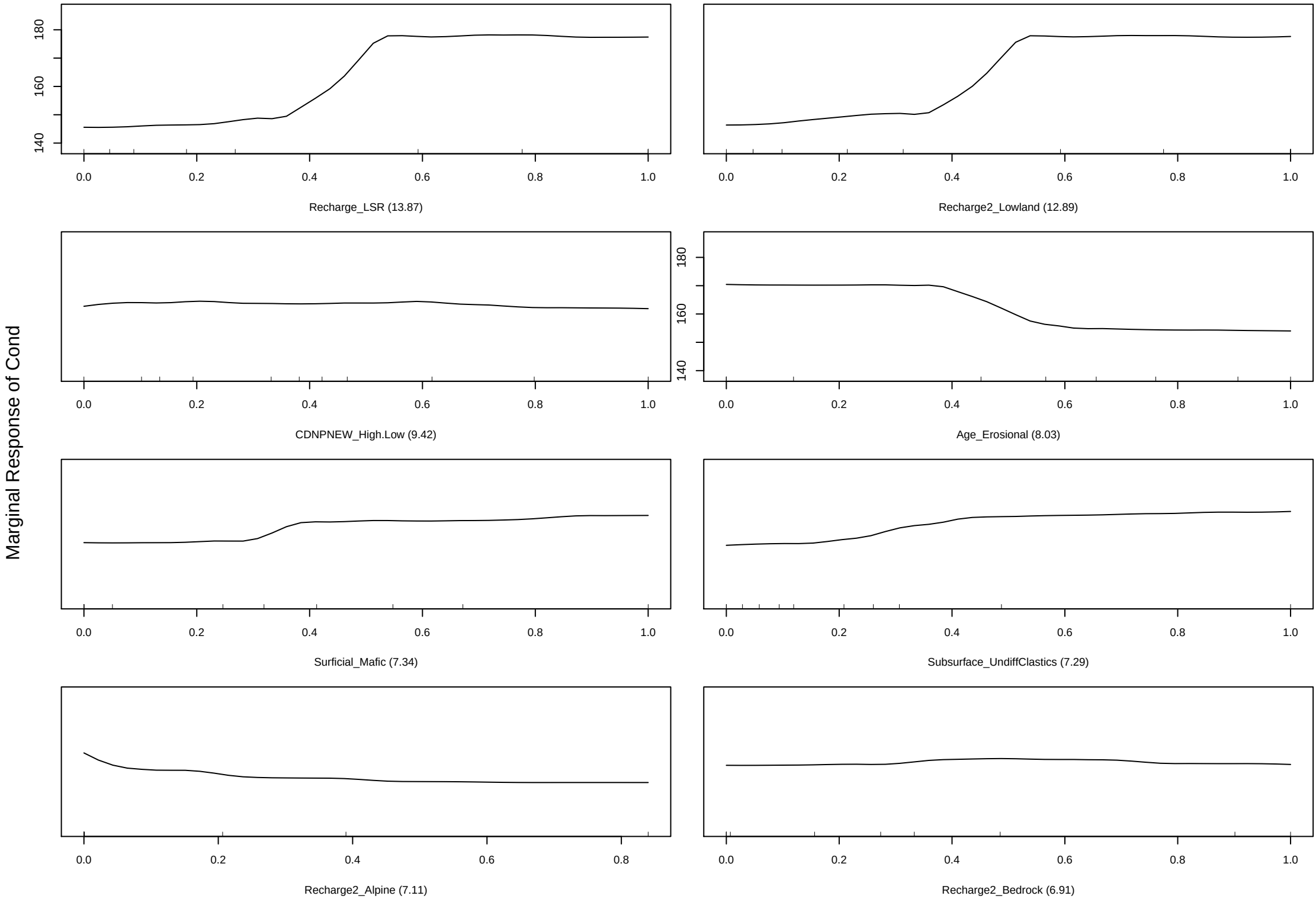

Marginal Response of d2HH2O
−48
−50
−52
−54
−56
0.0
0.2
0.4
0.6
0.8
1.0
PPT_inland (14.68)
Surficial_Mafic (13.21)
Bypass_Likely (12.39)
PPT_coastal_2 (10.99)
0.00
0.05
0.10
0.15
0.20
0.25
Recharge2_Bedrock1 (10.88)
CDNPNEW_High.Low (9.79)
Recharge_ARR (9.61)
CDNPNEW_Moderate.Low (9.22)

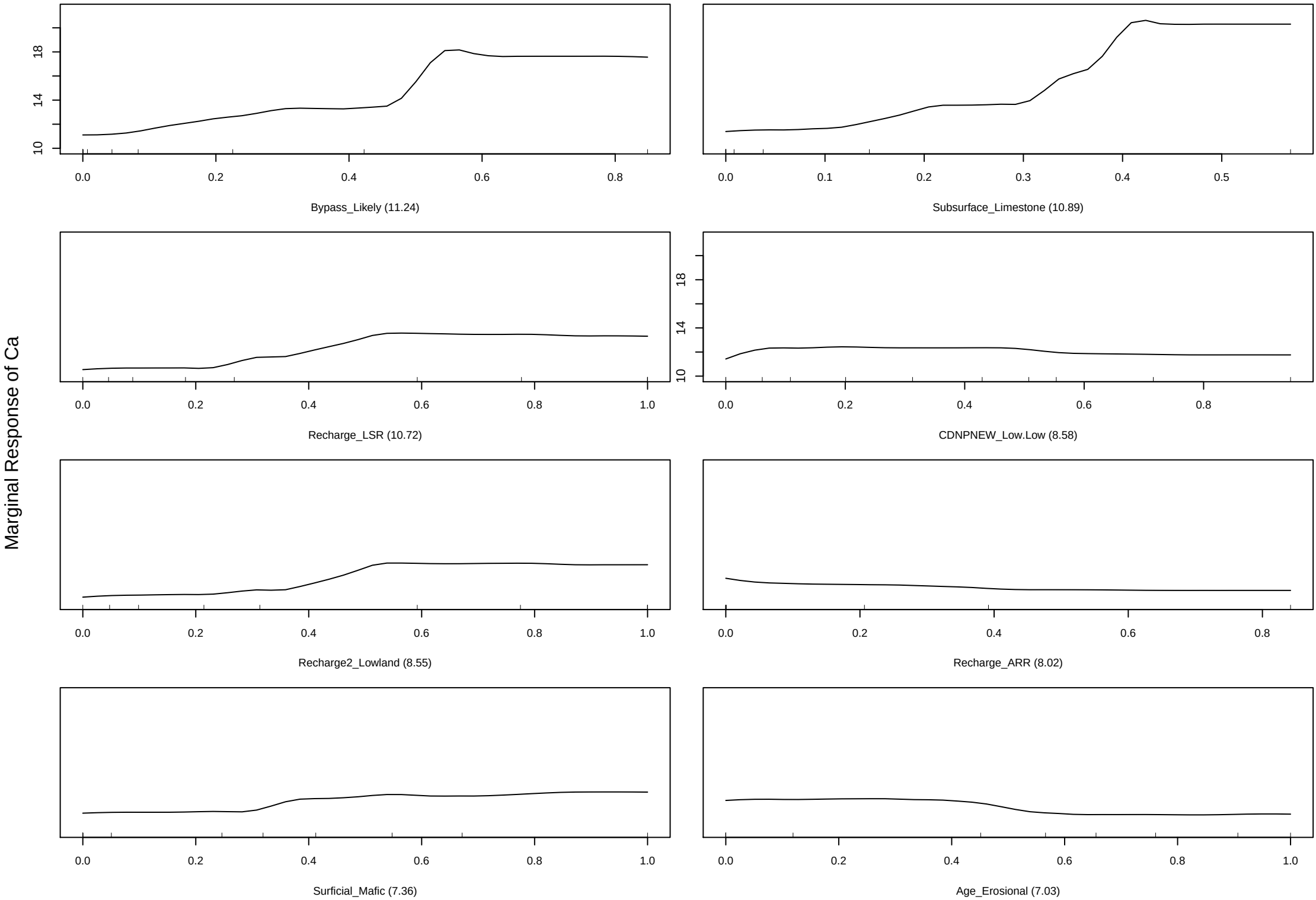

Marginal Response of CI
PPT_coastal (12.88)
PPT_inland (11.75)
PPT_coastal_2 (11.17)
Bypass_Likely (10.88)
CDNPNEW_Low.Low (10.79)
Recharge_LSR (10)
Recharge2_Lowland (9.41)
CDNPNEW_Moderate.Low (8.79)

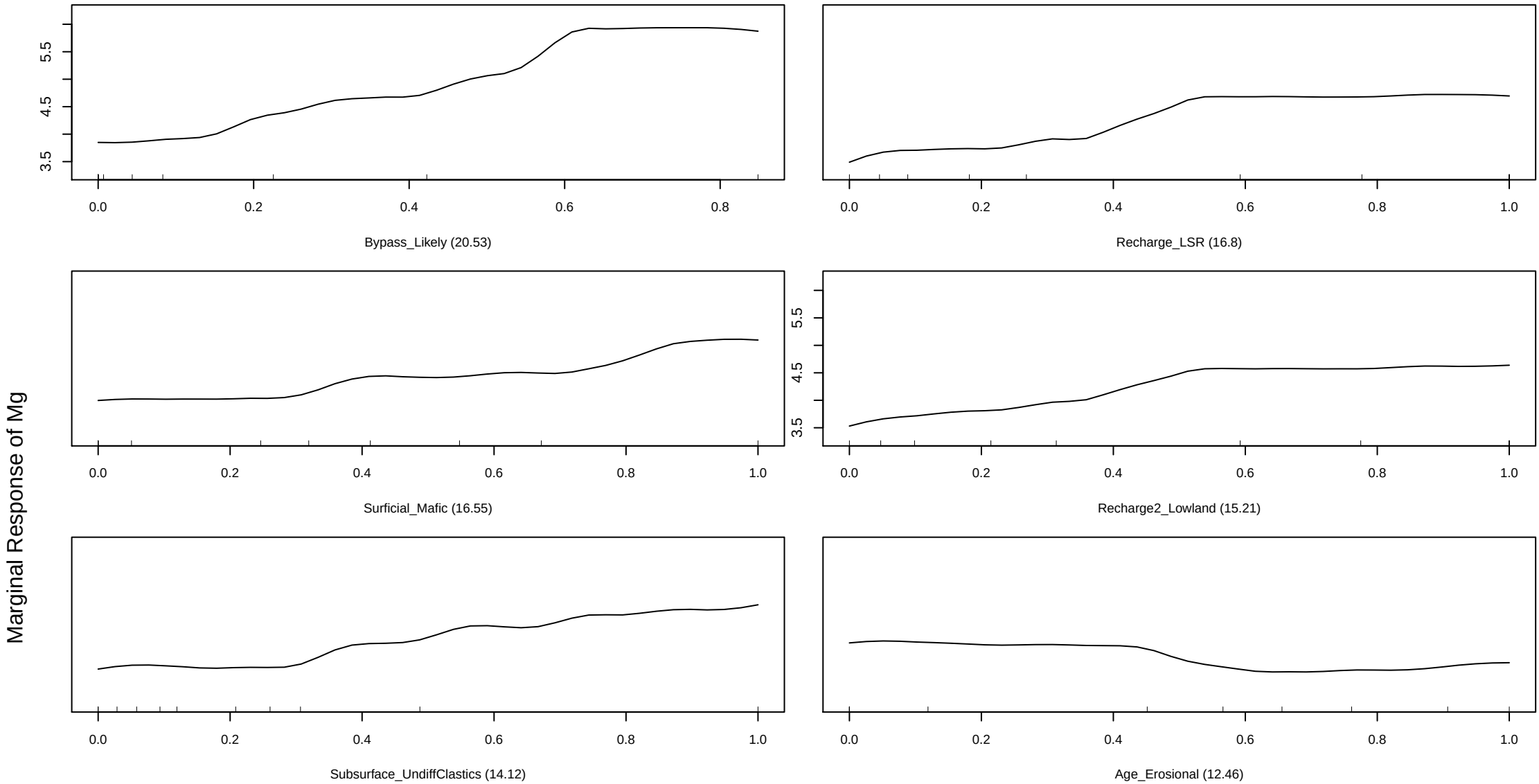

Marginal Response of Br
0.070
0.080
0.090
0.0
0.2
0.4
0.6
0.8
1.0
CDNPNEW_High.Low (13.89)
Subsurface_Felsic (10.21)
PPT_inland (9.94)
Surficial_Mafic (9.84)
Age_Q8.Q10 (9.72)
Bypass_Likely (8.76)
Recharge2_Alpine (8.55)
CDNPNEW_Low.Low (8.47)

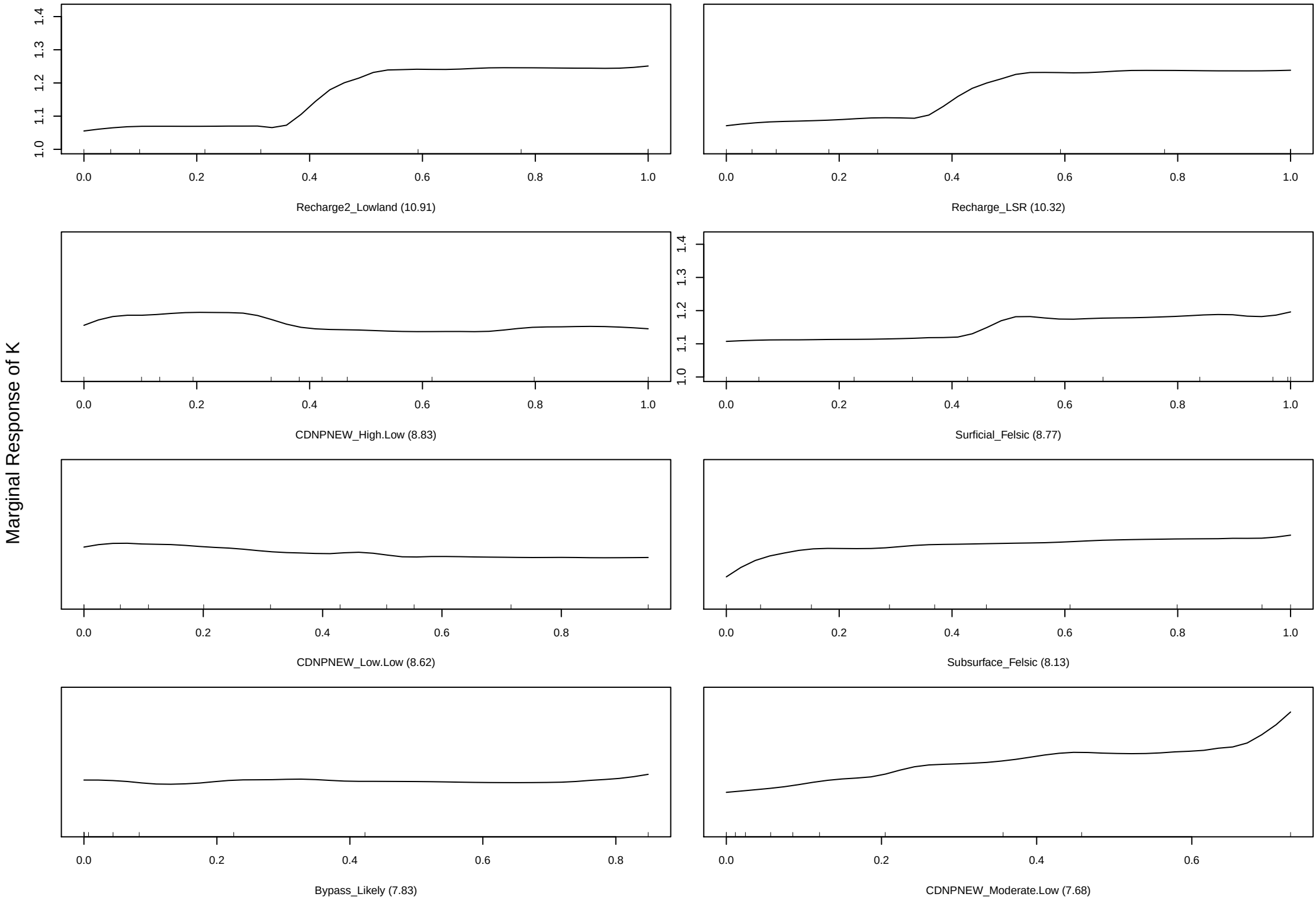

Marginal Response of SO4
Recharge_LSR (16.04)
Recharge2_Lowland (14.26)
CDNPNEW_High.Low (13.45)
Age_Erosional (9.75)
CDNPNEW_Moderate.Low (9.6)
Subsurface_Limestone (8.27)
PPT_inland (7.9)
Recharge_BRR (7.63)

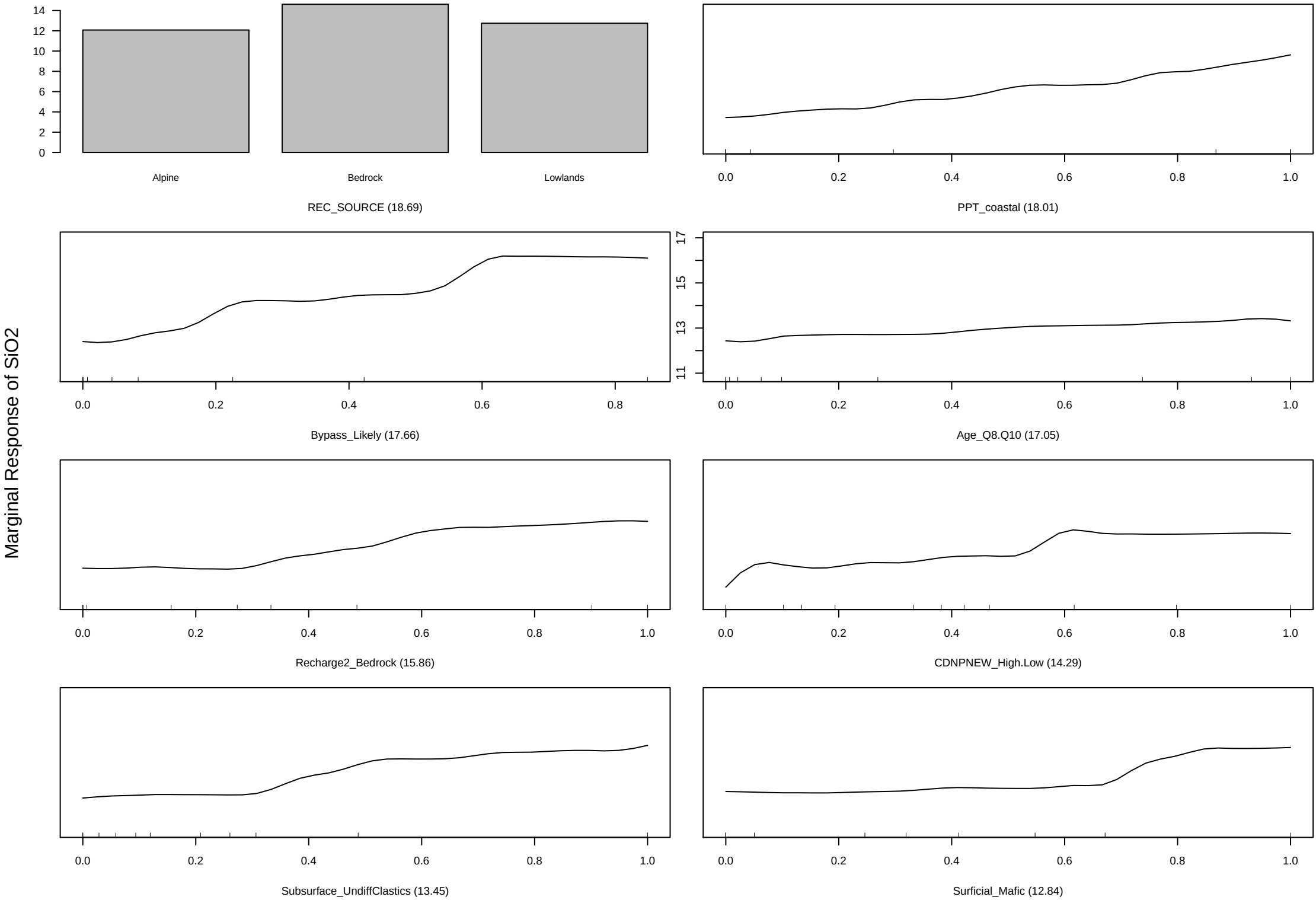

Marginal Response of Na
Recharge_ARR (11.62)
PPT_inland (11.28)
PPT_coastal (10.12)
Recharge2_Bedrock1 (9.73)
PPT_coastal_2 (9.72)
Recharge2_Alpine (8.75)
CDNPNEW_High.Low (8.28)
Recharge_LSR (8.02)

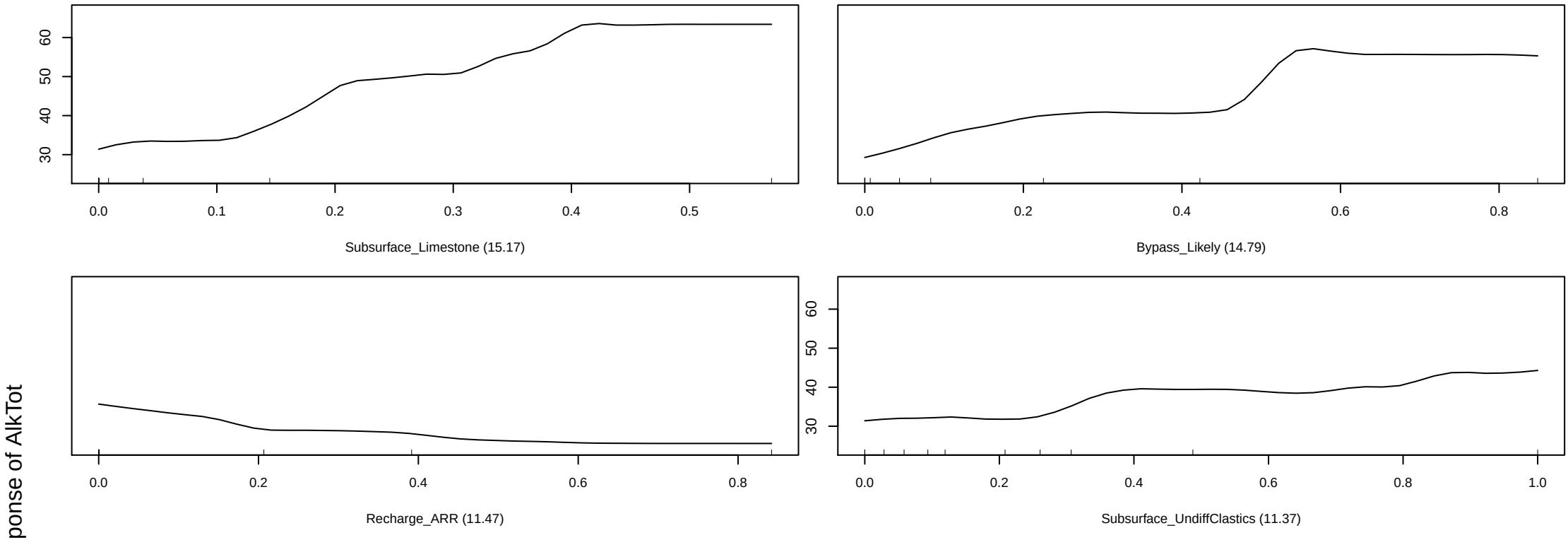

Marginal Response of Cond
140
160
180
0.0
0.2
0.4
0.6
0.8
1.0
Recharge_LSR (13.87)
Recharge2_Lowland (12.89)
CDNPNEW_High.Low (9.42)
Age_Erosional (8.03)
Surficial_Mafic (7.34)
Subsurface_UndiffClastics (7.29)
Recharge2_Alpine (7.11)
Recharge2_Bedrock (6.91)

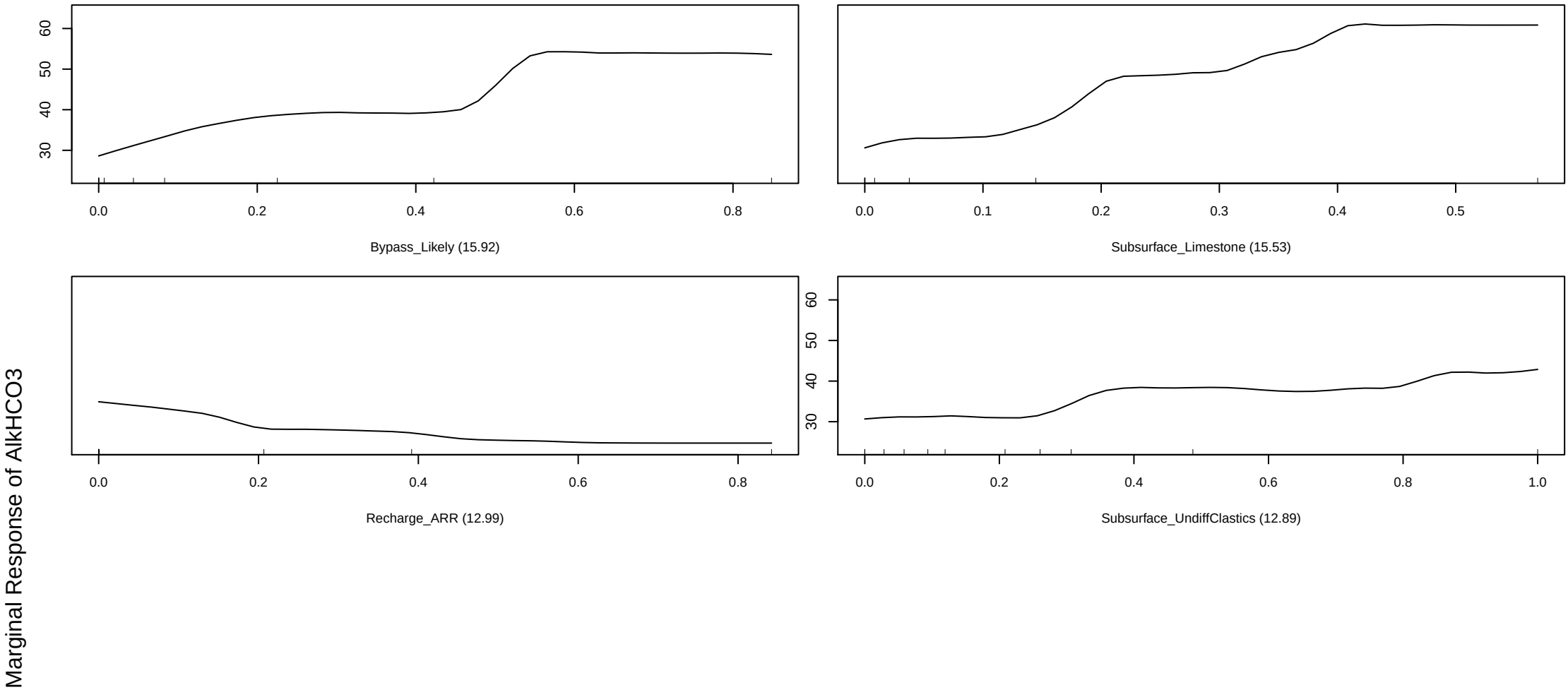

Marginal Response of Ca
10
14
18
0.0
0.2
0.4
0.6
0.8
Bypass_Likely (11.24)
0.0
0.1
0.2
0.3
0.4
0.5
Subsurface_Limestone (10.89)
0.0
0.2
0.4
0.6
0.8
1.0
Recharge_LSR (10.72)
CDNPNEW_Low.Low (8.58)
Recharge2_Lowland (8.55)
Recharge_ARR (8.02)
Surficial_Mafic (7.36)
Age_Erosional (7.03)

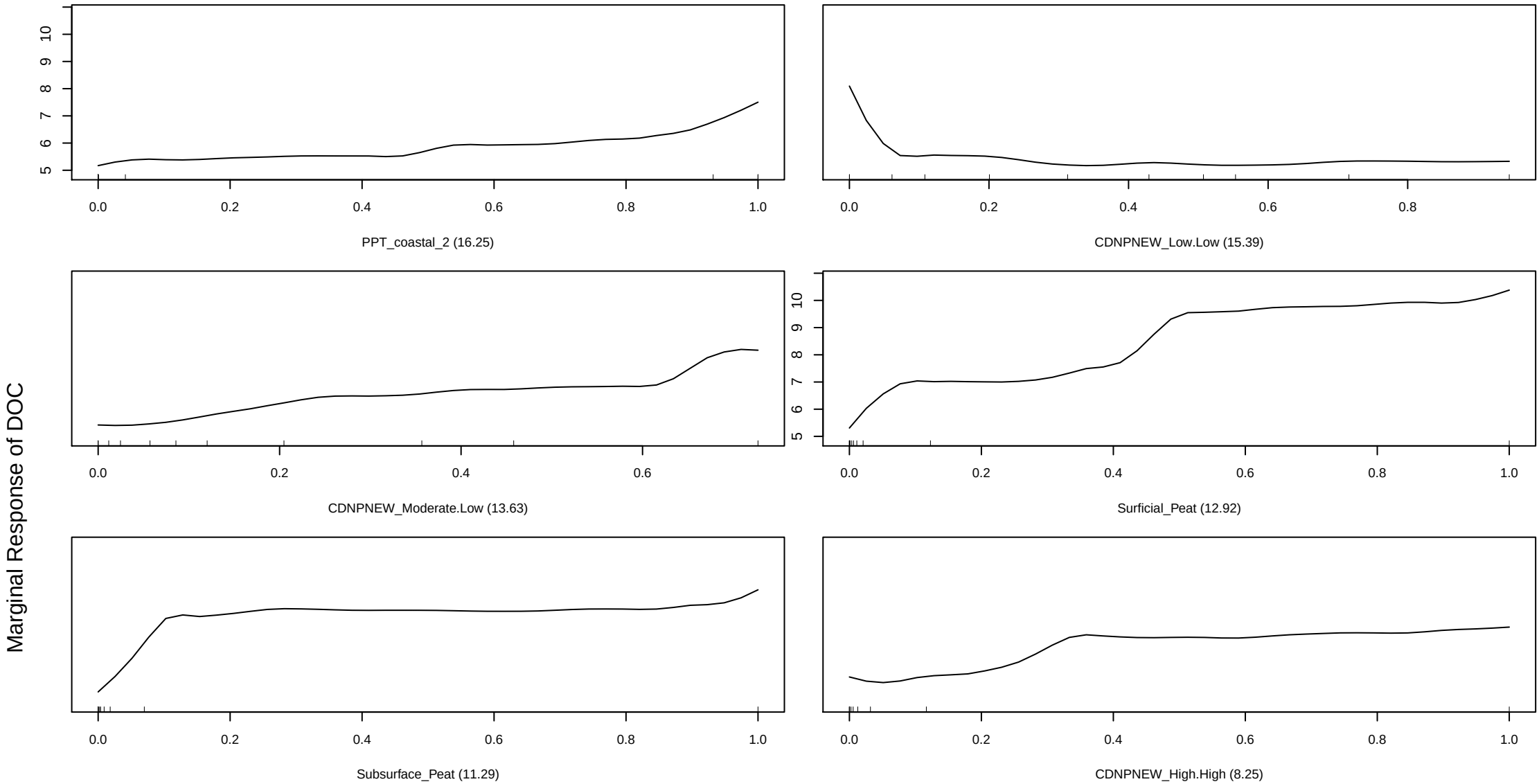

Marginal Response of Mg
3.5
4.5
5.5
0.0
0.2
0.4
0.6
0.8
1.0
Bypass_Likely (20.53)
Recharge_LSR (16.8)
Surficial_Mafic (16.55)
Recharge2_Lowland (15.21)
Subsurface_UndiffClastics (14.12)
Age_Erosional (12.46)

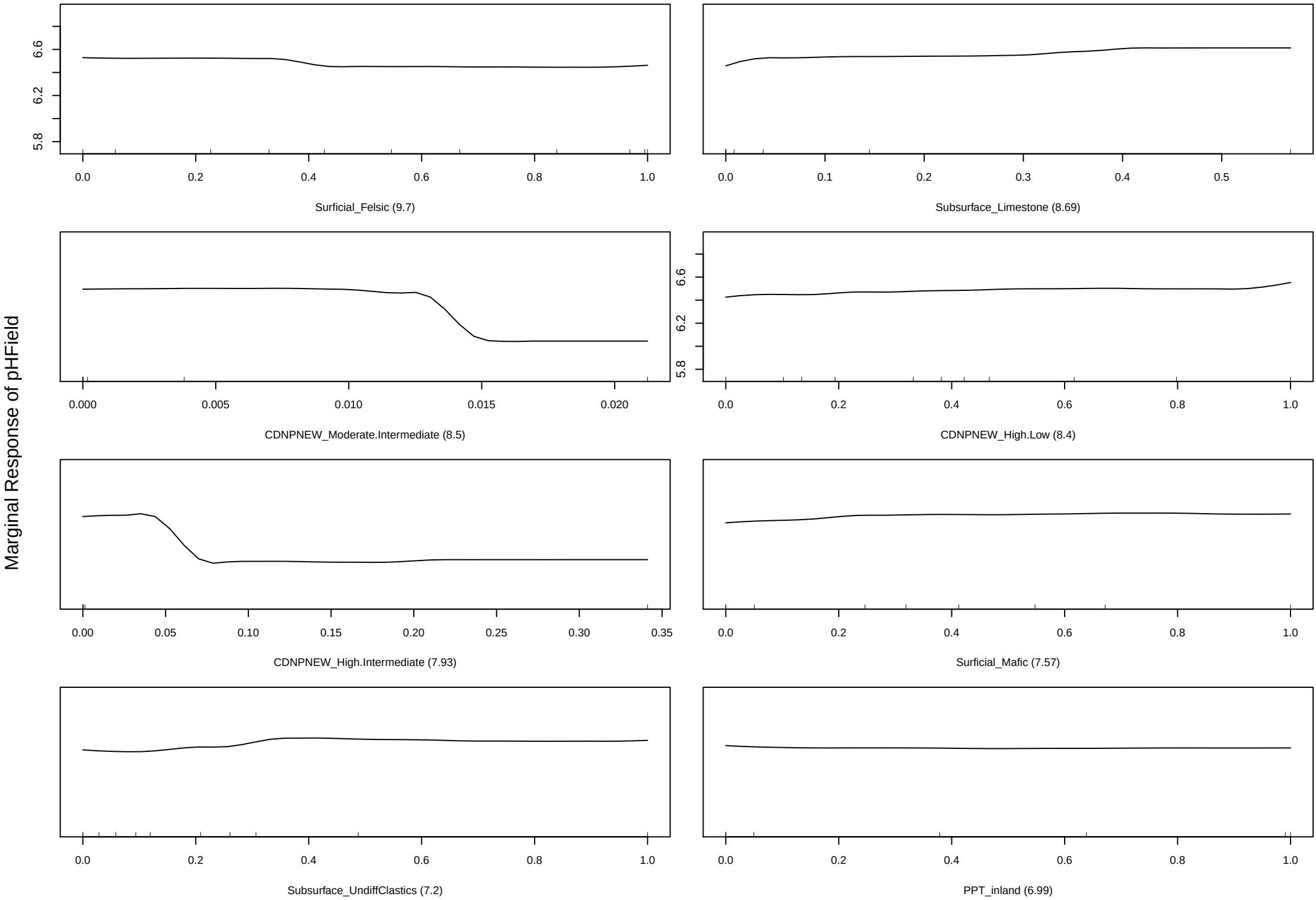

Marginal Response of K
Recharge2_Lowland (10.91)
Recharge_LSR (10.32)
CDNPNEW_High.Low (8.83)
Surficial_Felsic (8.77)
CDNPNEW_Low.Low (8.62)
Subsurface_Felsic (8.13)
Bypass_Likely (7.83)
CDNPNEW_Moderate.Low (7.68)

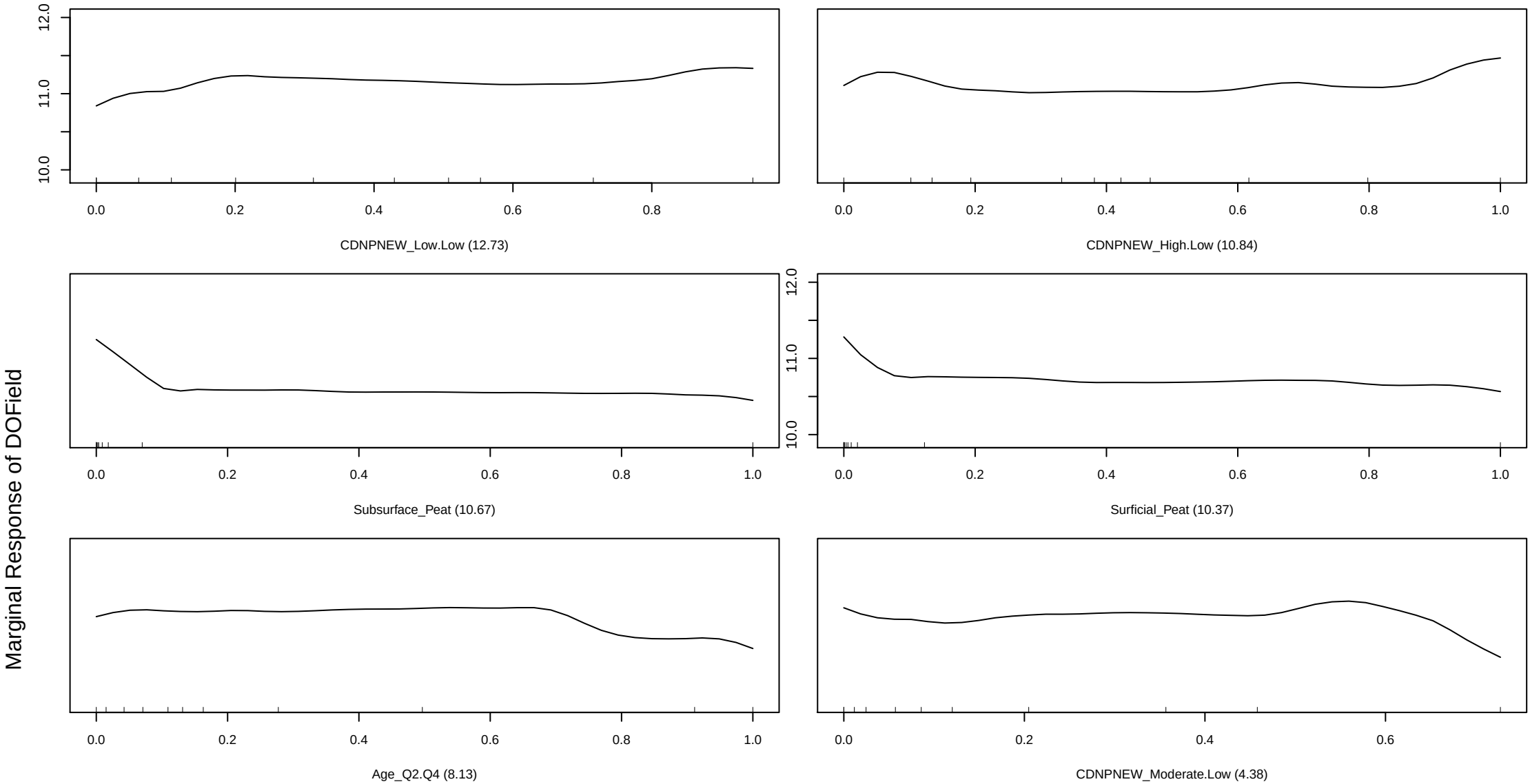

Marginal Response of SiO2
14
12
10
8
6
4
2
0
Alpine
Bedrock
Lowlands
REC_SOURCE (18.69)
PPT_coastal (18.01)
Bypass_Likely (17.66)
17
15
13
11
Age_Q8.Q10 (17.05)
Recharge2_Bedrock (15.86)
CDNPNEW_High.Low (14.29)
Subsurface_UndiffClastics (13.45)
Surficial_Mafic (12.84)
0.0
0.2
0.4
0.6
0.8
1.0

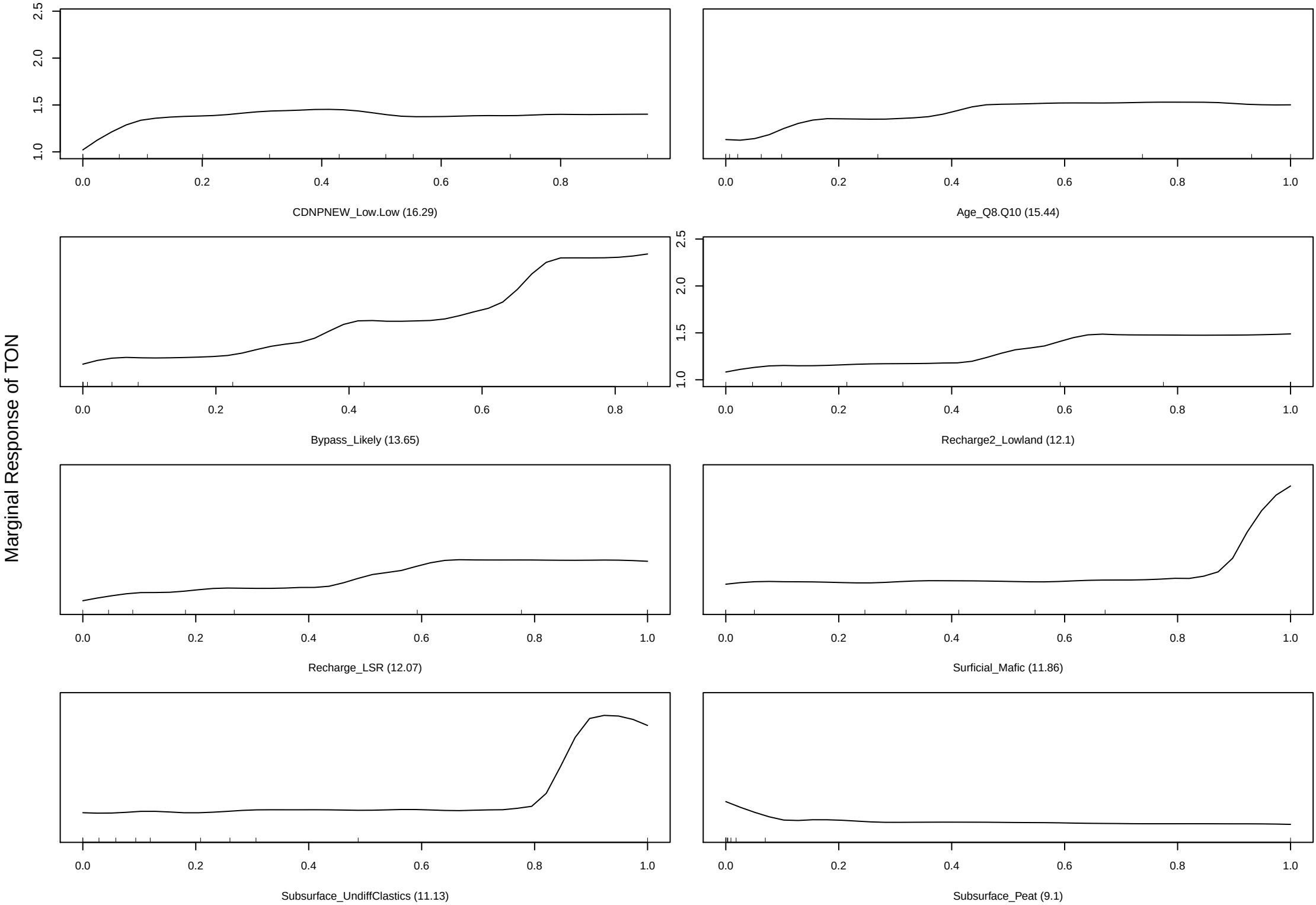

Marginal Response of AlkTot
30
40
50
60
0.0
0.1
0.2
0.3
0.4
0.5
Subsurface_Limestone (15.17)
0.0
0.2
0.4
0.6
0.8
Bypass_Likely (14.79)
0.0
0.2
0.4
0.6
0.8
Recharge_ARR (11.47)
30
40
50
60
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_UndiffClastics (11.37)

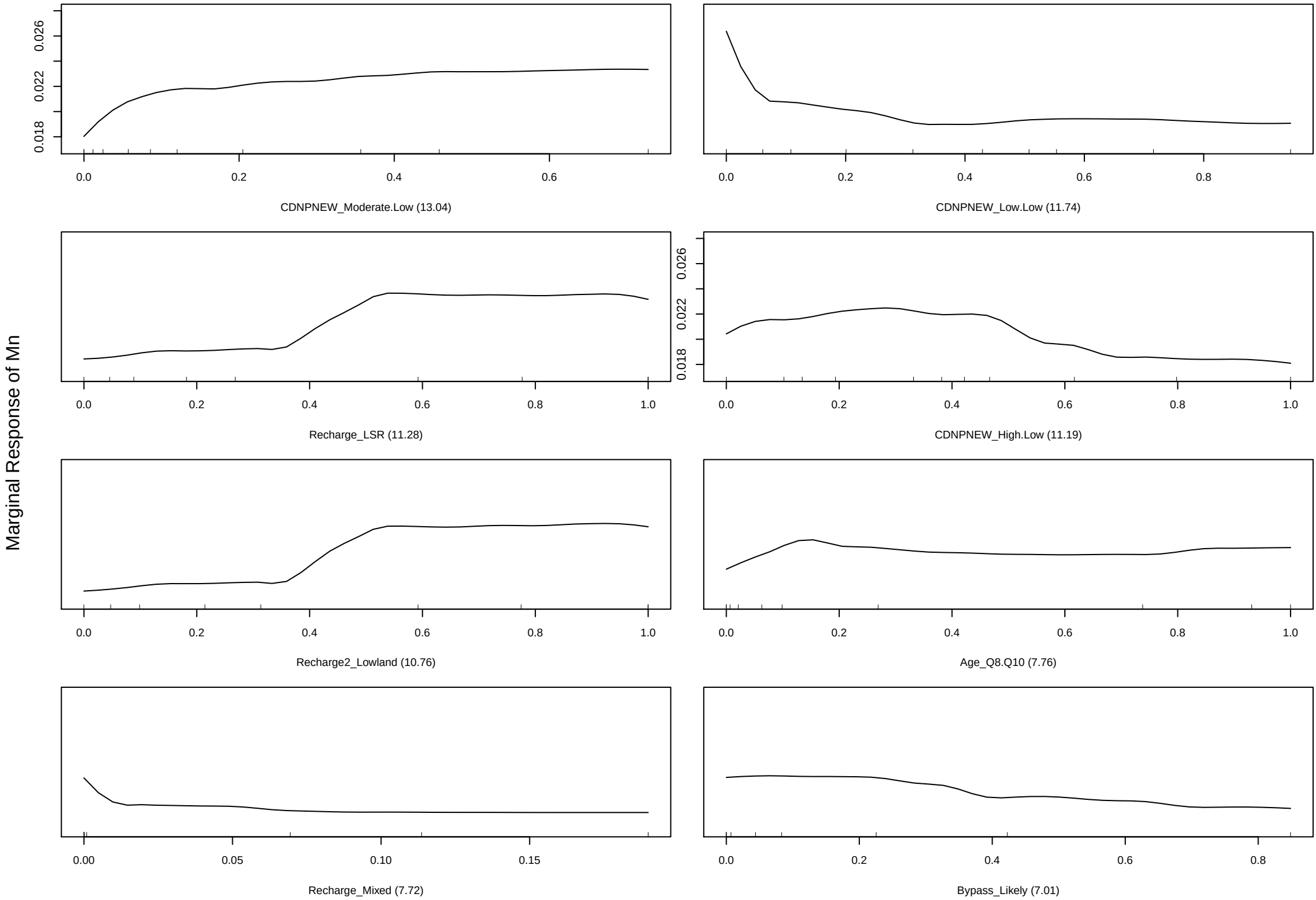

Marginal Response of AlkHCO3
30
40
50
60
0.0
0.2
0.4
0.6
0.8
Bypass_Likely (15.92)
0.0
0.1
0.2
0.3
0.4
0.5
Subsurface_Limestone (15.53)
Recharge_ARR (12.99)
Subsurface_UndiffClastics (12.89)
1.0

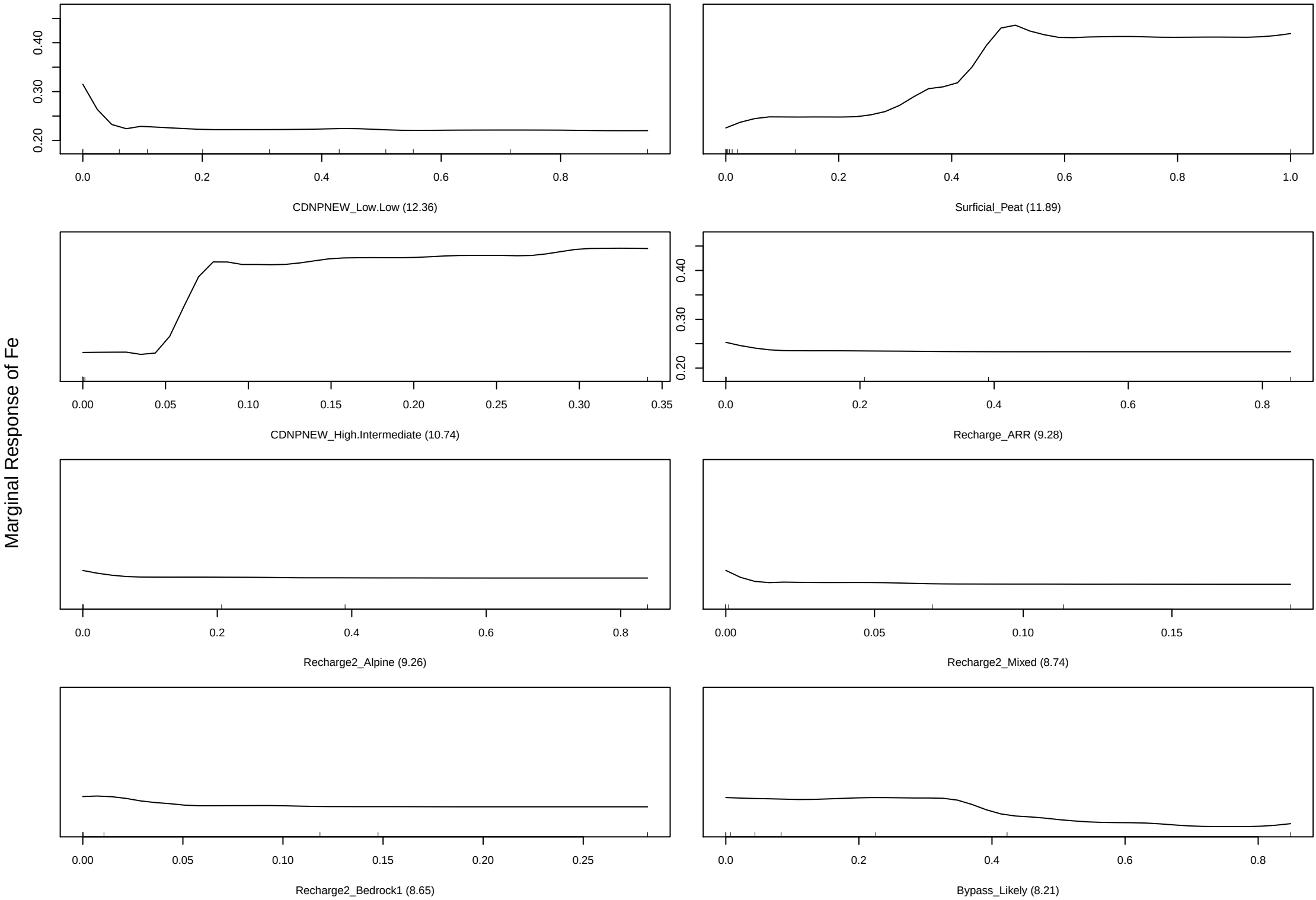

Marginal Response of DOC
PPT_coastal_2 (16.25)
CDNPNEW_Low.Low (15.39)
CDNPNEW_Moderate.Low (13.63)
Surficial_Peat (12.92)
Subsurface_Peat (11.29)
CDNPNEW_High.High (8.25)

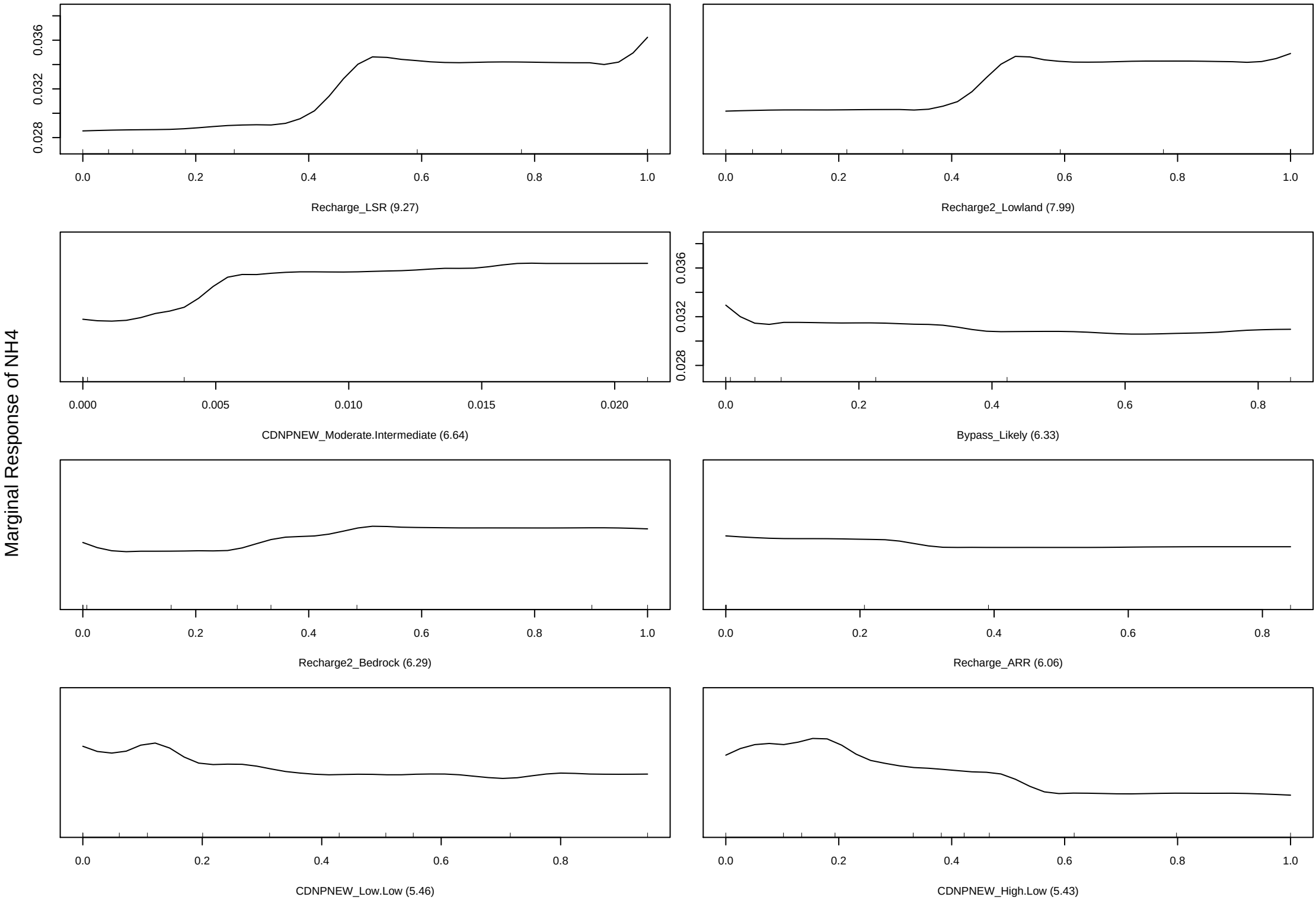

Marginal Response of pHField
6.6
6.2
5.8
0.0
0.2
0.4
0.6
0.8
1.0
Surficial_Felsic (9.7)
0.0
0.1
0.2
0.3
0.4
0.5
Subsurface_Limestone (8.69)
0.000
0.005
0.010
0.015
0.020
CDNPNEW_Moderate.Intermediate (8.5)
CDNPNEW_High.Low (8.4)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
CDNPNEW_High.Intermediate (7.93)
Surficial_Mafic (7.57)
Subsurface_UndiffClastics (7.2)
PPT_inland (6.99)

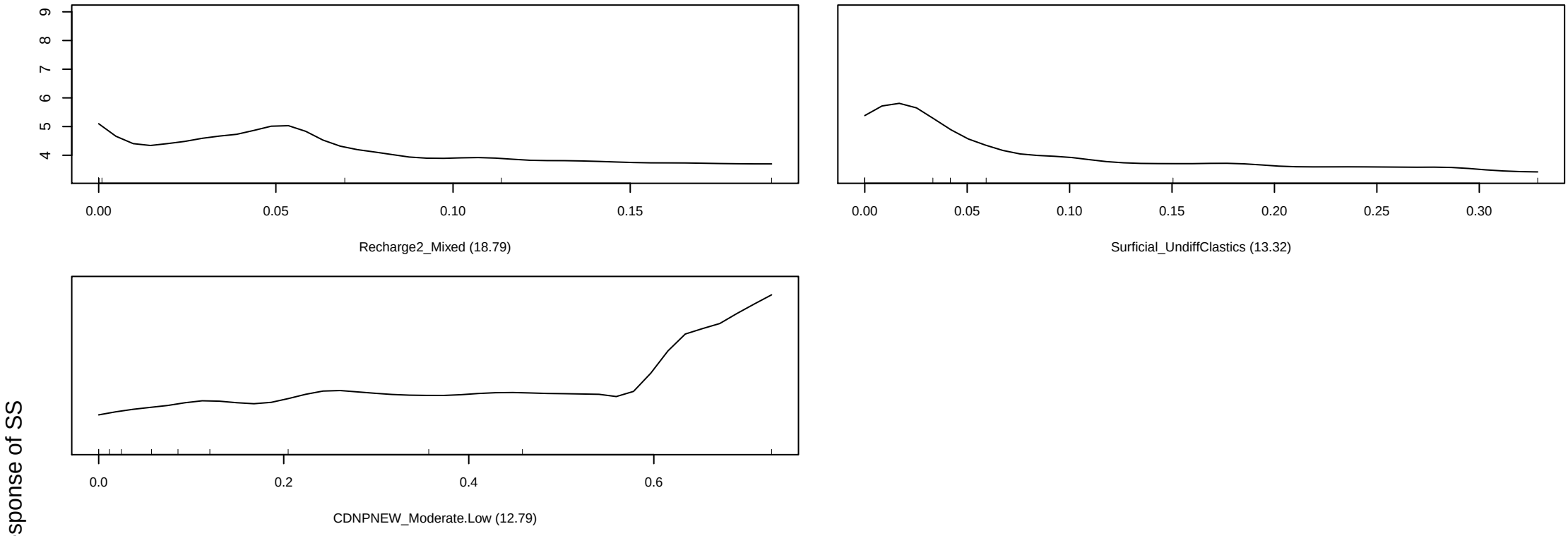

Marginal Response of DOField
12.0
11.0
10.0
0.0
0.2
0.4
0.6
0.8
1.0
CDNPNEW_Low.Low (12.73)
CDNPNEW_High.Low (10.84)
Subsurface_Peat (10.67)
Surficial_Peat (10.37)
Age_Q2.Q4 (8.13)
CDNPNEW_Moderate.Low (4.38)

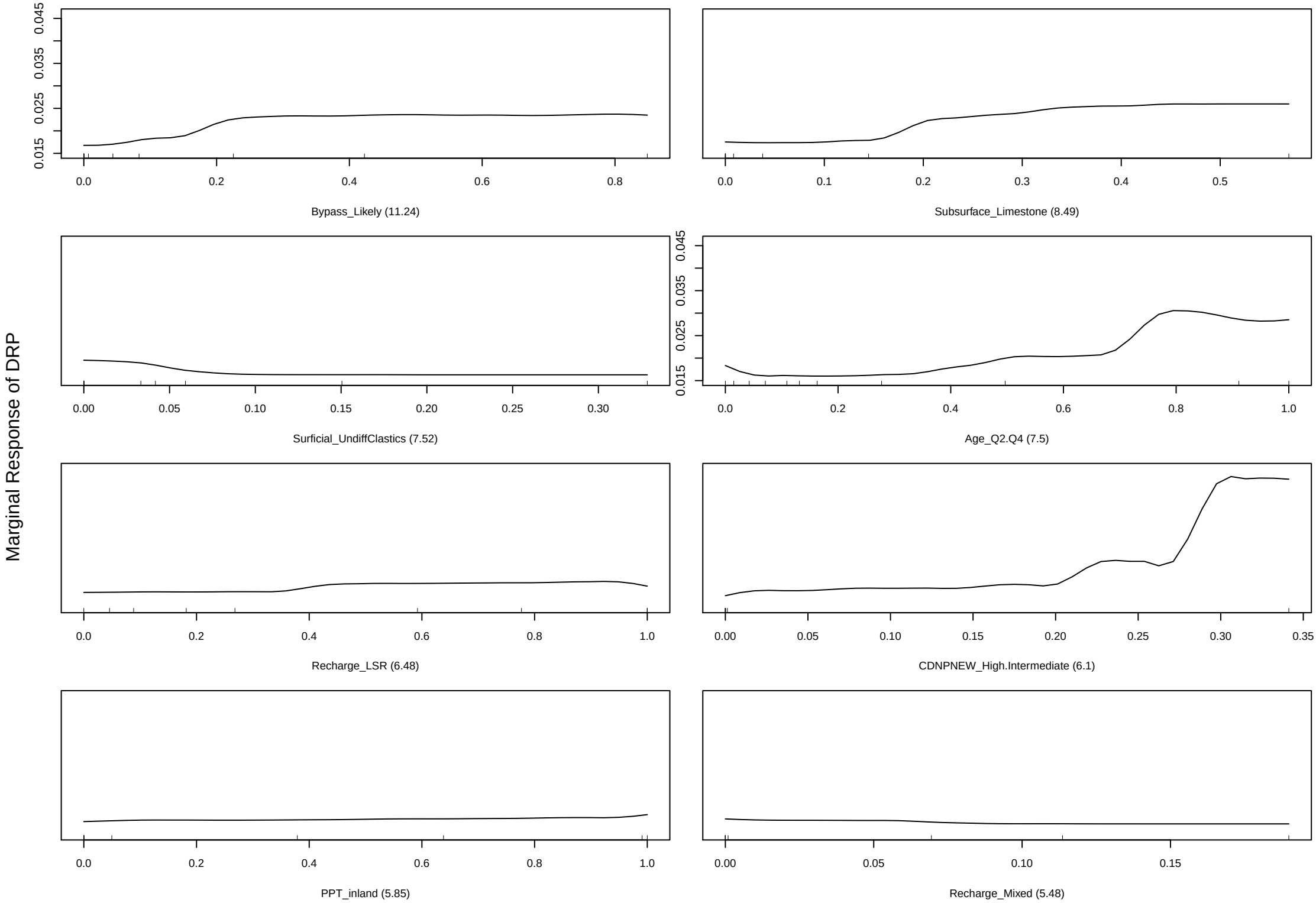

Marginal Response of TON
CDNPNEW_Low.Low (16.29)
Age_Q8.Q10 (15.44)
Bypass_Likely (13.65)
Recharge2_Lowland (12.1)
Recharge_LSR (12.07)
Surficial_Mafic (11.86)
Subsurface_UndiffClastics (11.13)
Subsurface_Peat (9.1)

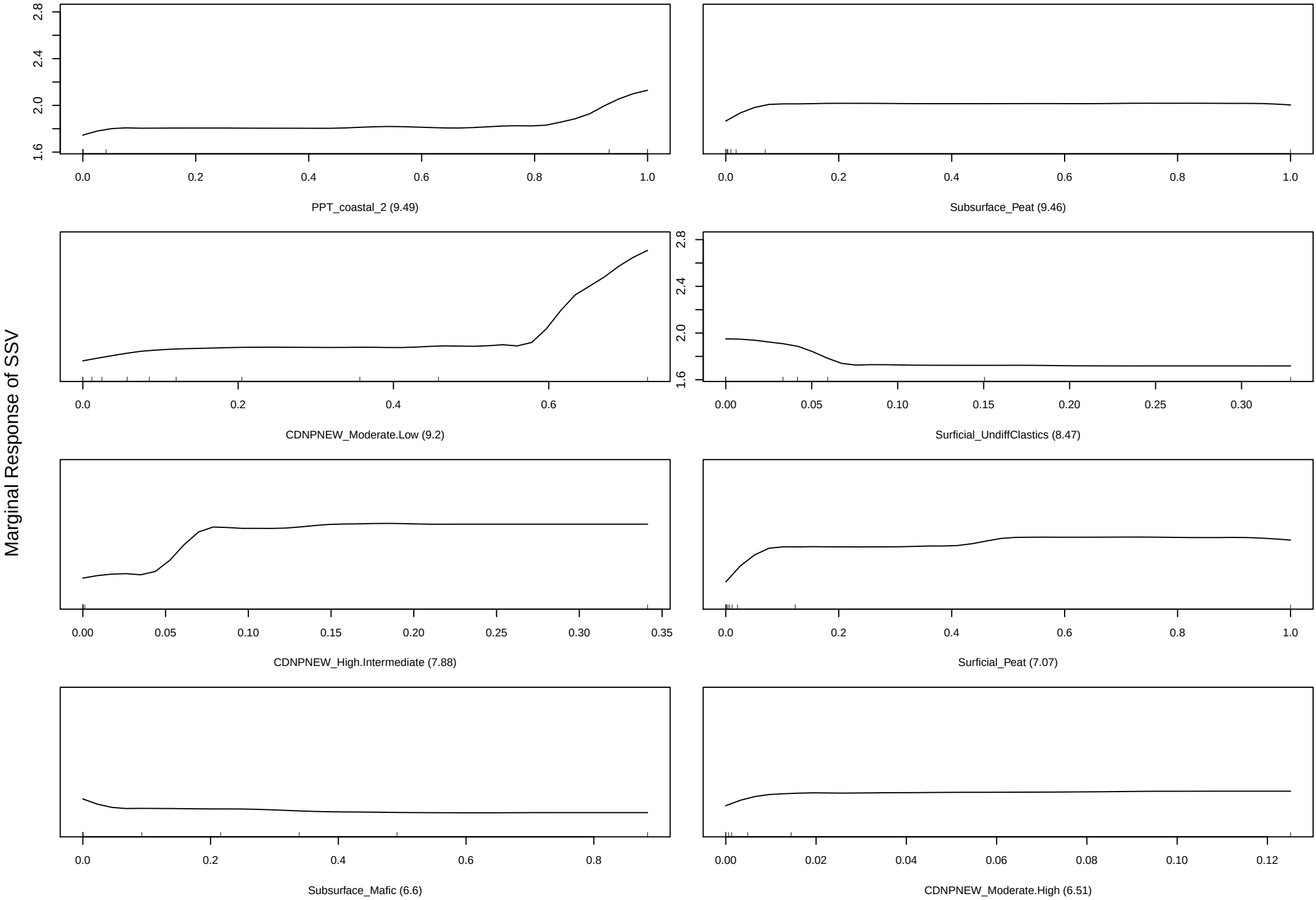

Marginal Response of Mn
0.018
0.022
0.026
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.05
0.10
0.15
CDNPNEW_Moderate.Low (13.04)
CDNPNEW_Low.Low (11.74)
Recharge_LSR (11.28)
CDNPNEW_High.Low (11.19)
Recharge2_Lowland (10.76)
Age_Q8.Q10 (7.76)
Recharge_Mixed (7.72)
Bypass_Likely (7.01)

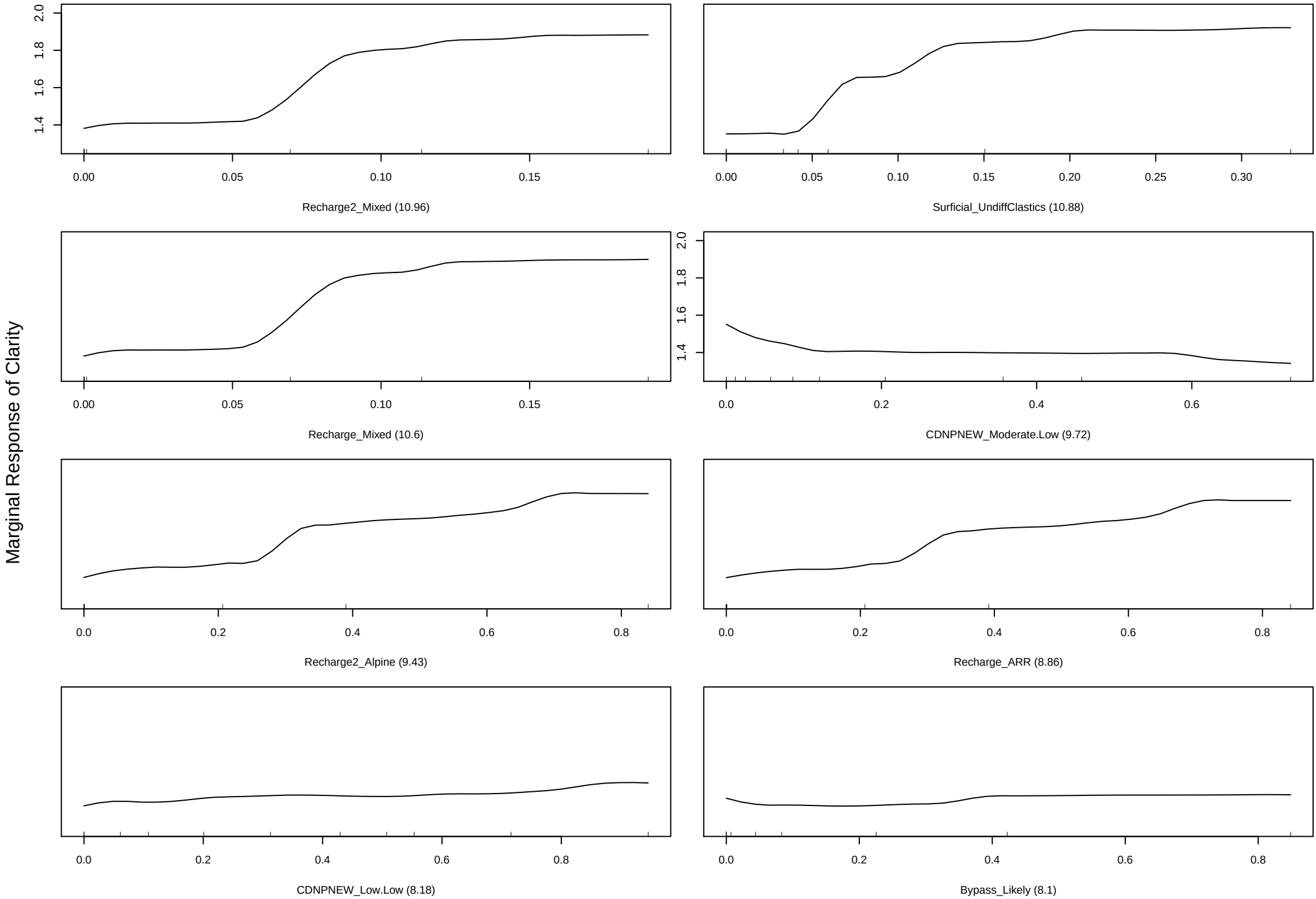

Marginal Response of Fe
0.20
0.30
0.40
0.0
0.2
0.4
0.6
0.8
CDNPNEW_Low.Low (12.36)
1.0
Surficial_Peat (11.89)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
CDNPNEW_High.Intermediate (10.74)
Recharge_ARR (9.28)
Recharge2_Alpine (9.26)
Recharge2_Mixed (8.74)
Recharge2_Bedrock1 (8.65)
Bypass_Likely (8.21)

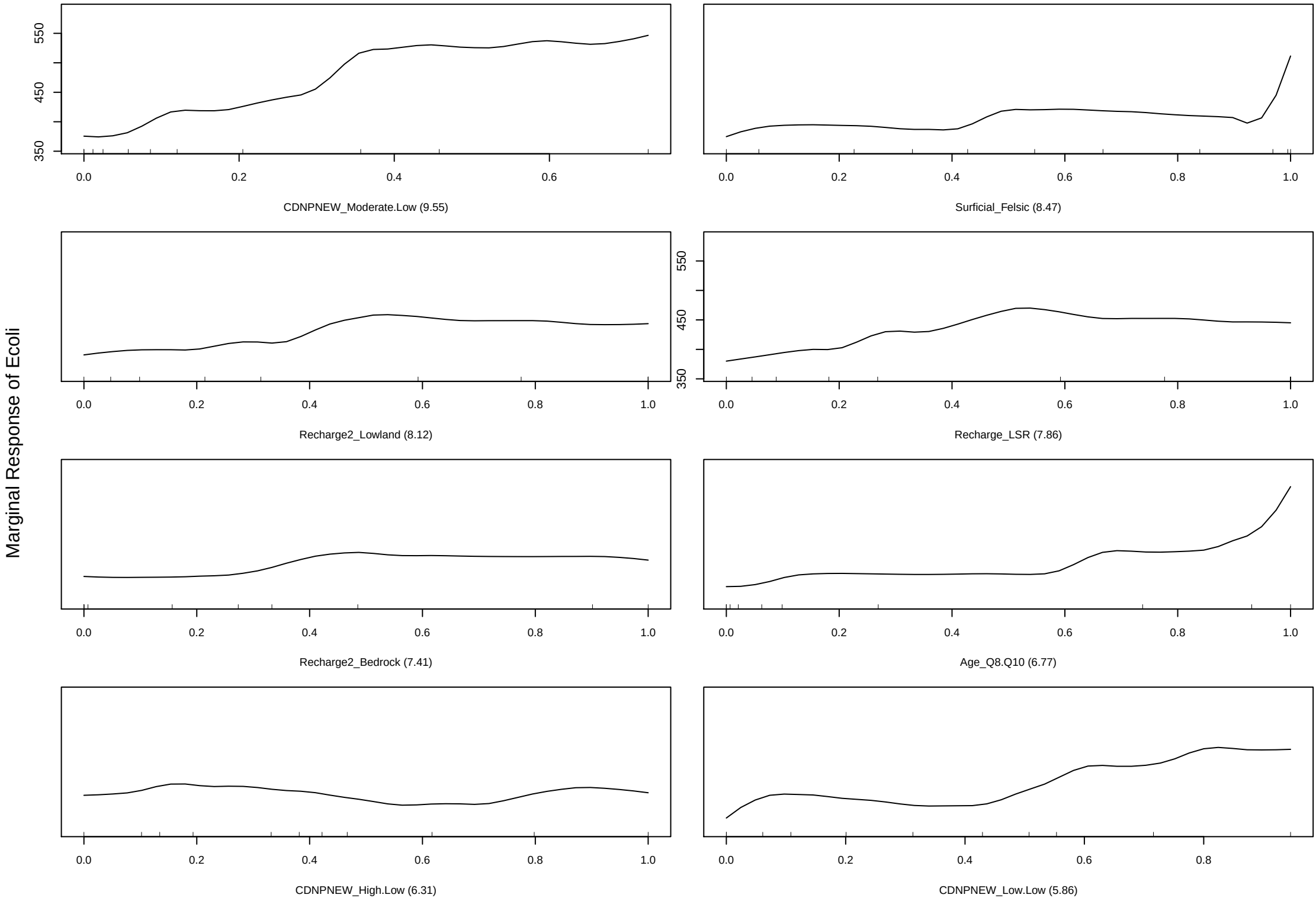

Marginal Response of NH4
0.028
0.032
0.036
0.0
0.2
0.4
0.6
0.8
1.0
Recharge_LSR (9.27)
Recharge2_Lowland (7.99)
0.000
0.005
0.010
0.015
0.020
CDNPNEW_Moderate.Intermediate (6.64)
Bypass_Likely (6.33)
Recharge2_Bedrock (6.29)
Recharge_ARR (6.06)
CDNPNEW_Low.Low (5.46)
CDNPNEW_High.Low (5.43)

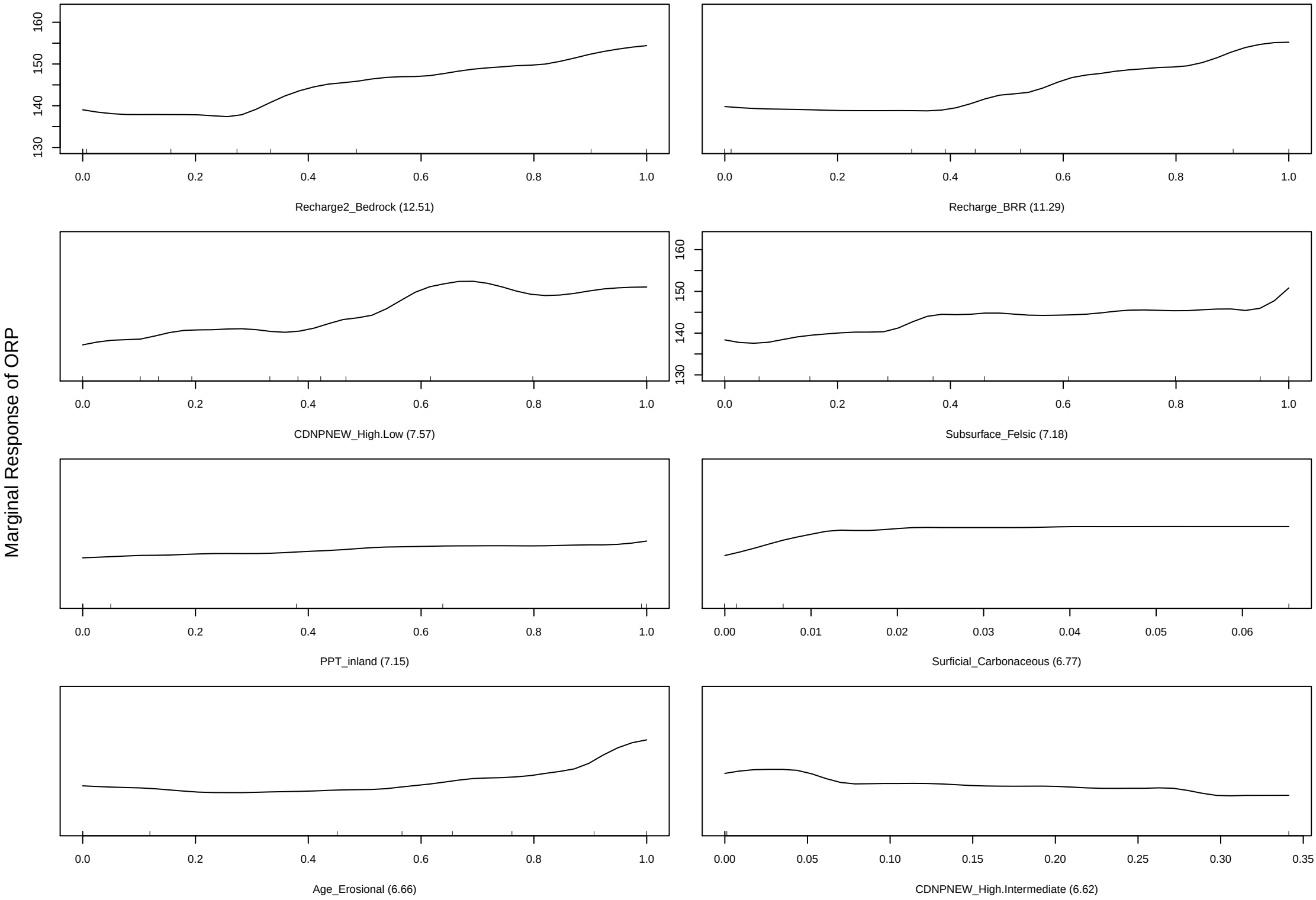

Marginal Response of SS
4
5
6
7
8
9
0.00
0.05
0.10
0.15
Recharge2_Mixed (18.79)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Surficial_UndiffClastics (13.32)
0.0
0.2
0.4
0.6
CDNPNEW_Moderate.Low (12.79)

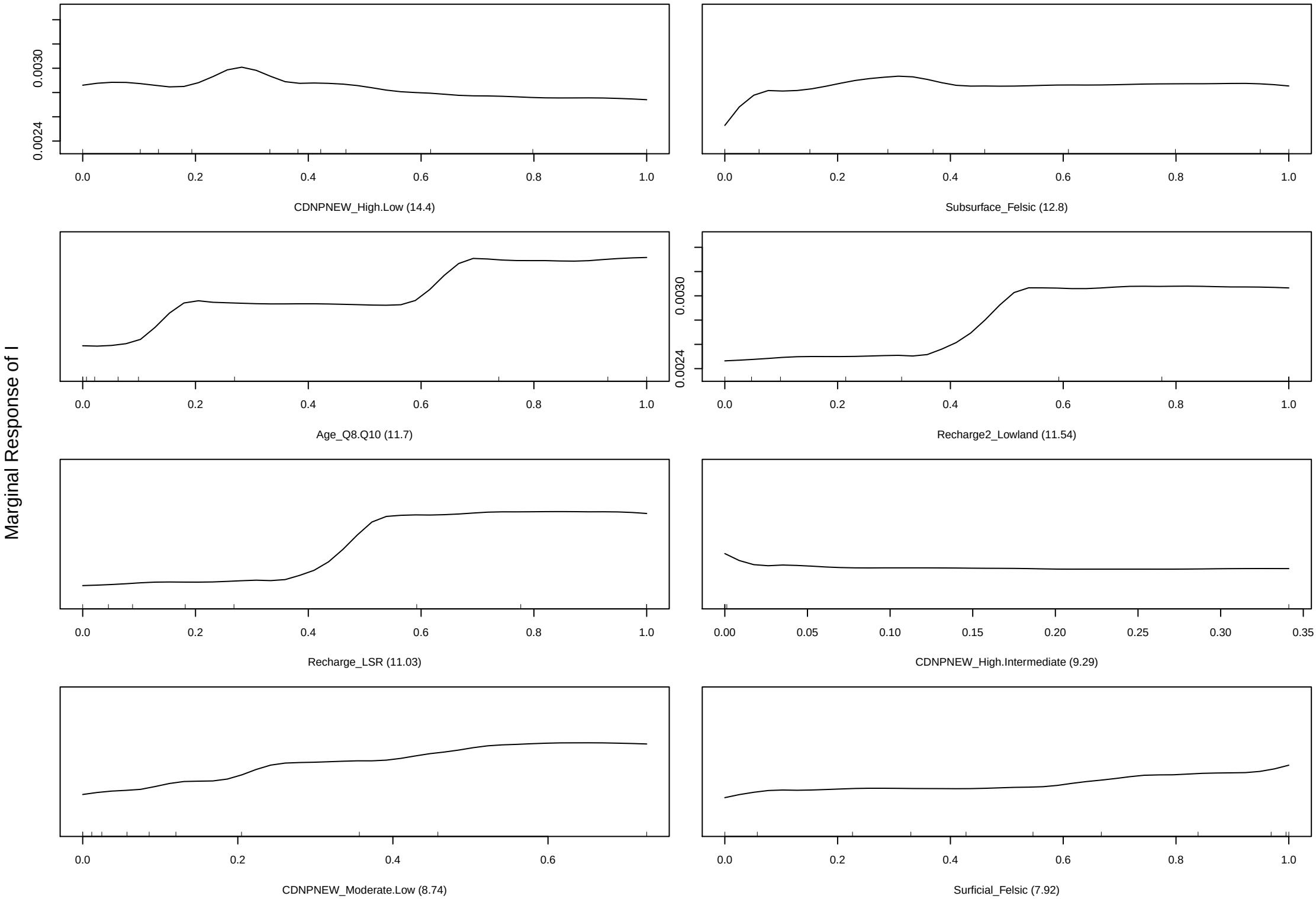

Marginal Response of DRP
0.015
0.025
0.035
0.045
0.0
0.2
0.4
0.6
0.8
Bypass_Likely (11.24)
0.1
0.3
0.5
Subsurface_Limestone (8.49)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Surficial_UndiffClastics (7.52)
1.0
Age_Q2.Q4 (7.5)
Recharge_LSR (6.48)
0.35
CDNPNEW_High.Intermediate (6.1)
PPT_inland (5.85)
Recharge_Mixed (5.48)

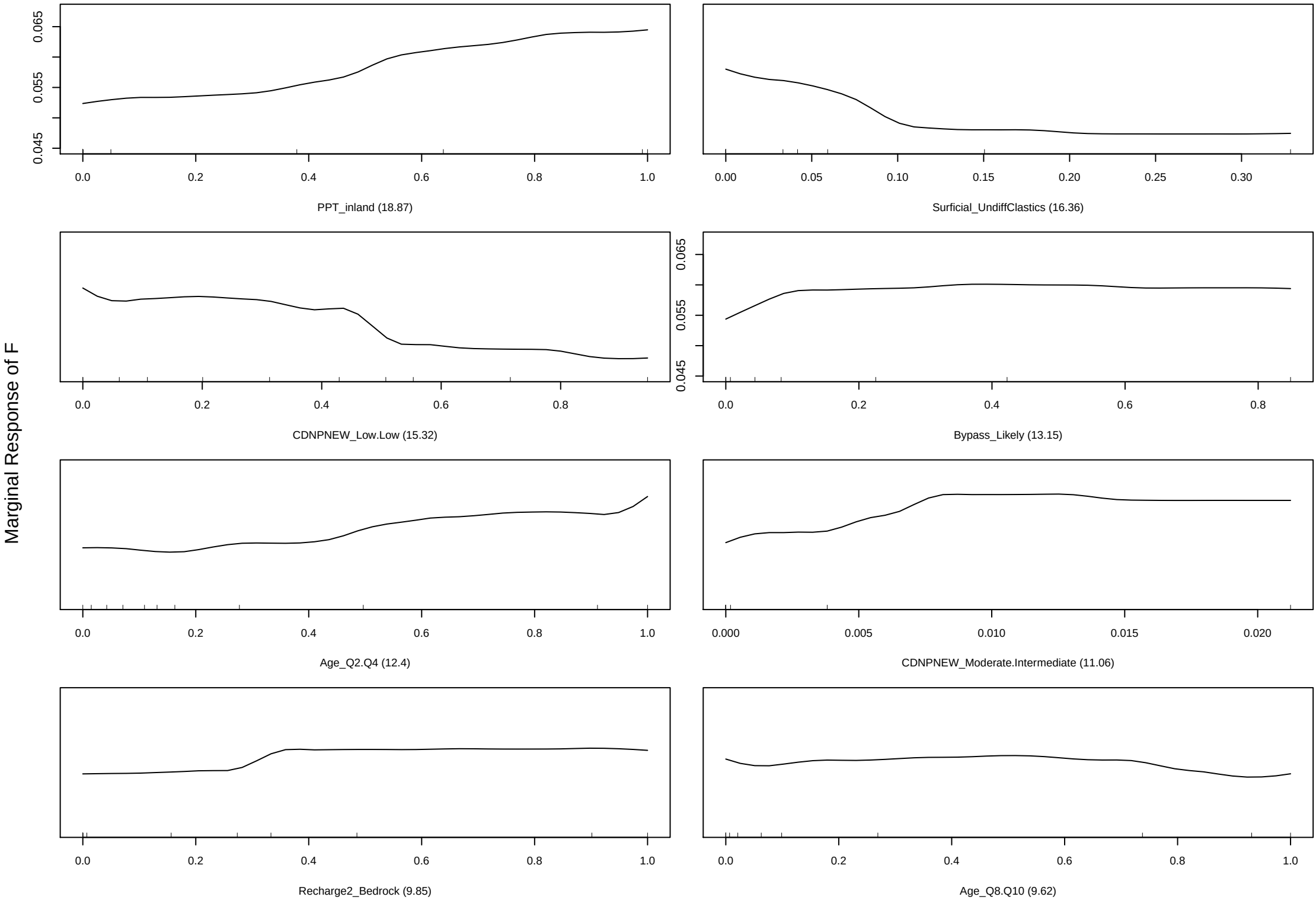

Marginal Response of SSV
PPT_coastal_2 (9.49)
Subsurface_Peat (9.46)
CDNPNEW_Moderate.Low (9.2)
Surficial_UndiffClastics (8.47)
CDNPNEW_High.Intermediate (7.88)
Surficial_Peat (7.07)
Subsurface_Mafic (6.6)
CDNPNEW_Moderate.High (6.51)

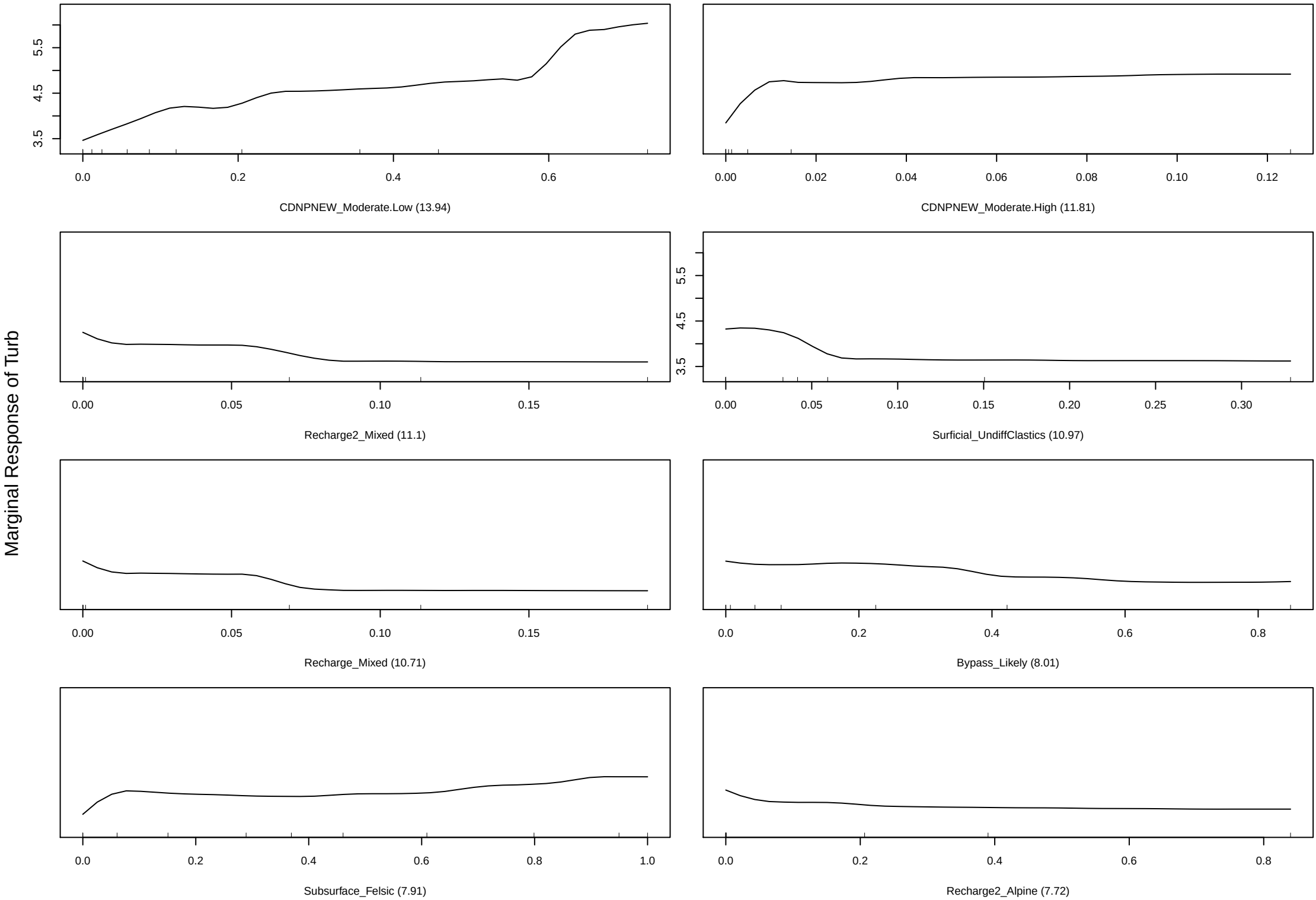

Marginal Response of Clarity
Recharge2_Mixed (10.96)
Surficial_UndiffClastics (10.88)
Recharge_Mixed (10.6)
CDNPNEW_Moderate.Low (9.72)
Recharge2_Alpine (9.43)
Recharge_ARR (8.86)
CDNPNEW_Low.Low (8.18)
Bypass_Likely (8.1)

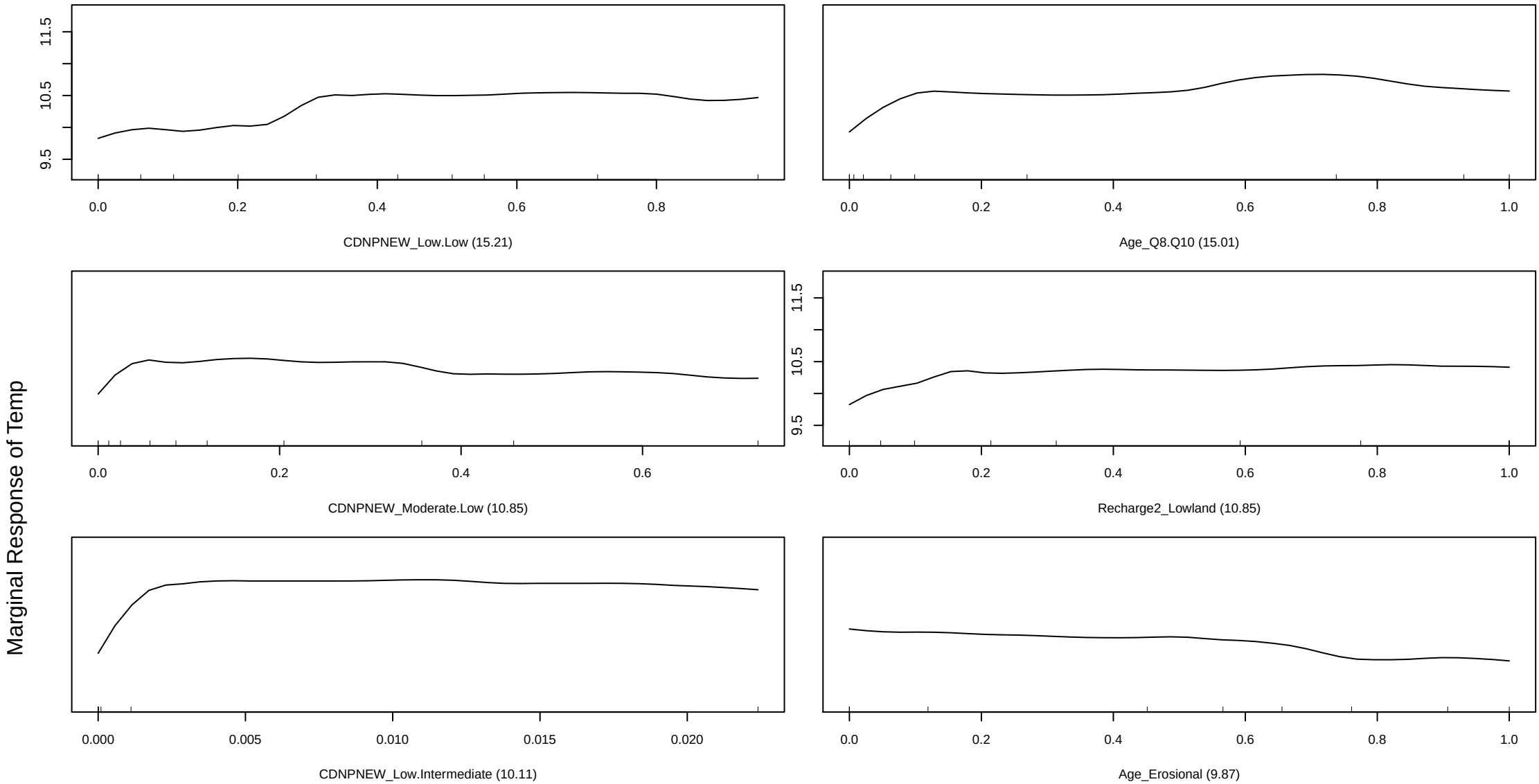

Marginal Response of Ecoli
350
450
550
0.0
0.2
0.4
0.6
0.8
1.0
CDNPNEW_Moderate.Low (9.55)
Surficial_Felsic (8.47)
Recharge2_Lowland (8.12)
Recharge_LSR (7.86)
Recharge2_Bedrock (7.41)
Age_Q8.Q10 (6.77)
CDNPNEW_High.Low (6.31)
CDNPNEW_Low.Low (5.86)

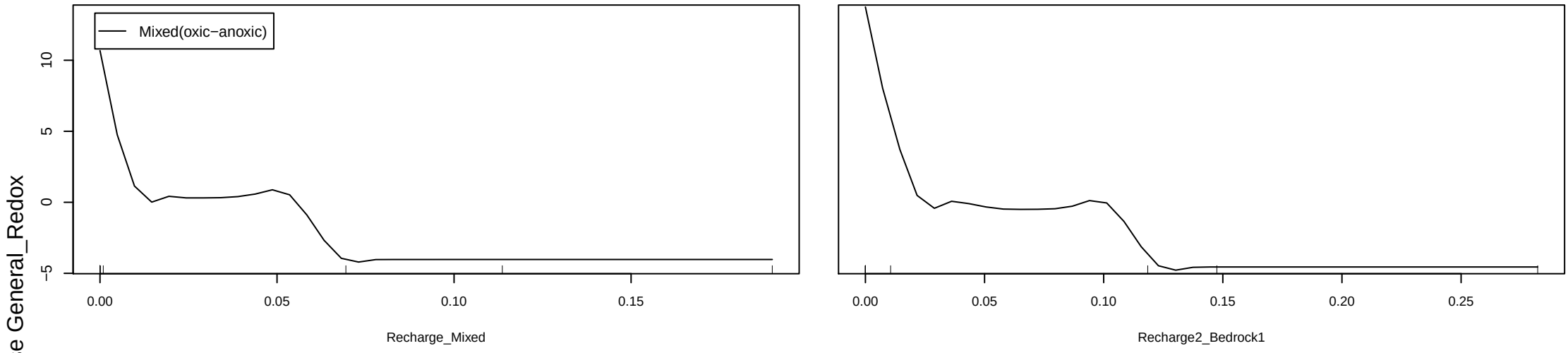

Marginal Response of ORP
130
140
150
160
0.0
0.2
0.4
0.6
0.8
1.0
Recharge2_Bedrock (12.51)
Recharge_BRR (11.29)
CDNPNEW_High.Low (7.57)
Subsurface_Felsic (7.18)
PPT_inland (7.15)
0.00
0.01
0.02
0.03
0.04
0.05
0.06
Surficial_Carbonaceous (6.77)
Age_Erosional (6.66)
0.10
0.15
0.20
0.25
0.30
0.35
CDNPNEW_High.Intermediate (6.62)

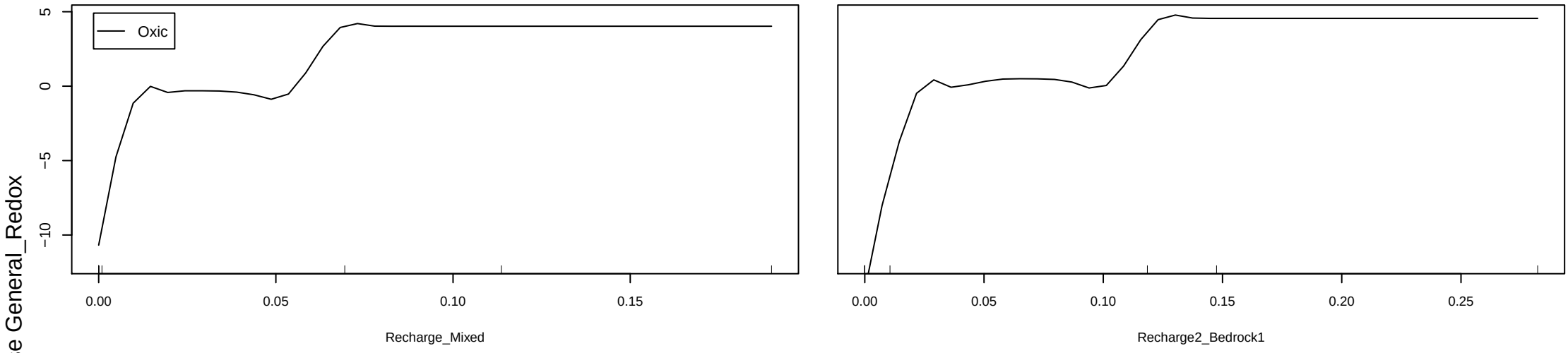

Marginal Response of I
0.0030
0.0024
0.0
0.2
0.4
0.6
0.8
1.0
CDNPNEW_High.Low (14.4)
Subsurface_Felsic (12.8)
Age_Q8.Q10 (11.7)
Recharge2_Lowland (11.54)
Recharge_LSR (11.03)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
CDNPNEW_High.Intermediate (9.29)
CDNPNEW_Moderate.Low (8.74)
Surficial_Felsic (7.92)

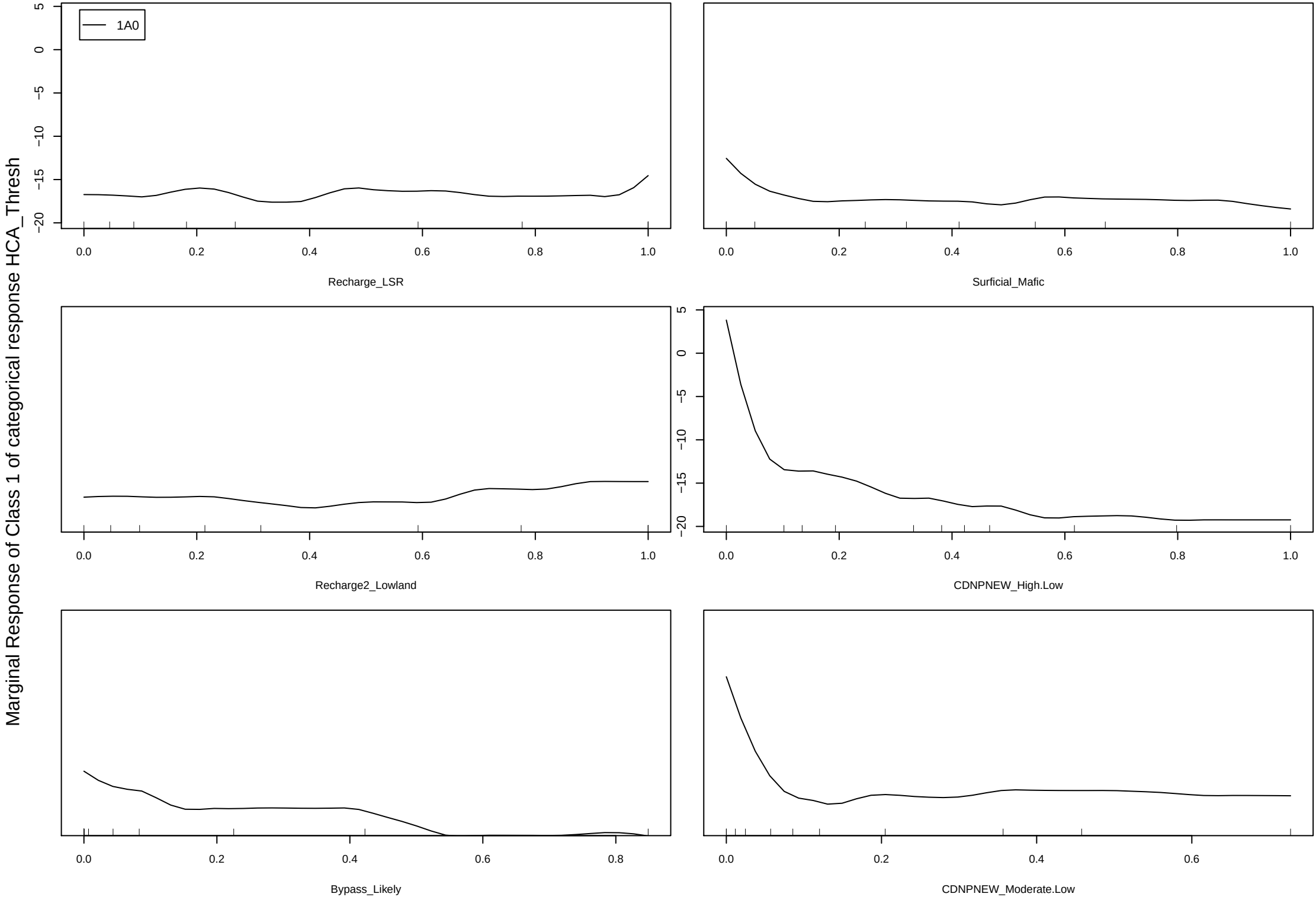

Marginal Response of F
0.045
0.055
0.065
0.0
0.2
0.4
0.6
0.8
1.0
PPT_inland (18.87)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Surficial_UndiffClastics (16.36)
CDNPNEW_Low.Low (15.32)
Bypass_Likely (13.15)
Age_Q2.Q4 (12.4)
0.000
0.005
0.010
0.015
0.020
CDNPNEW_Moderate.Intermediate (11.06)
Recharge2_Bedrock (9.85)
Age_Q8.Q10 (9.62)

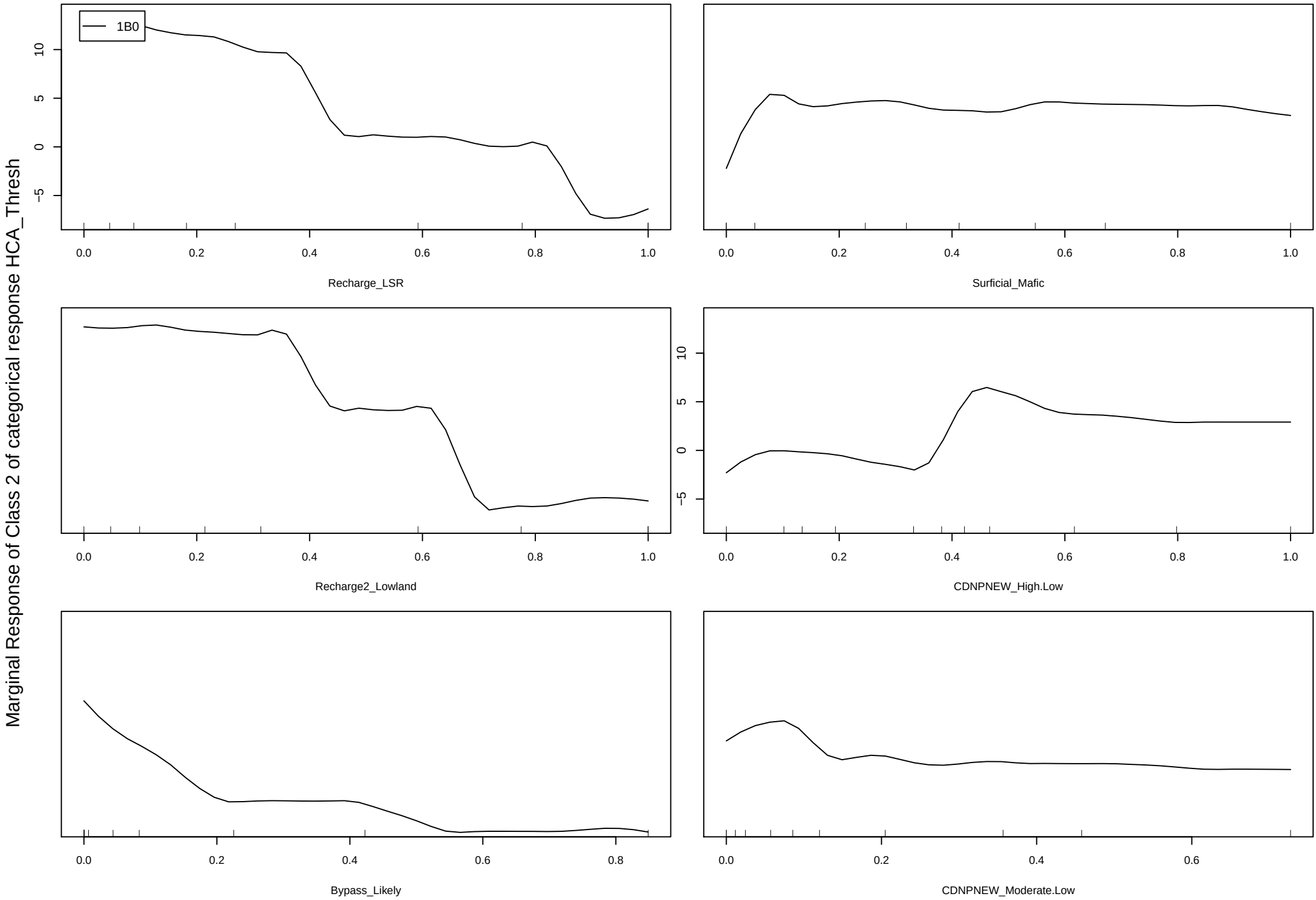

Marginal Response of Turb
3.5
4.5
5.5
0.0
0.2
0.4
0.6
CDNPNEW_Moderate.Low (13.94)
0.00
0.02
0.04
0.06
0.08
0.10
0.12
CDNPNEW_Moderate.High (11.81)
0.00
0.05
0.10
0.15
Recharge2_Mixed (11.1)
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Surficial_UndiffClastics (10.97)
0.00
0.05
0.10
0.15
Recharge_Mixed (10.71)
0.0
0.2
0.4
0.6
0.8
Bypass_Likely (8.01)
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_Felsic (7.91)
0.0
0.2
0.4
0.6
0.8
Recharge2_Alpine (7.72)

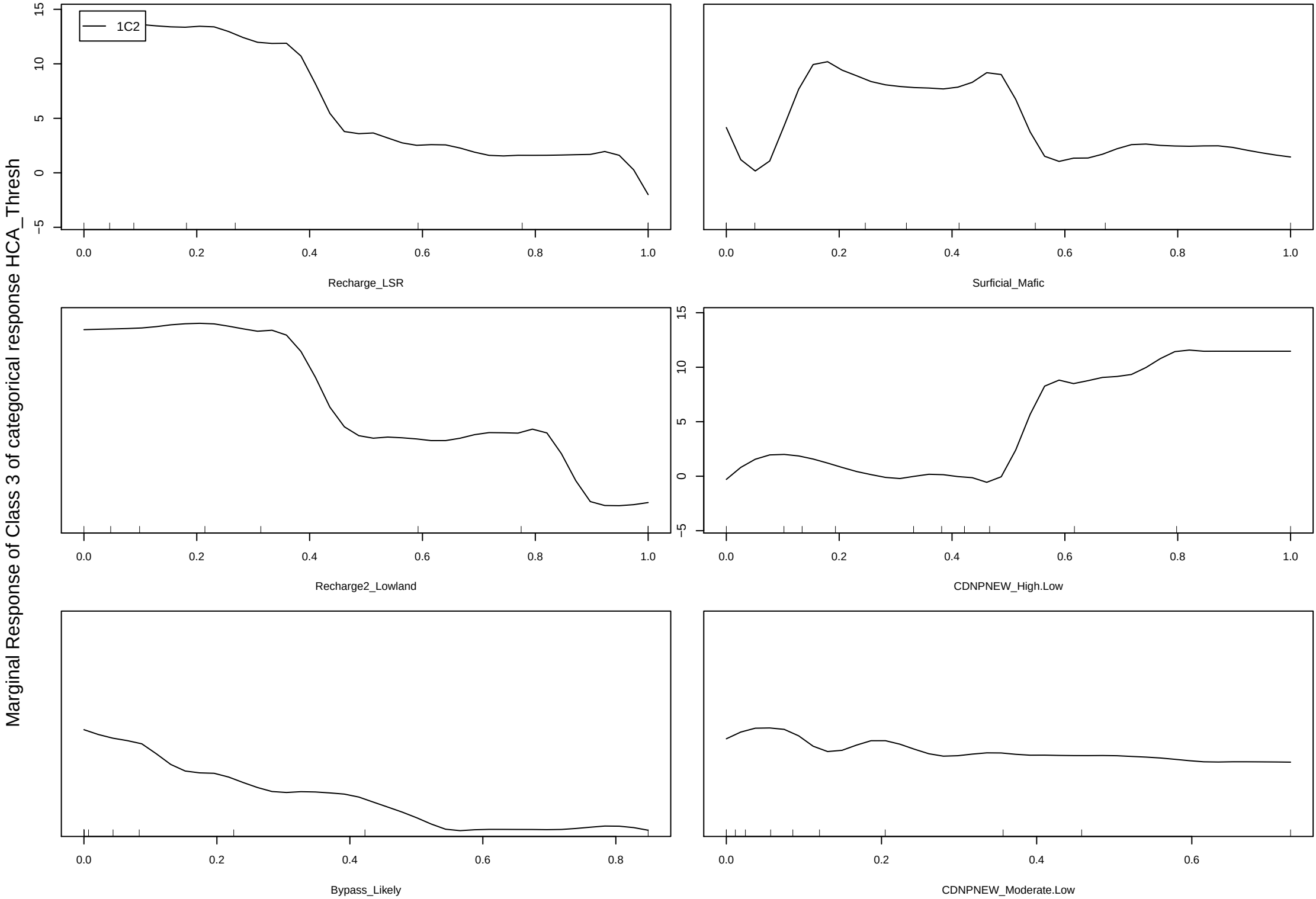

Marginal Response of Temp
11.5
10.5
9.5
0.0
0.2
0.4
0.6
0.8
CDNPNEW_Low.Low (15.21)
1.0
Age_Q8.Q10 (15.01)
CDNPNEW_Moderate.Low (10.85)
Recharge2_Lowland (10.85)
0.000
0.005
0.010
0.015
0.020
CDNPNEW_Low.Intermediate (10.11)
Age_Erosional (9.87)

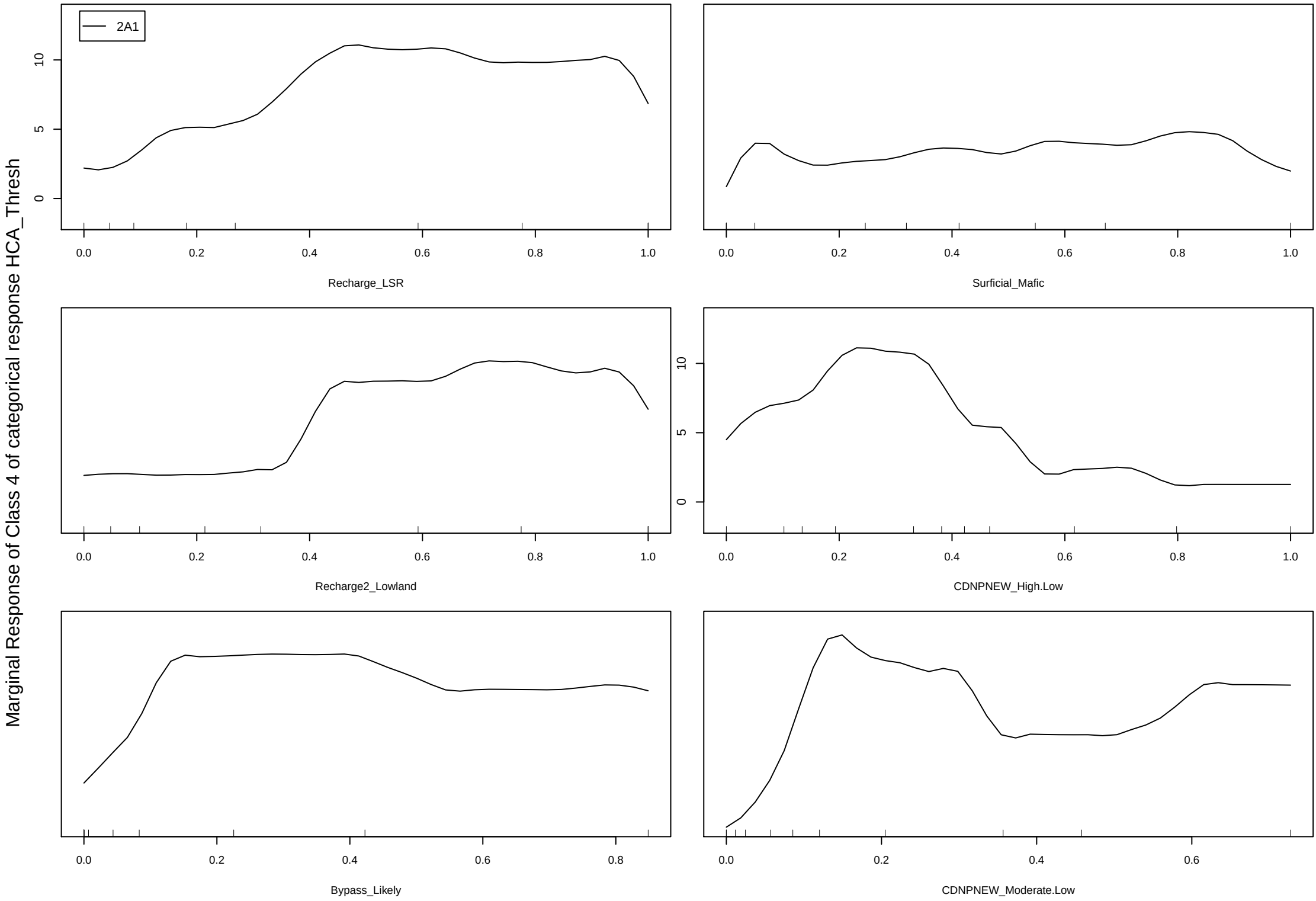

Marginal Response of Class 1 of categorical response General_Redox
Mixed(oxic−anoxic)
10
5
0
−5
0.00
0.05
0.10
0.15
Recharge_Mixed
0.00
0.05
0.10
0.15
0.20
0.25
Recharge2_Bedrock1

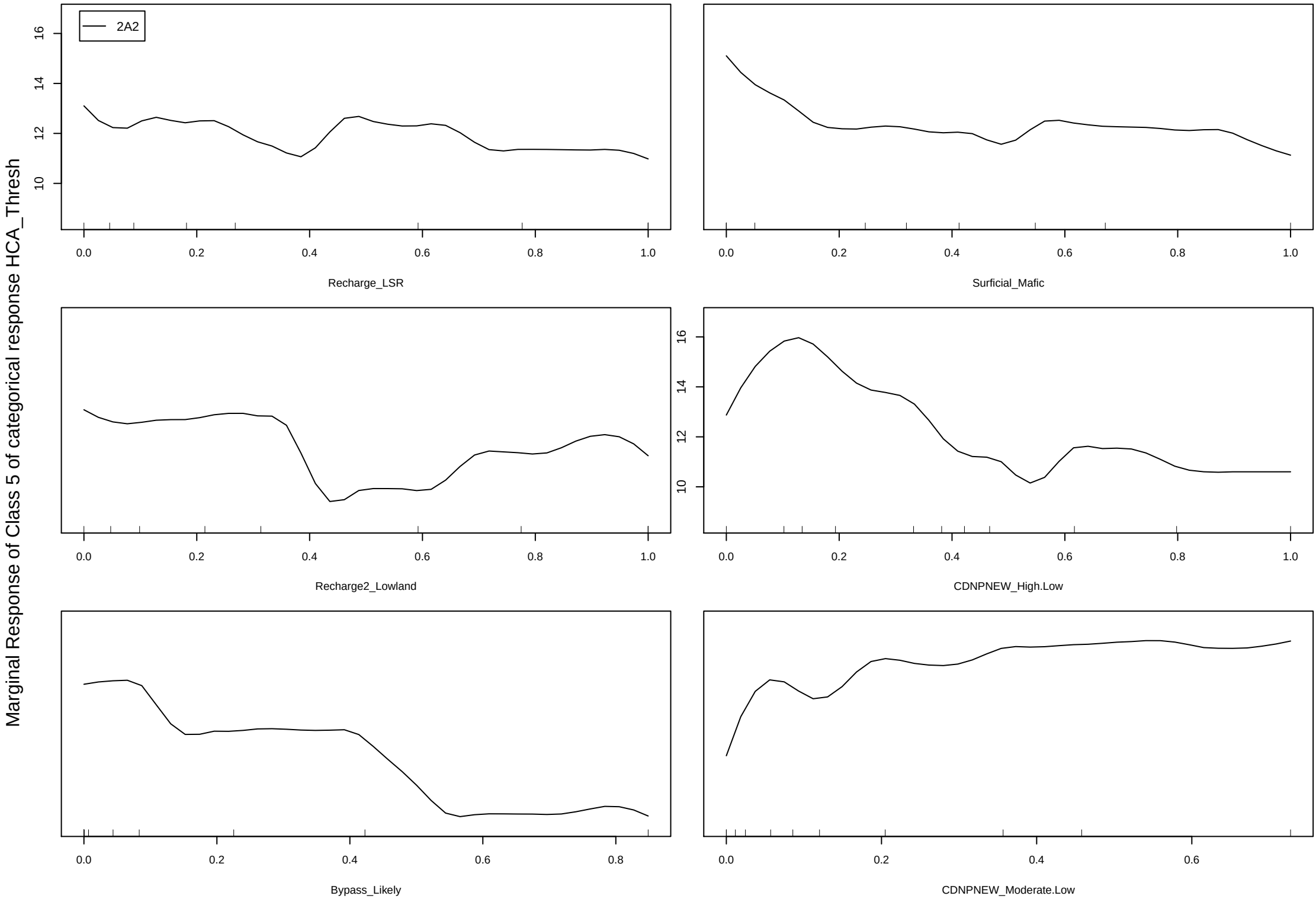

Marginal Response of Class 2 of categorical response General_Redox
Oxic
5
0
−5
−10
0.00
0.05
0.10
0.15
Recharge_Mixed
0.00
0.05
0.10
0.15
0.20
0.25
Recharge2_Bedrock1

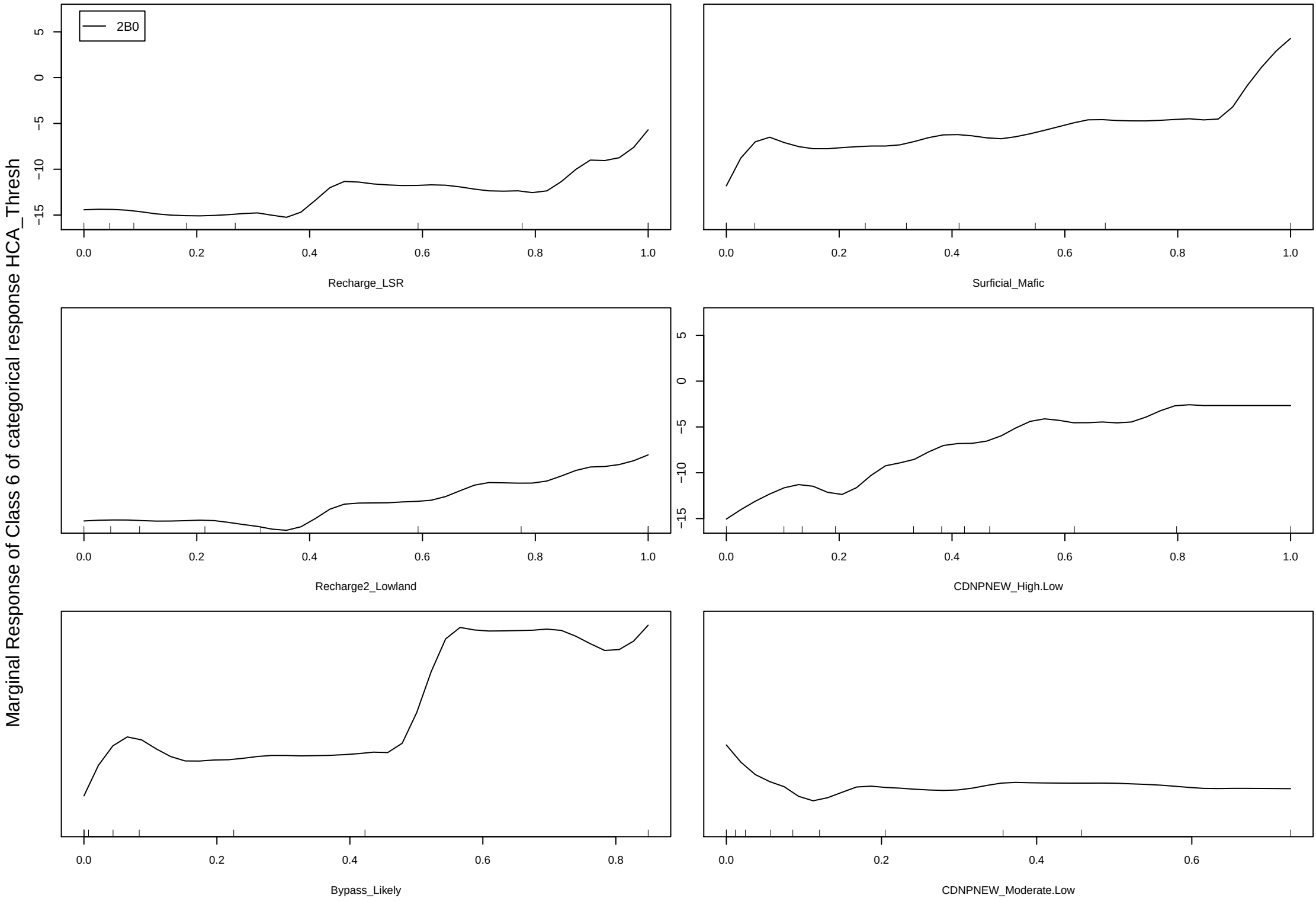

Marginal Response of Class 1 of categorical response HCA_Thresh
1A0
5
0
−5
−10
−15
−20
0.0
0.2
0.4
0.6
0.8
1.0
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

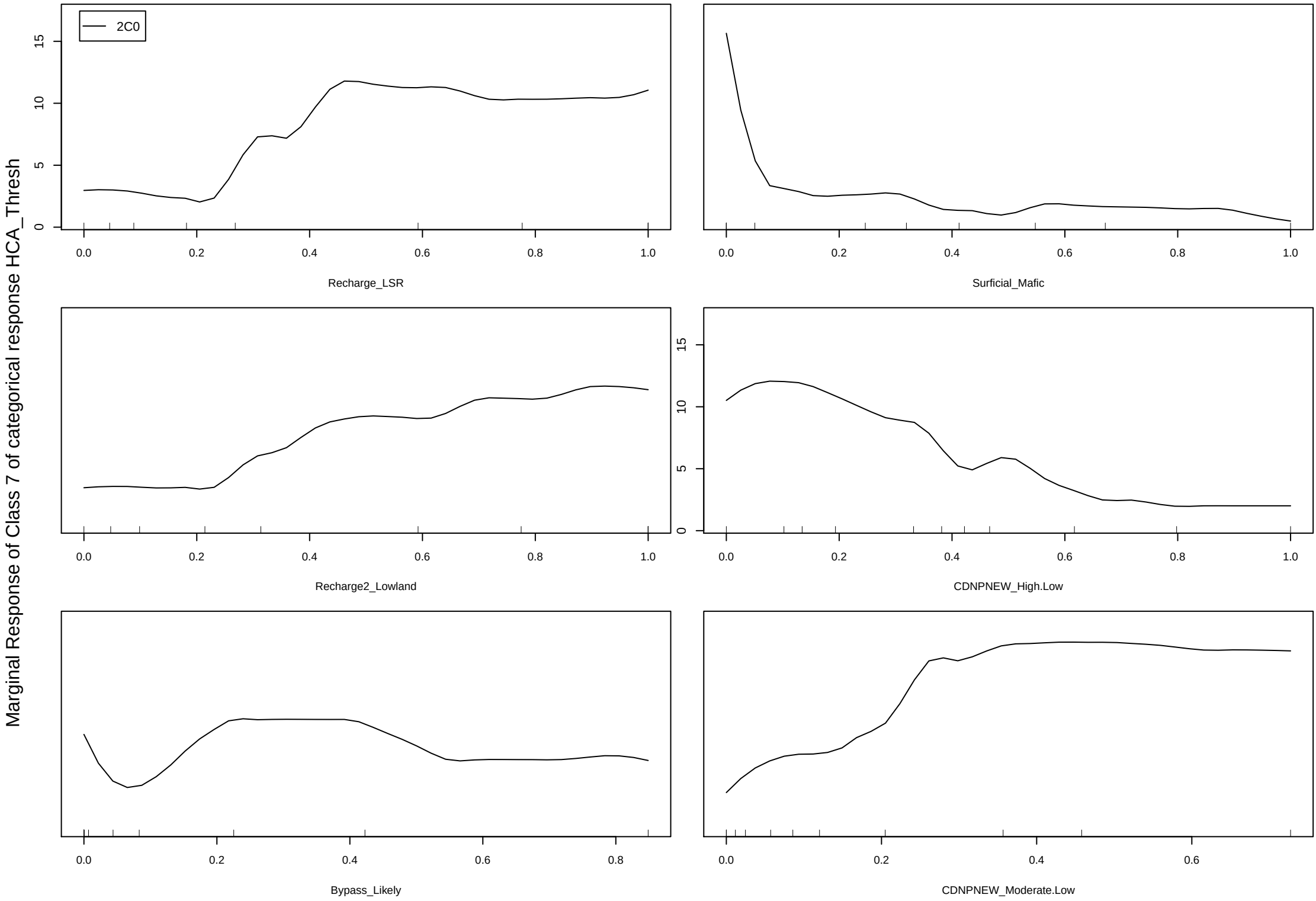

Marginal Response of Class 2 of categorical response HCA_Thresh
1B0
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

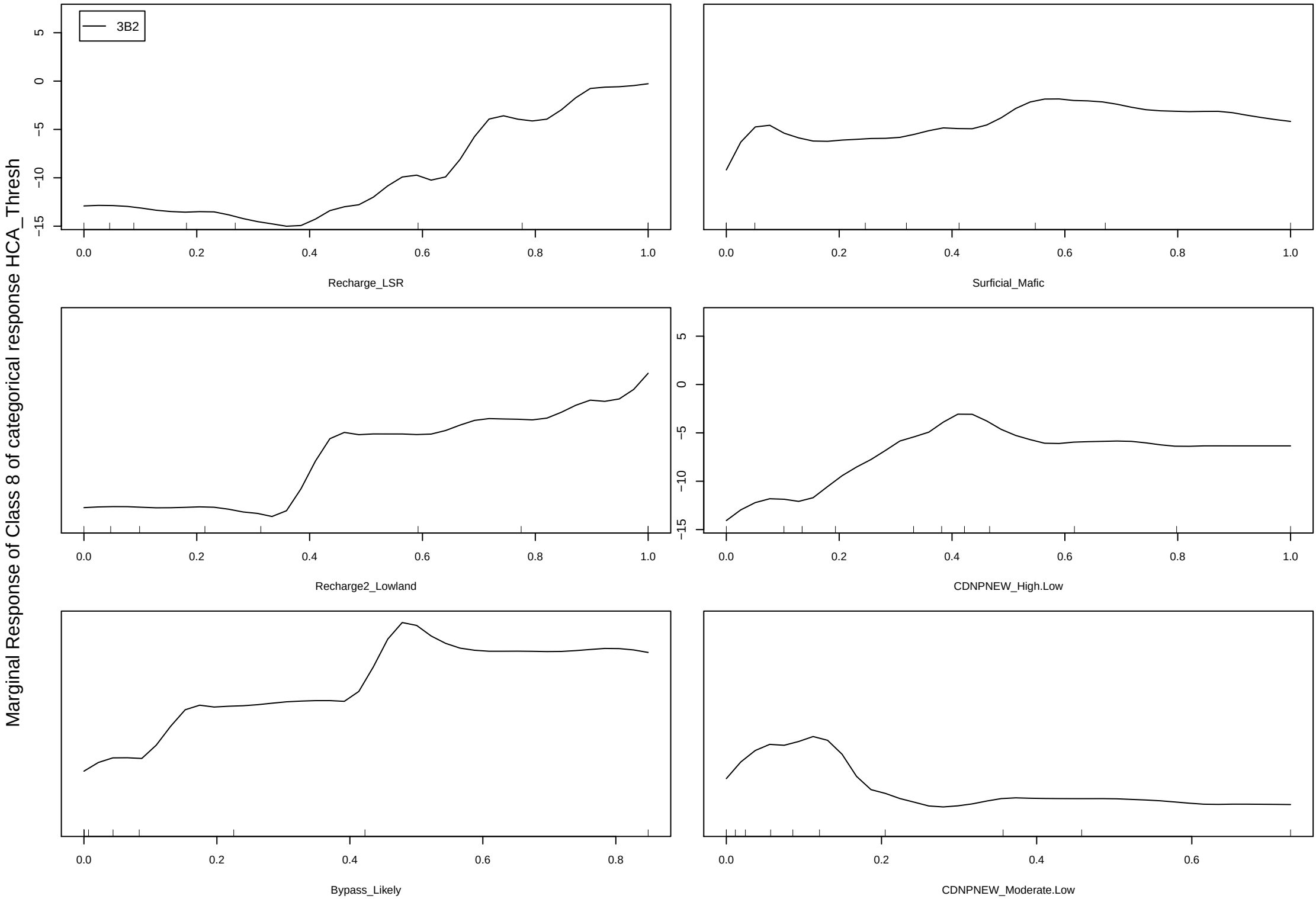

Marginal Response of Class 3 of categorical response HCA_Thresh
1C2
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

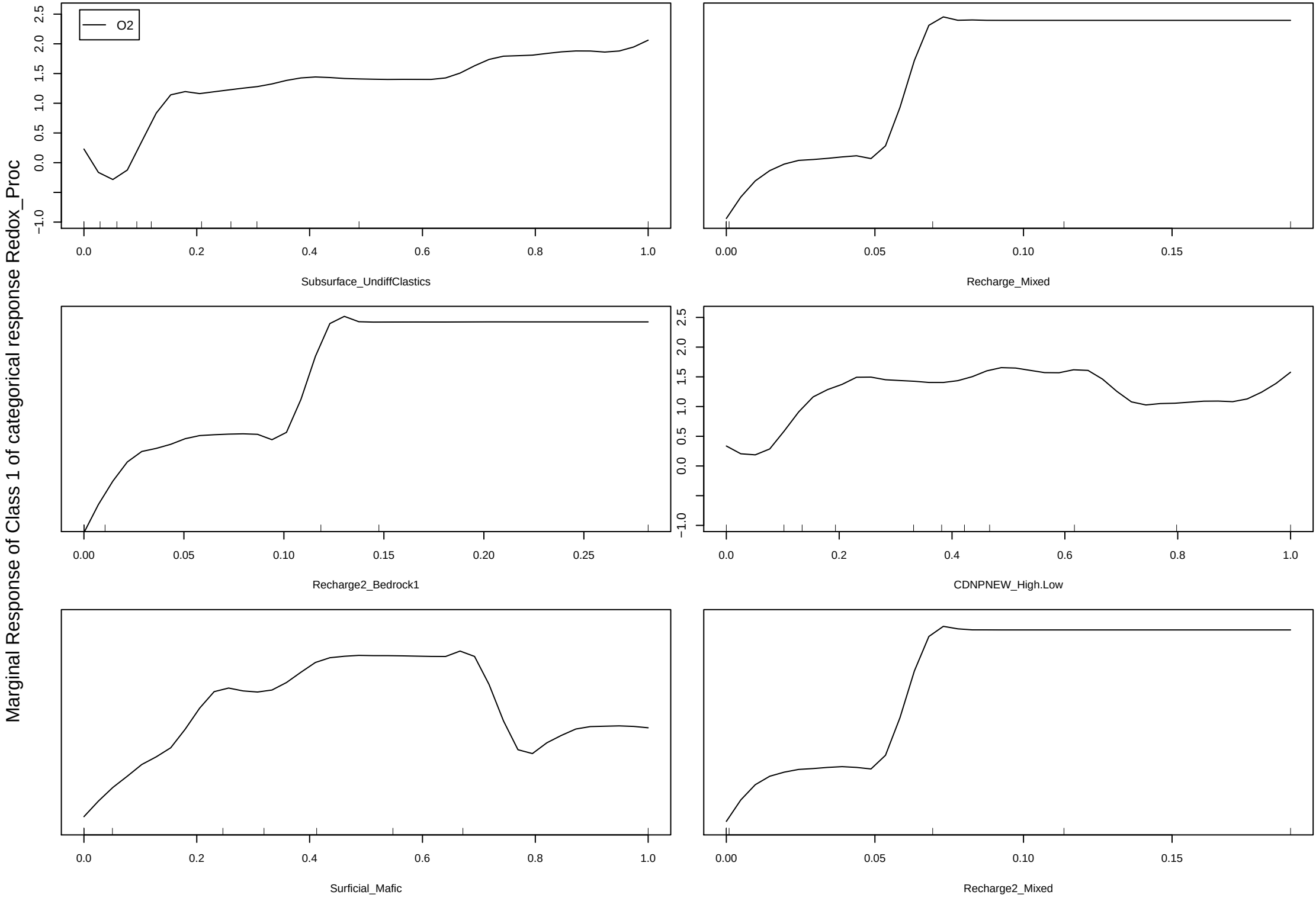

Marginal Response of Class 4 of categorical response HCA_Thresh
2A1
0
5
10
0.0
0.2
0.4
0.6
0.8
1.0
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

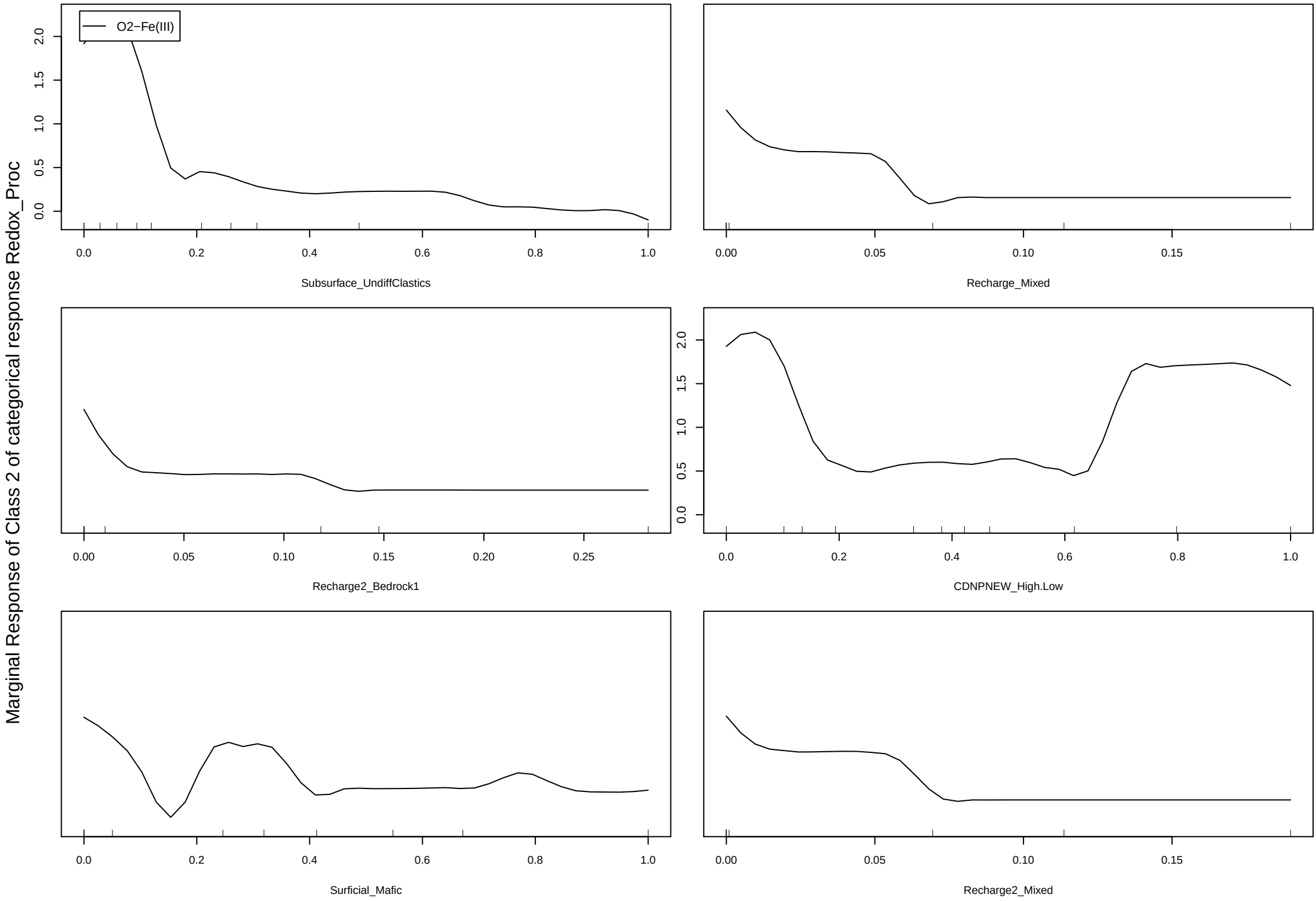

Marginal Response of Class 5 of categorical response HCA_Thresh
2A2
10
12
14
16
0.0
0.2
0.4
0.6
0.8
1.0
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

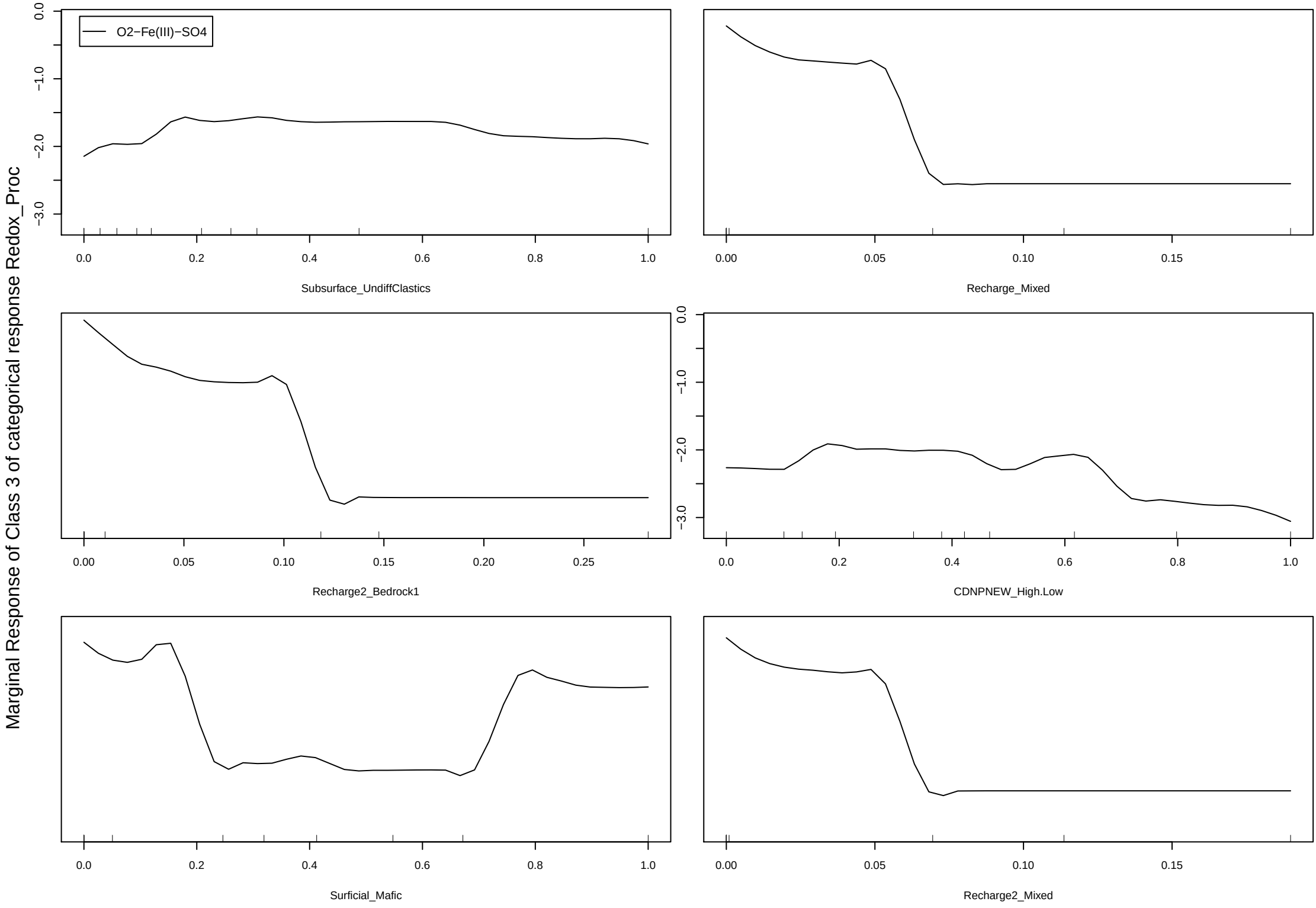

Marginal Response of Class 6 of categorical response HCA_Thresh
2B0
5
0
−5
−10
−15
0.0
0.2
0.4
0.6
0.8
1.0
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

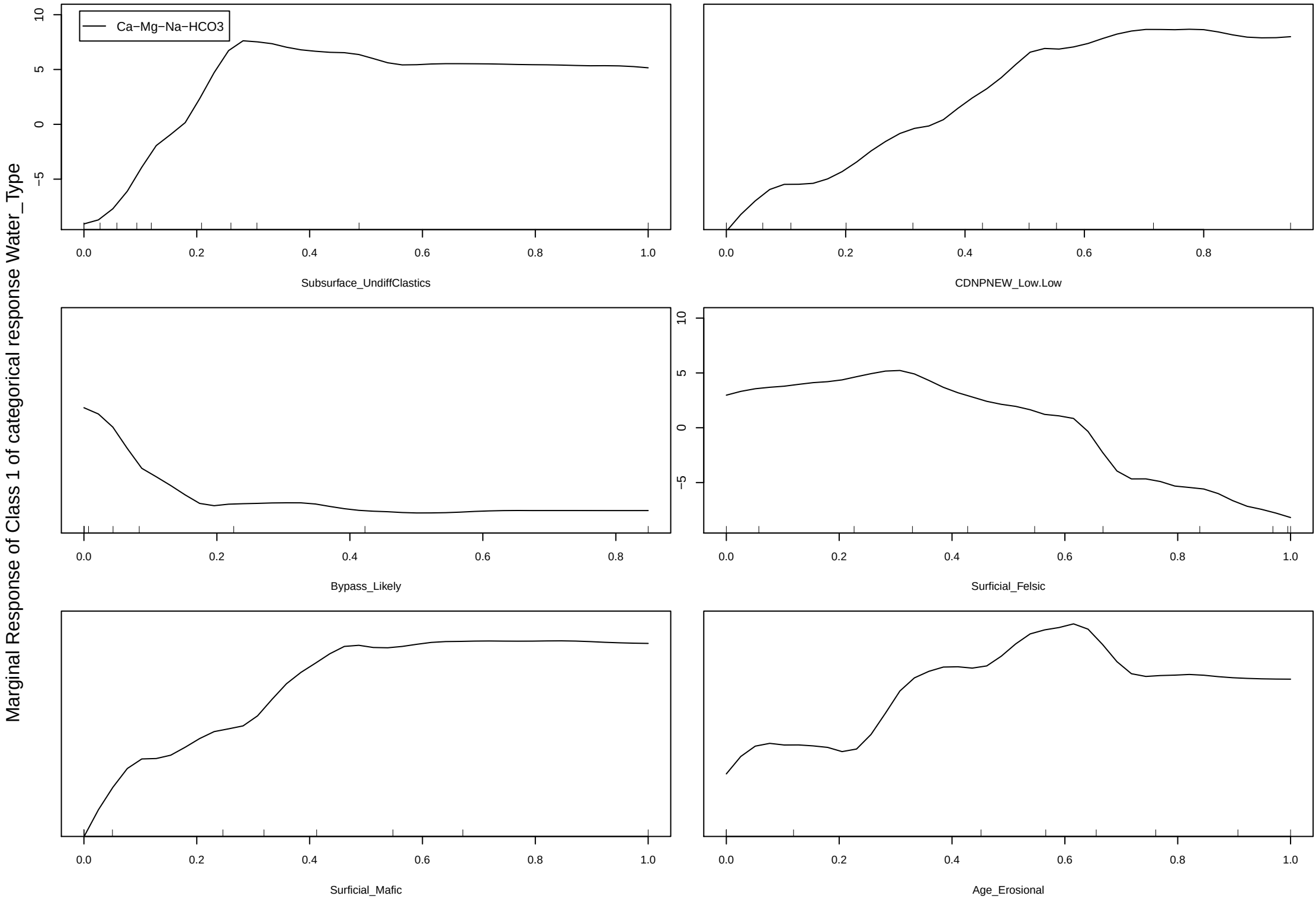

Marginal Response of Class 7 of categorical response HCA_Thresh
2C0
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

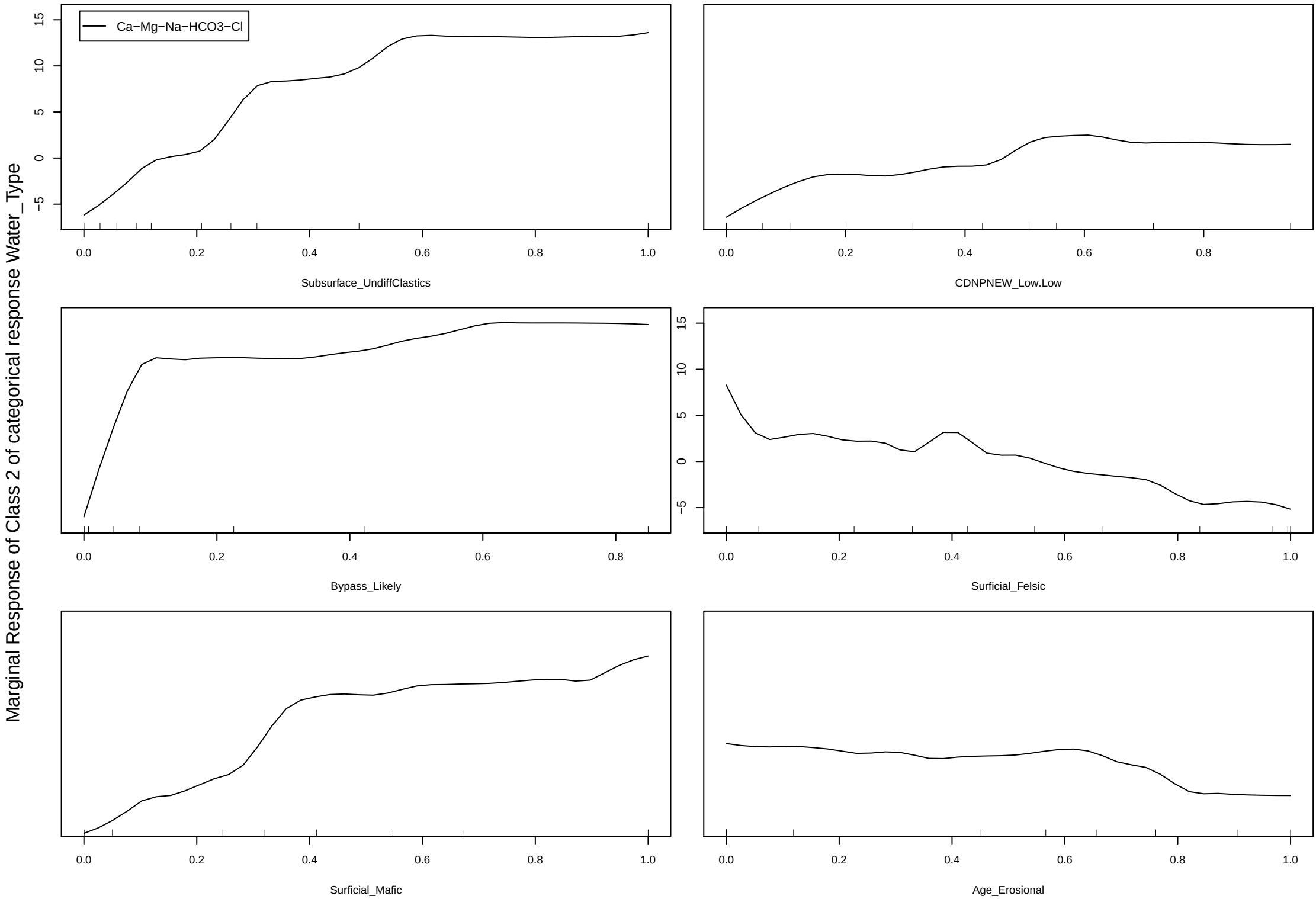

Marginal Response of Class 8 of categorical response HCA_Thresh
3B2
Recharge_LSR
Surficial_Mafic
Recharge2_Lowland
CDNPNEW_High.Low
Bypass_Likely
CDNPNEW_Moderate.Low

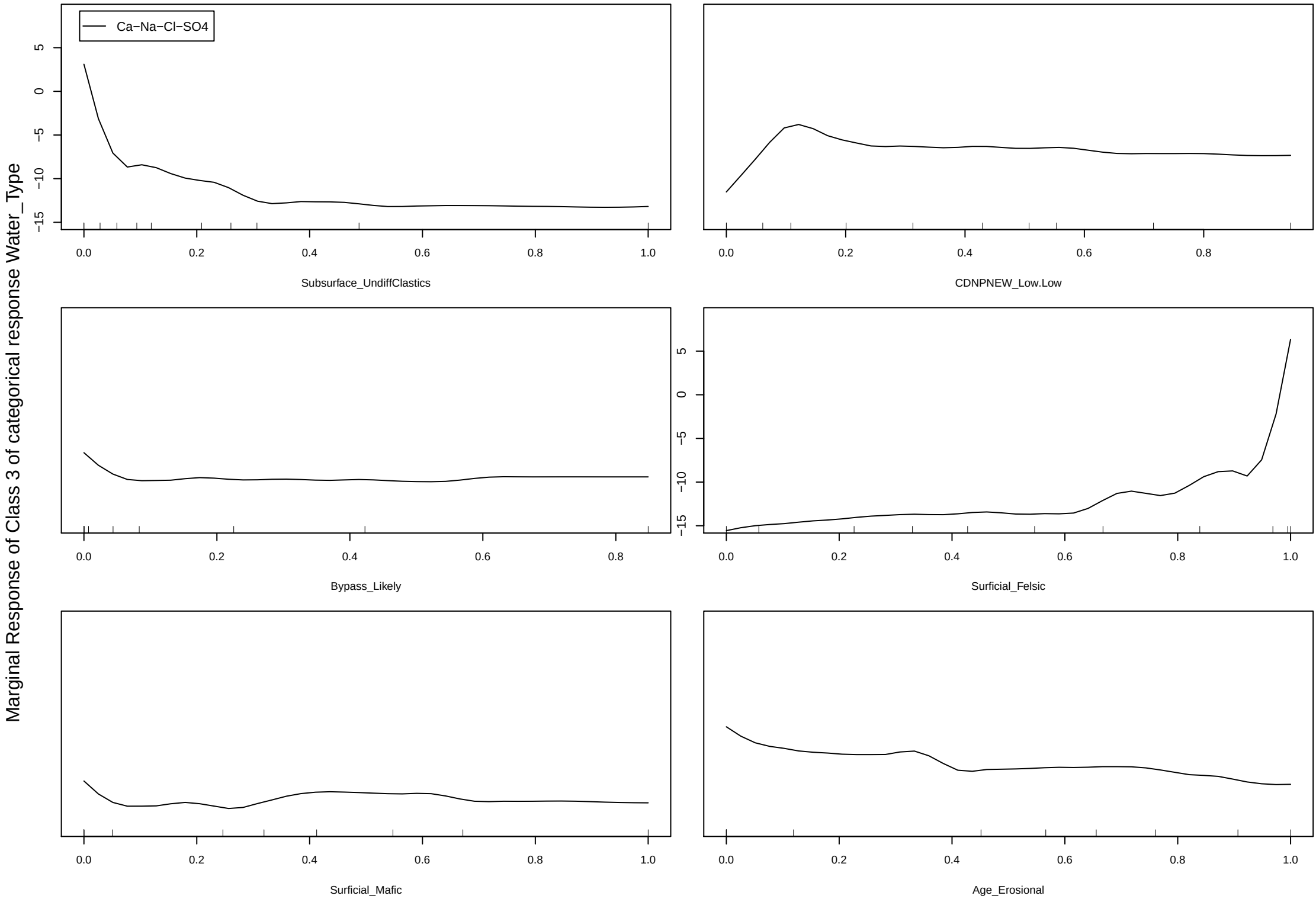

Marginal Response of Class 1 of categorical response Redox_Proc
O2
−1.0
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_UndiffClastics
0.00
0.05
0.10
0.15
Recharge_Mixed
0.00
0.05
0.10
0.15
0.20
0.25
Recharge2_Bedrock1
−1.0
0.0
0.5
1.0
1.5
2.0
2.5
0.0
0.2
0.4
0.6
0.8
1.0
CDNPNEW_High.Low
0.0
0.2
0.4
0.6
0.8
1.0
Surficial_Mafic
0.00
0.05
0.10
0.15
Recharge2_Mixed

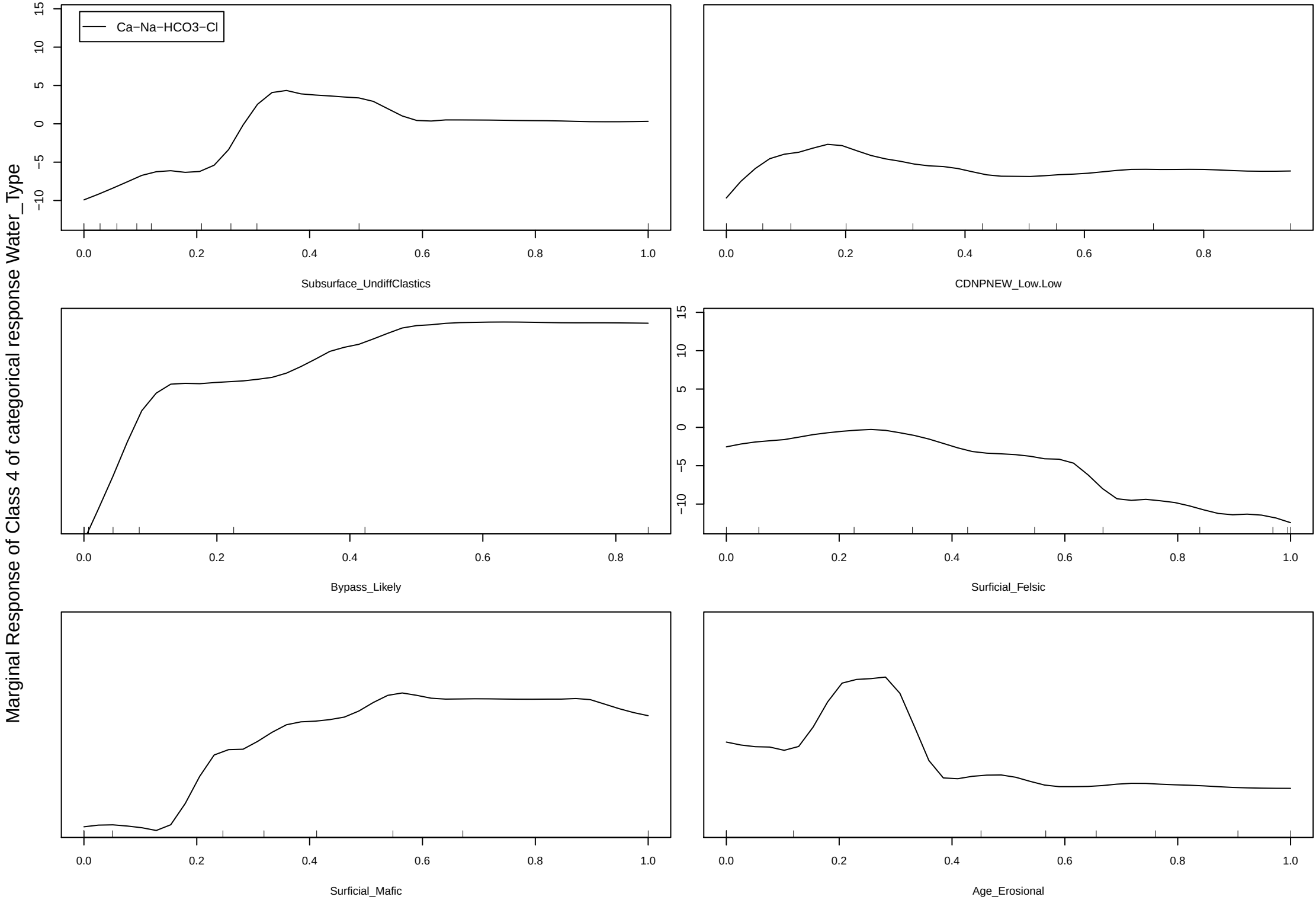

Marginal Response of Class 2 of categorical response Redox_Proc
O2−Fe(III)
0.0
0.5
1.0
1.5
2.0
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_UndiffClastics
0.00
0.05
0.10
0.15
Recharge_Mixed
0.00
0.05
0.10
0.15
0.20
0.25
Recharge2_Bedrock1
0.0
0.5
1.0
1.5
2.0
0.0
0.2
0.4
0.6
0.8
1.0
CDNPNEW_High.Low
0.0
0.2
0.4
0.6
0.8
1.0
Surficial_Mafic
0.00
0.05
0.10
0.15
Recharge2_Mixed

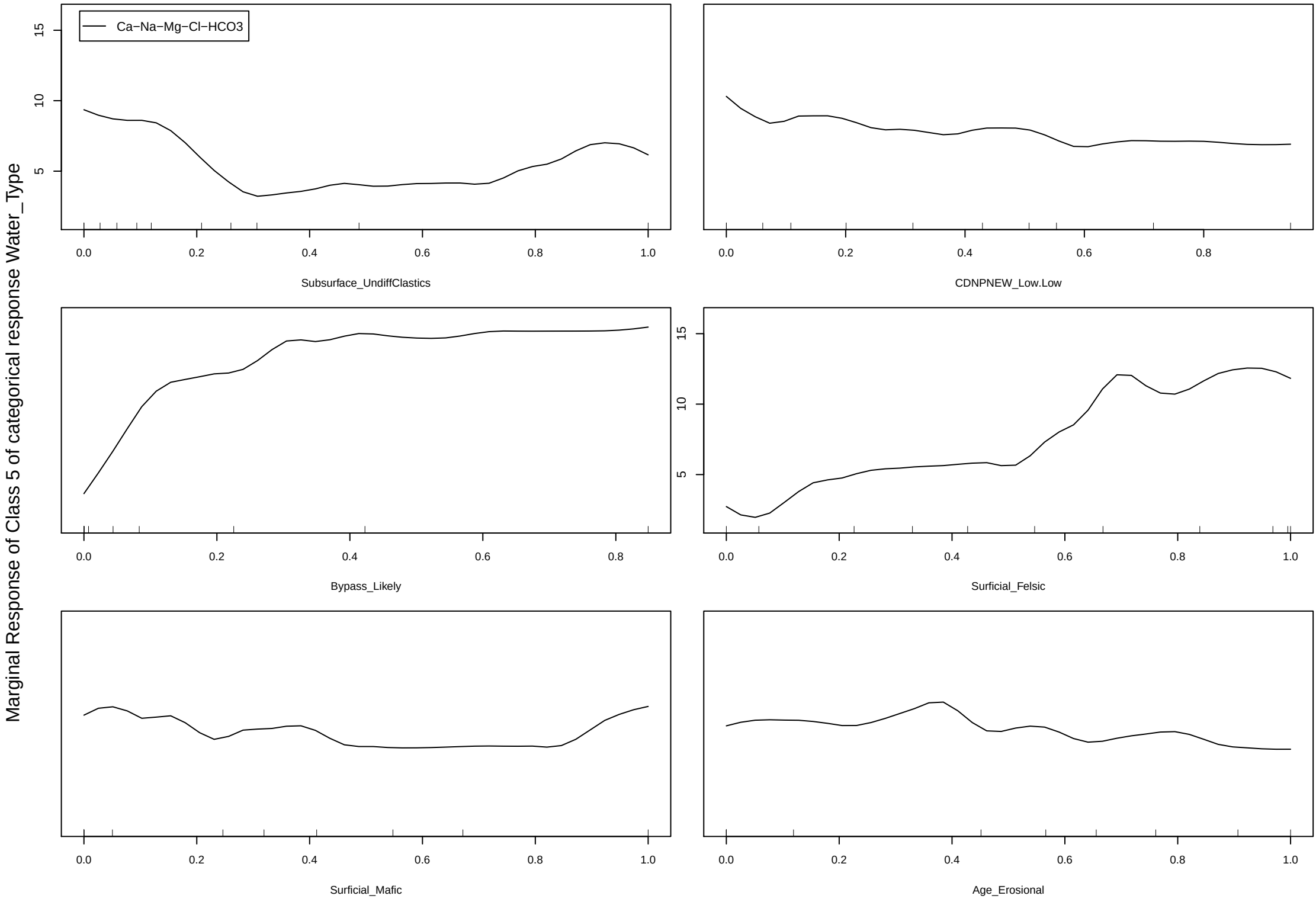

Marginal Response of Class 3 of categorical response Redox_Proc
O2−Fe(III)−SO4
0.0
−1.0
−2.0
−3.0
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_UndiffClastics
0.00
0.05
0.10
0.15
Recharge_Mixed
0.00
0.05
0.10
0.15
0.20
0.25
Recharge2_Bedrock1
0.0
0.2
0.4
0.6
0.8
1.0
CDNPNEW_High.Low
0.0
0.2
0.4
0.6
0.8
1.0
Surficial_Mafic
0.00
0.05
0.10
0.15
Recharge2_Mixed

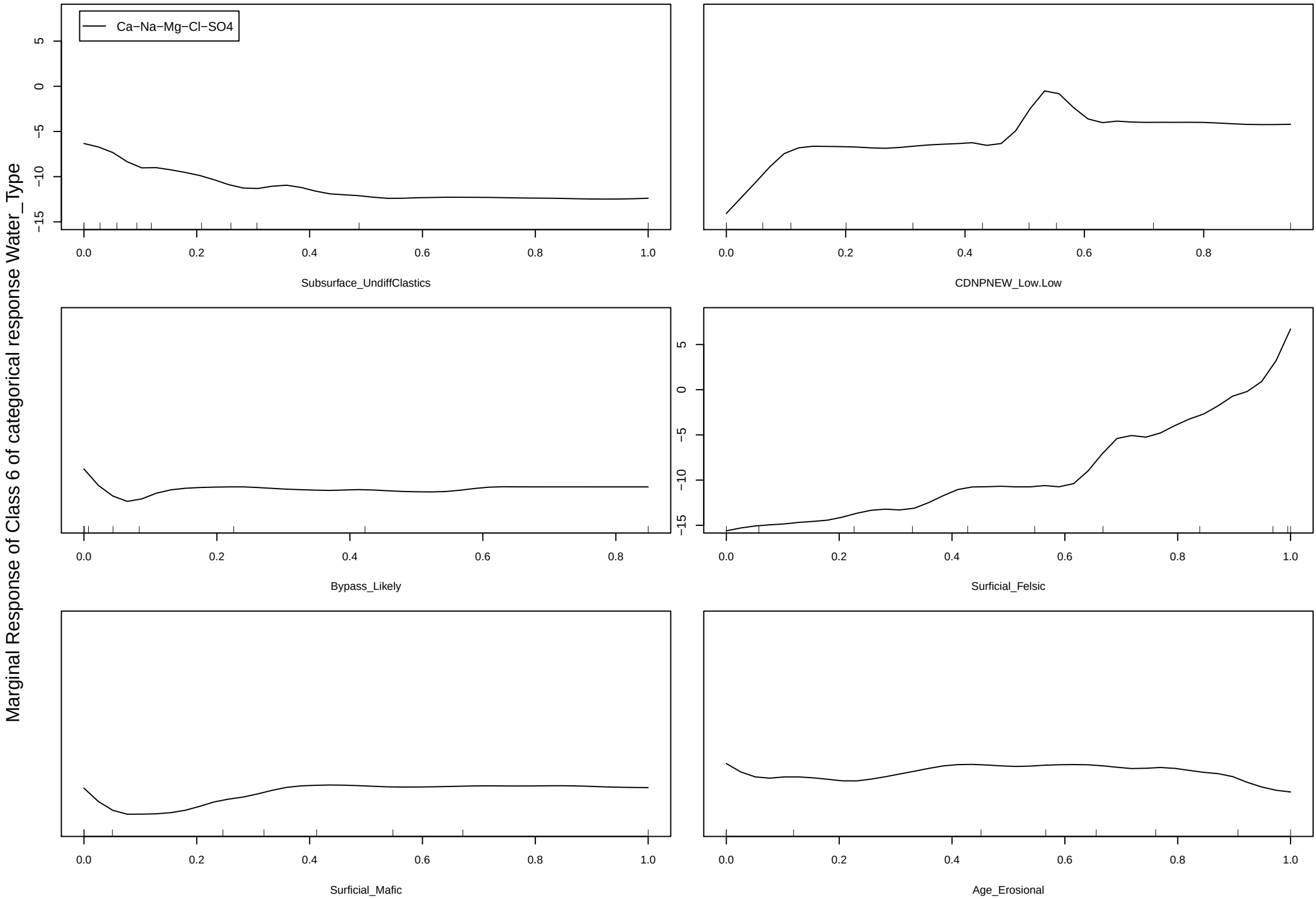

Marginal Response of Class 1 of categorical response Water_Type
Ca−Mg−Na−HCO3
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

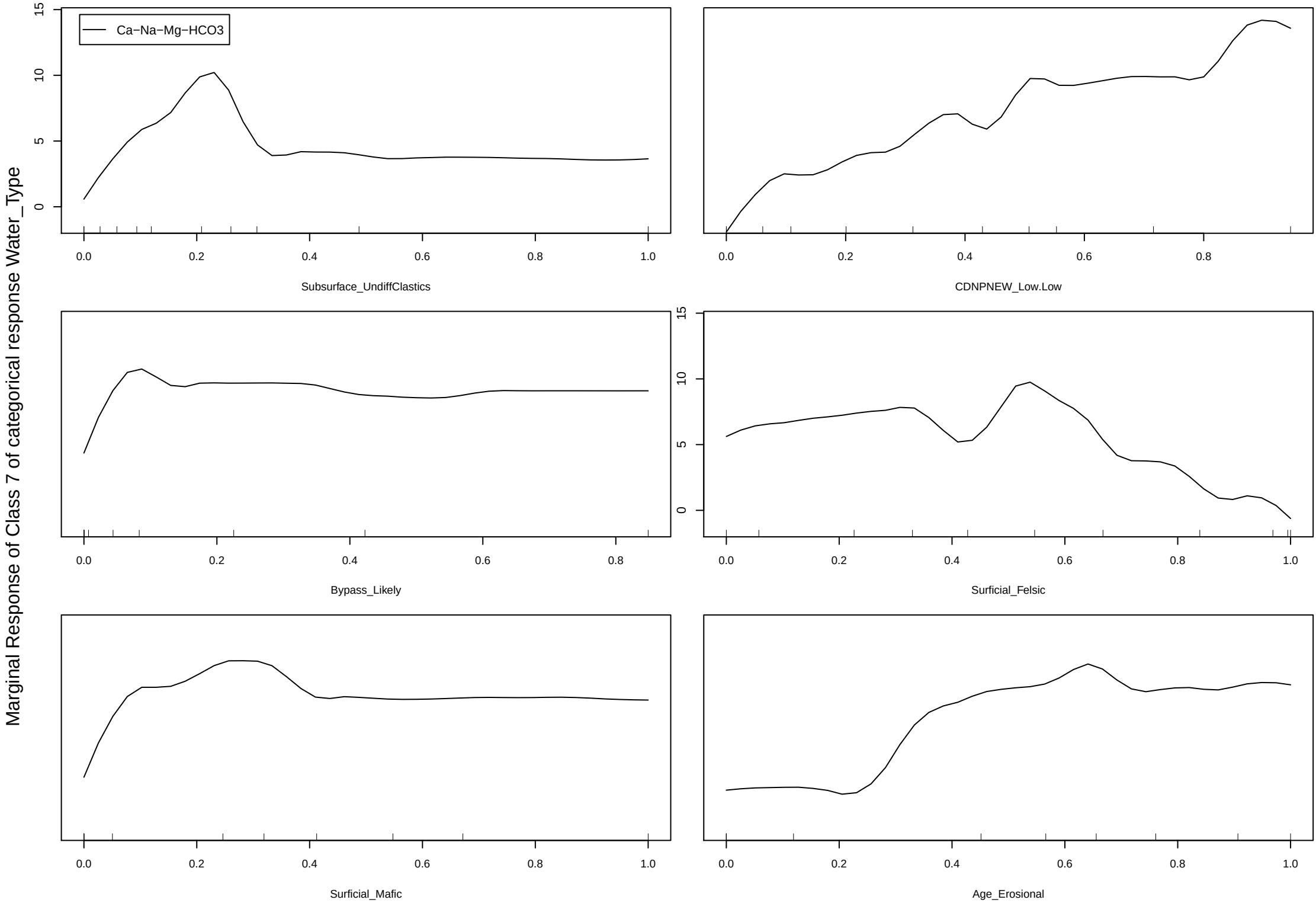

Marginal Response of Class 2 of categorical response Water_Type
Ca−Mg−Na−HCO3−Cl
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

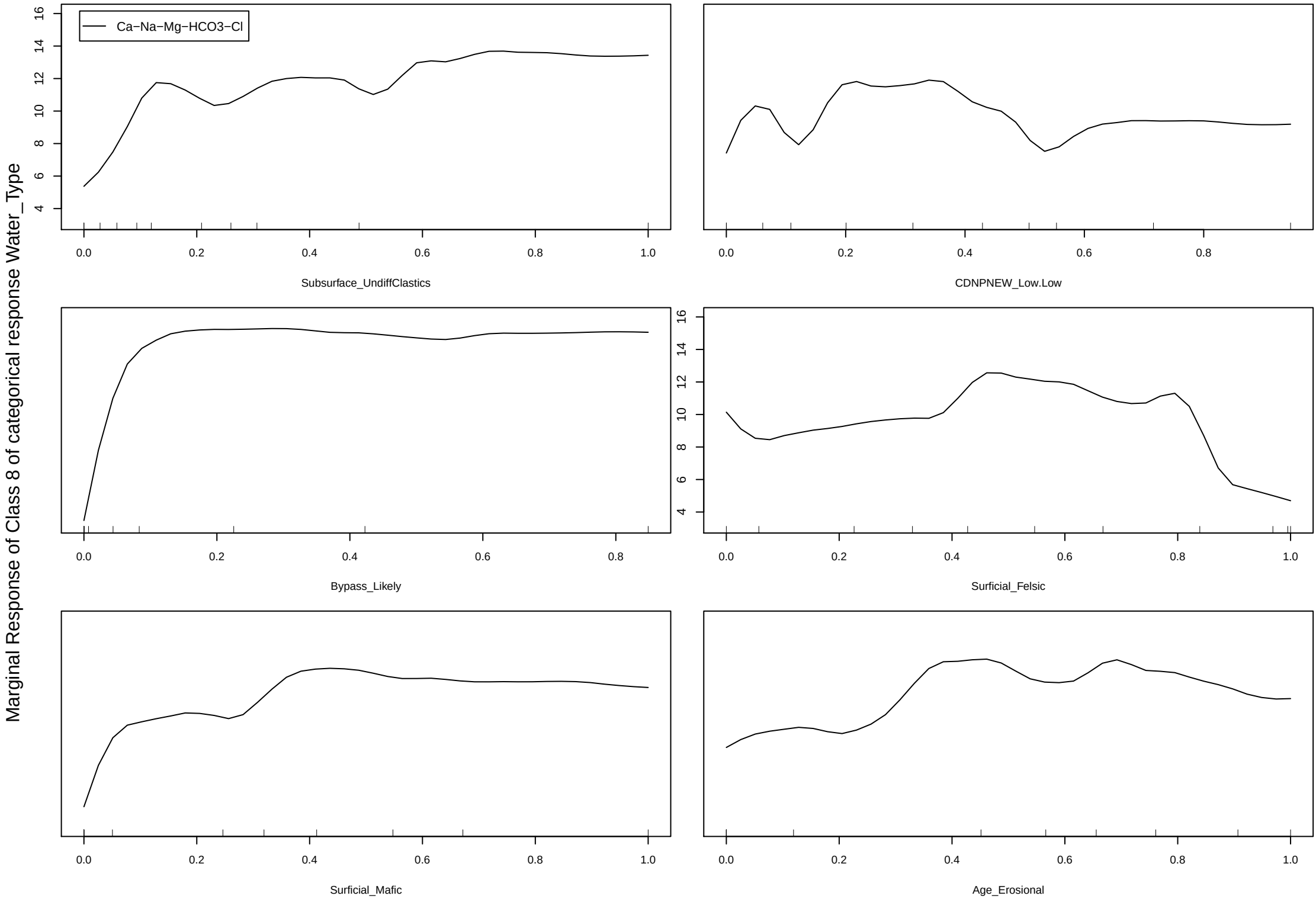

Marginal Response of Class 3 of categorical response Water_Type
Ca−Na−Cl−SO4
5
0
−5
−10
−15
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

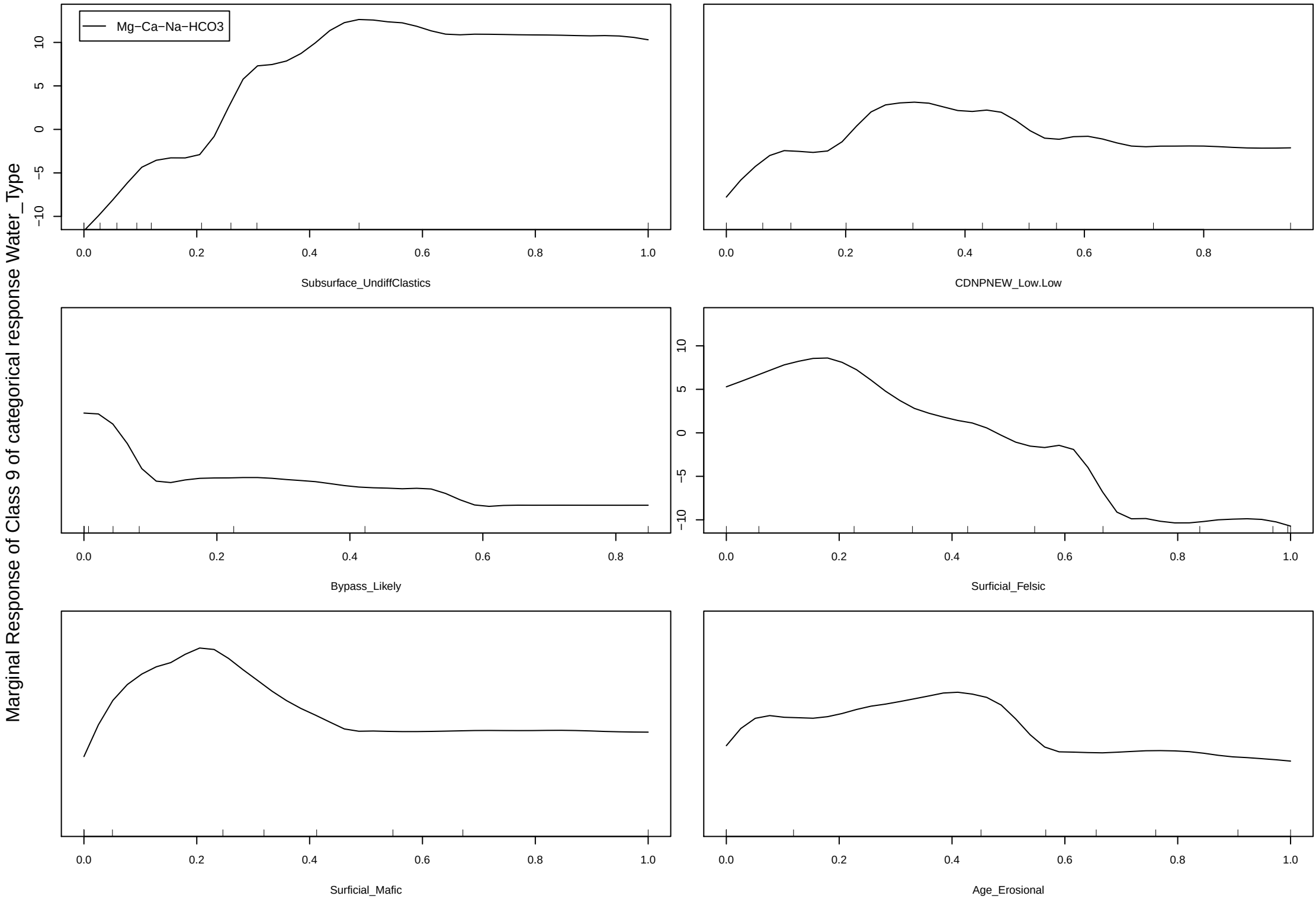

Marginal Response of Class 4 of categorical response Water_Type
Ca−Na−HCO3−Cl
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

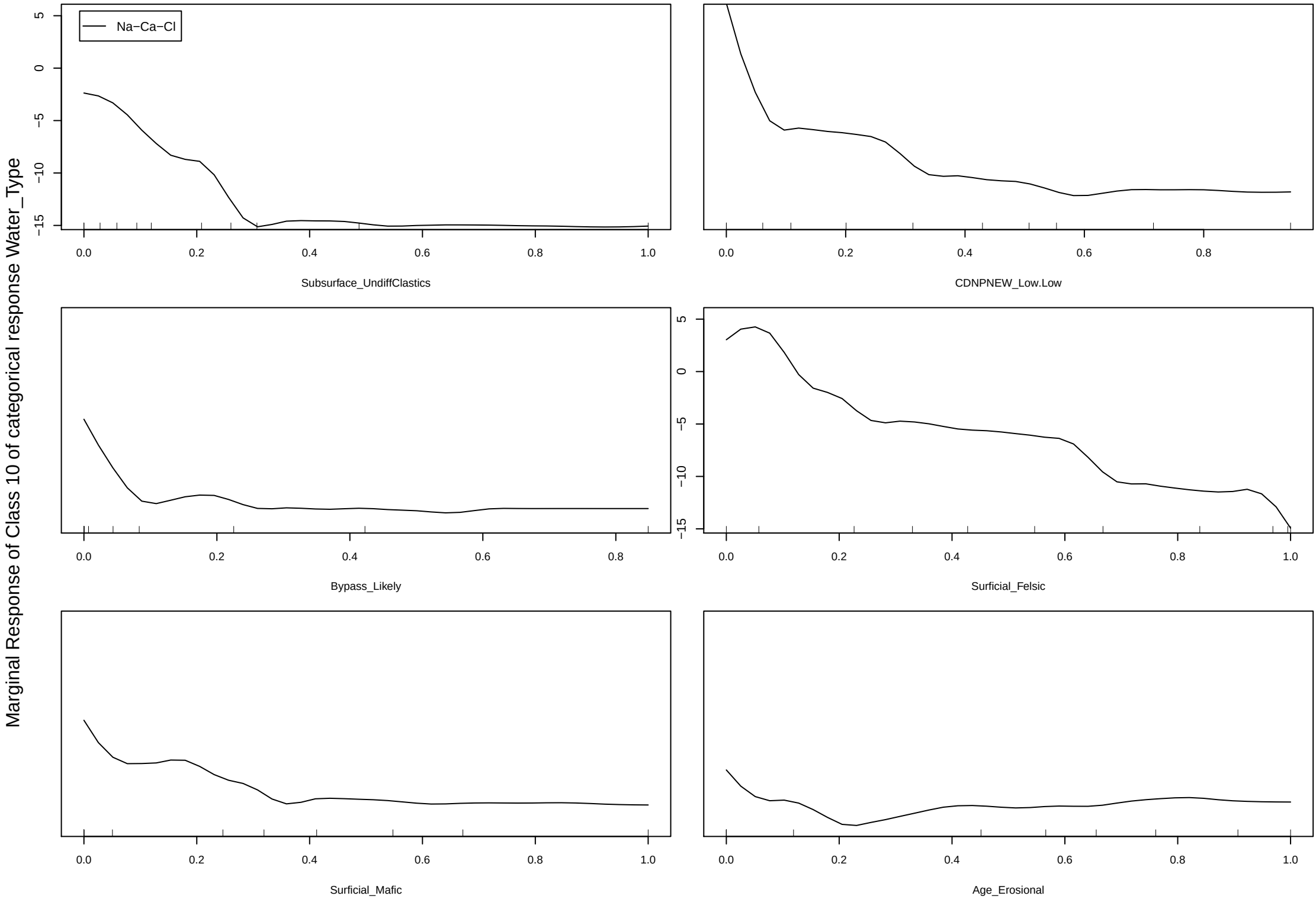

Ca−Na−Mg−Cl−HCO3
Marginal Response of Class 5 of categorical response Water_Type
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

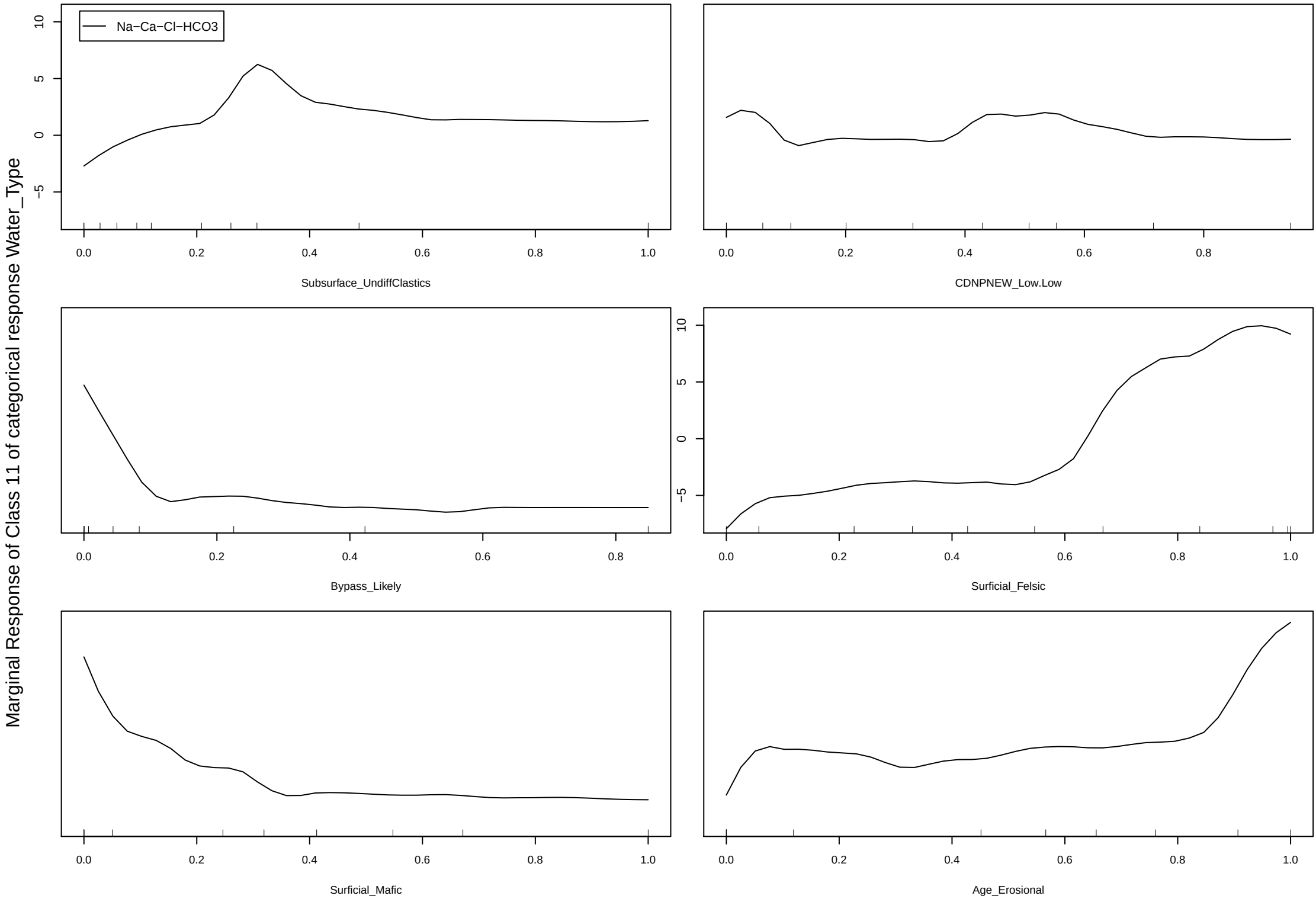

Marginal Response of Class 6 of categorical response Water_Type
Ca−Na−Mg−Cl−SO4
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

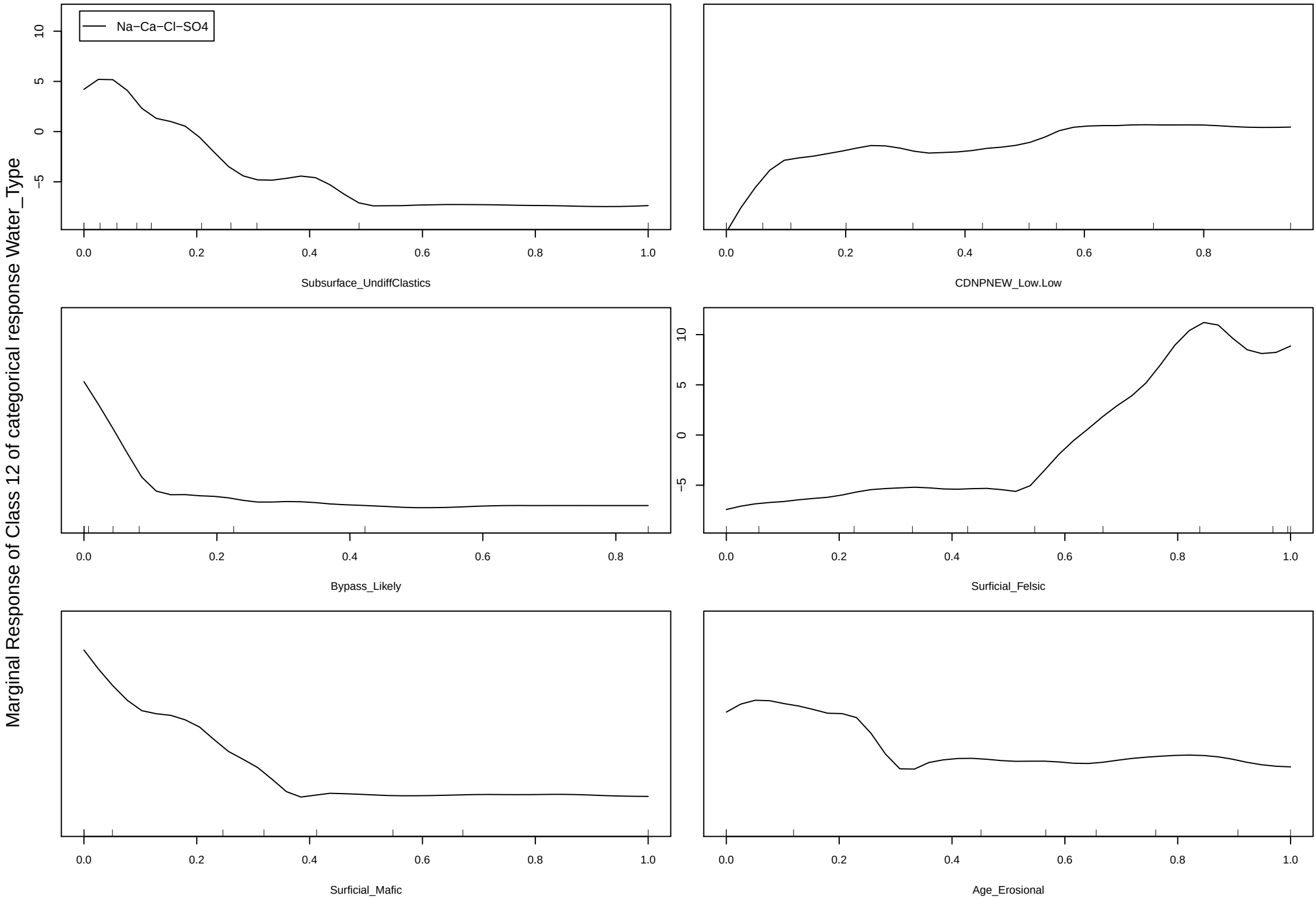

Marginal Response of Class 7 of categorical response Water_Type
Ca−Na−Mg−HCO3
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

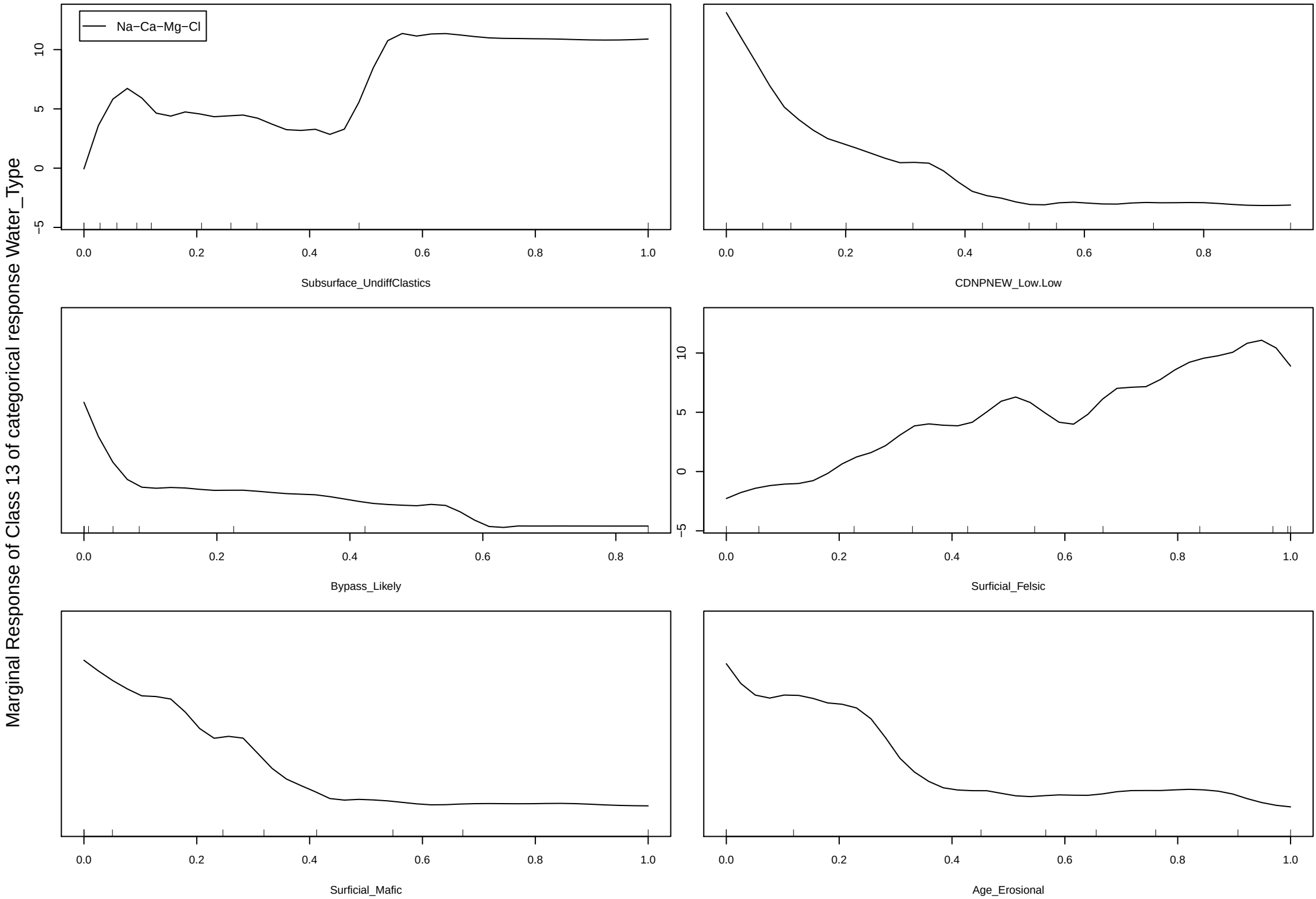

Marginal Response of Class 8 of categorical response Water_Type
Ca−Na−Mg−HCO3−Cl
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

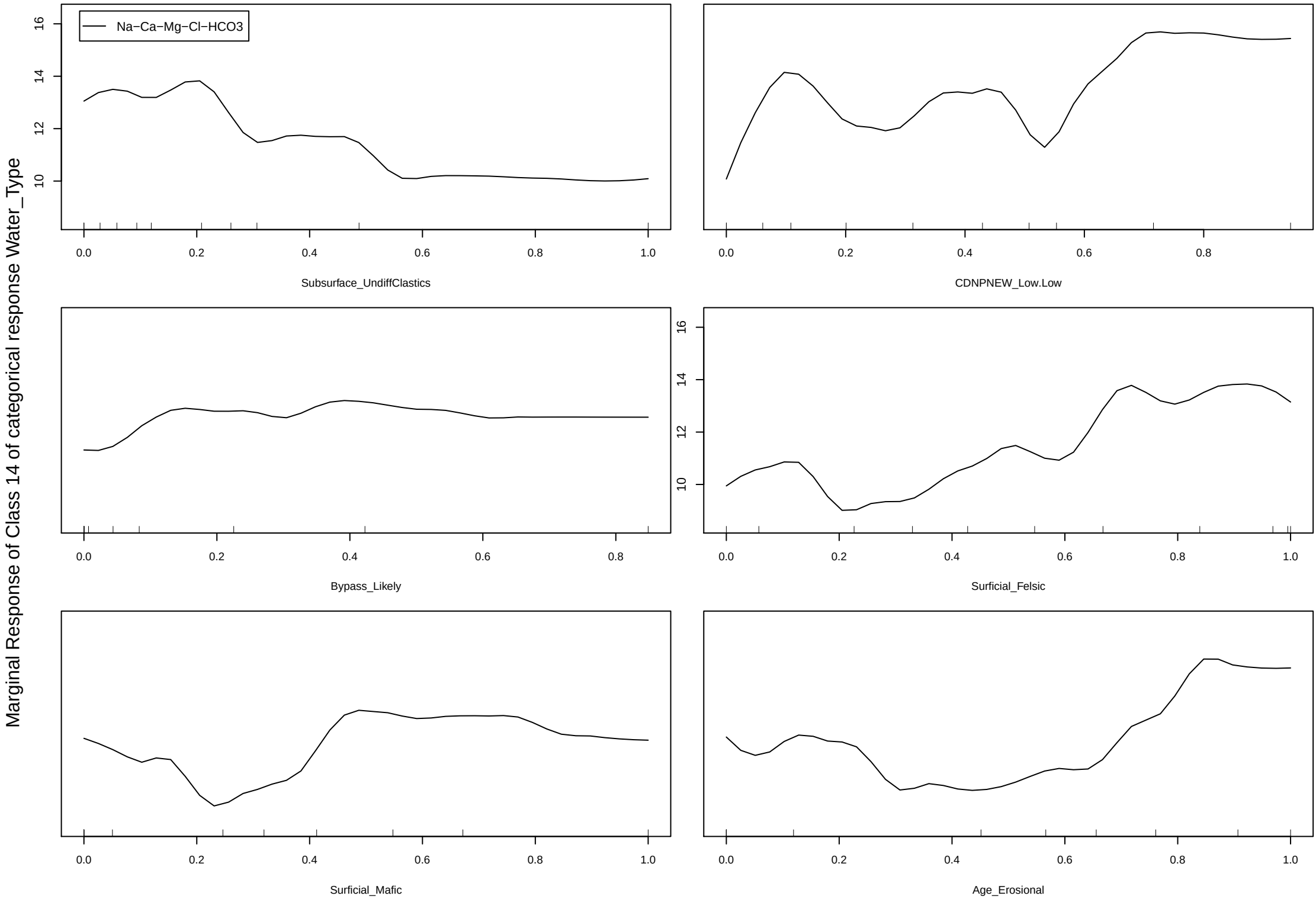

Marginal Response of Class 9 of categorical response Water_Type
Mg−Ca−Na−HCO3
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

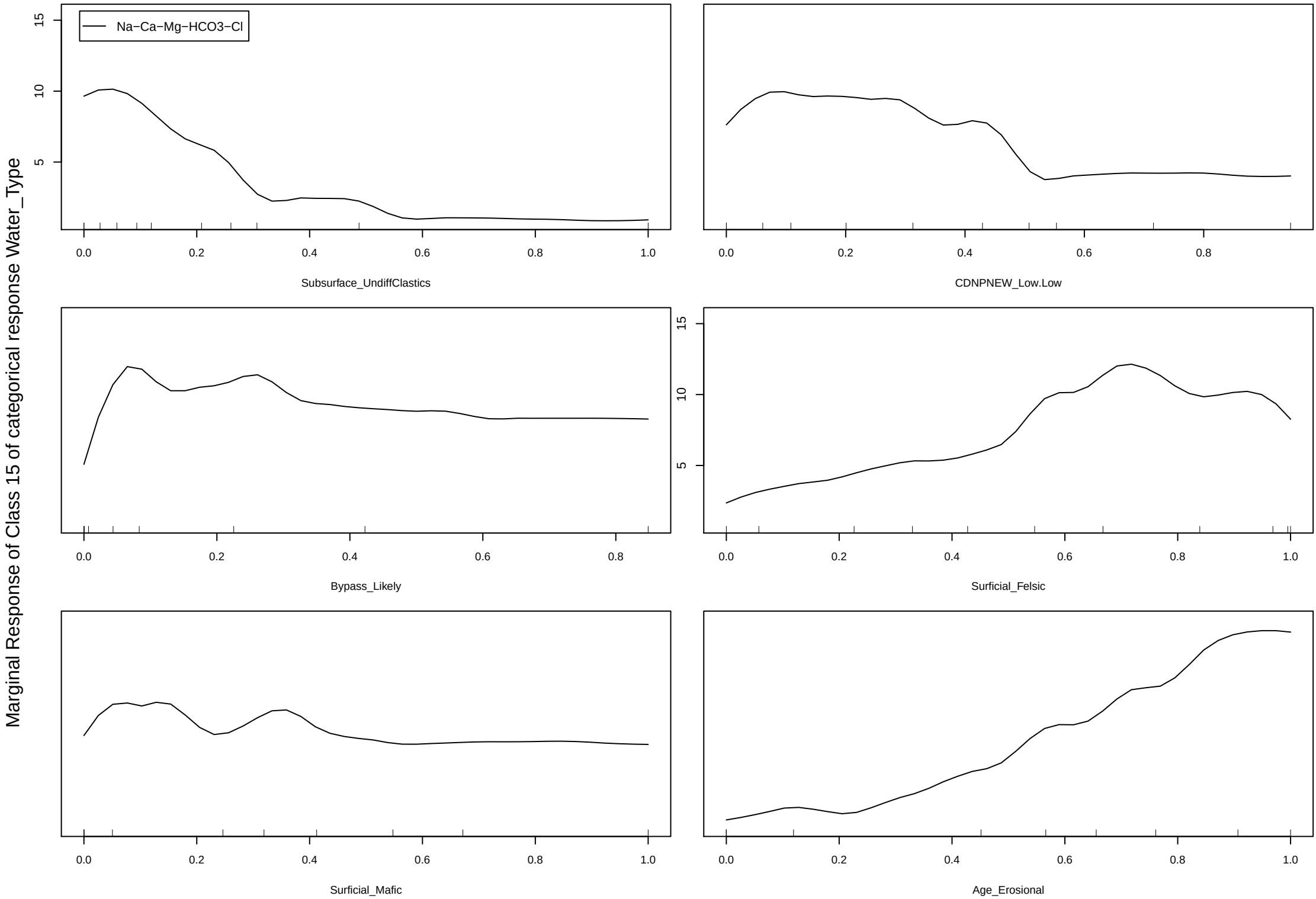

Marginal Response of Class 10 of categorical response Water_Type
Na−Ca−Cl
5
0
−5
−10
−15
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

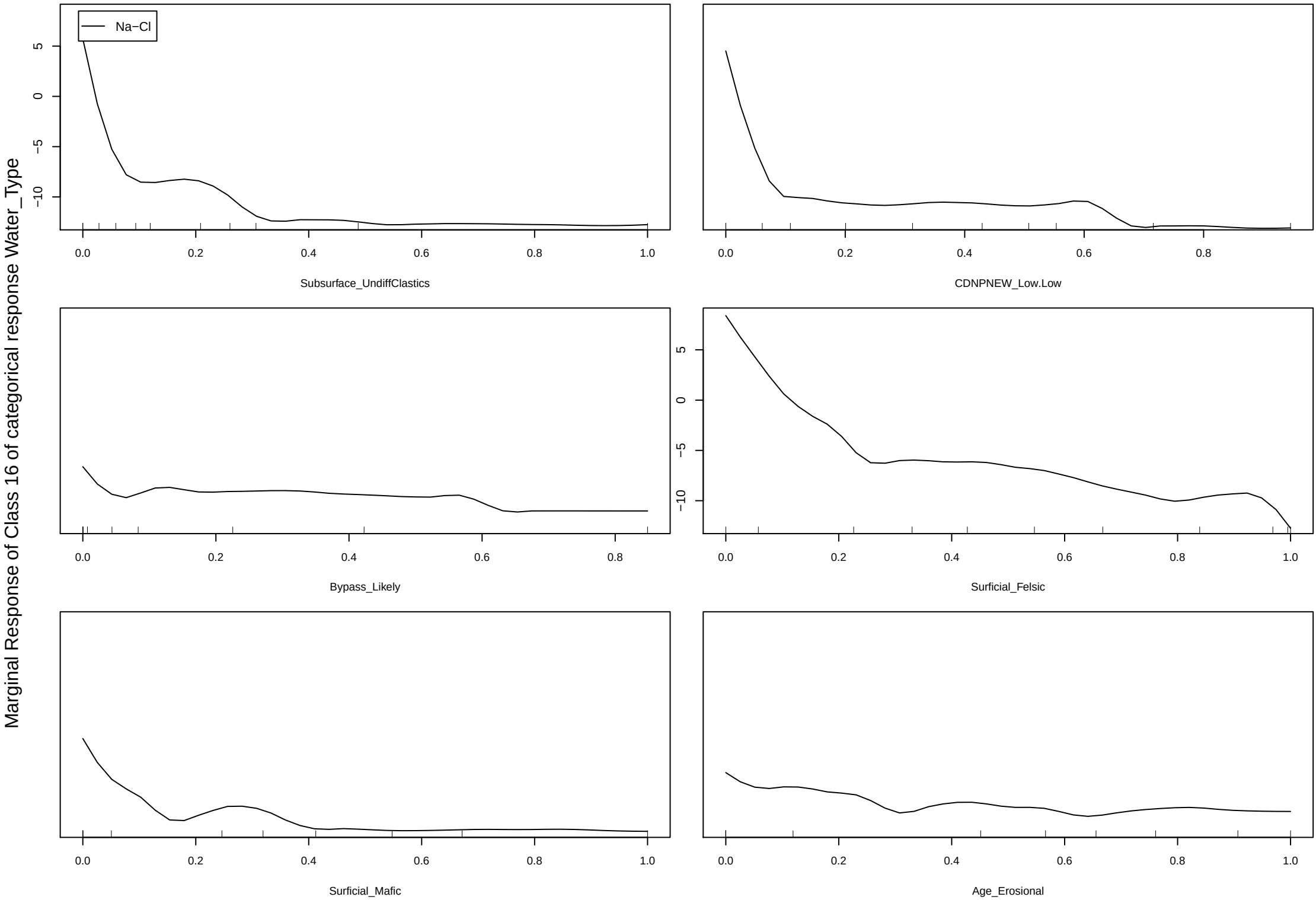

Marginal Response of Class 11 of categorical response Water_Type
Na−Ca−Cl−HCO3
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

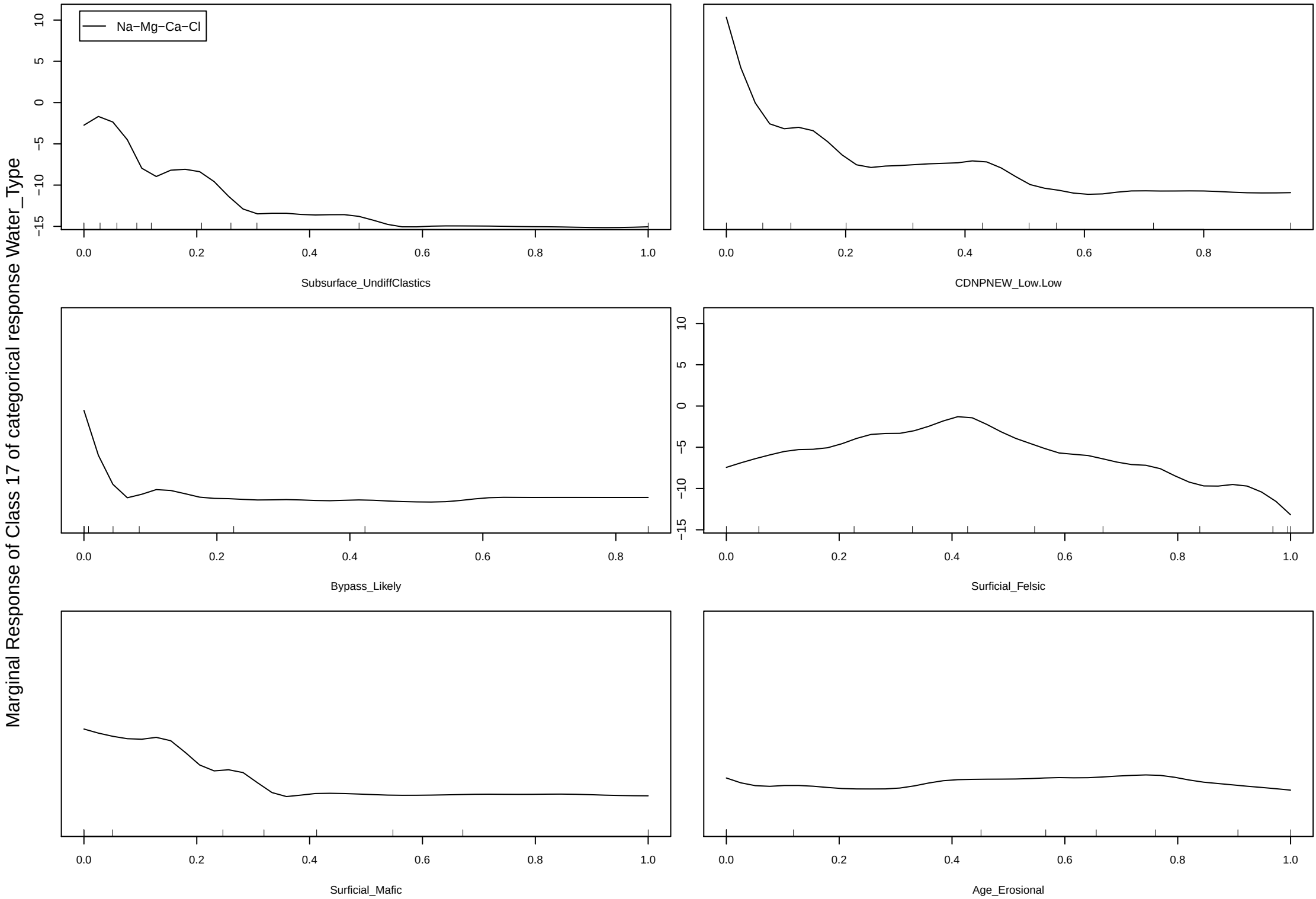

Marginal Response of Class 12 of categorical response Water_Type
Na−Ca−Cl−SO4
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

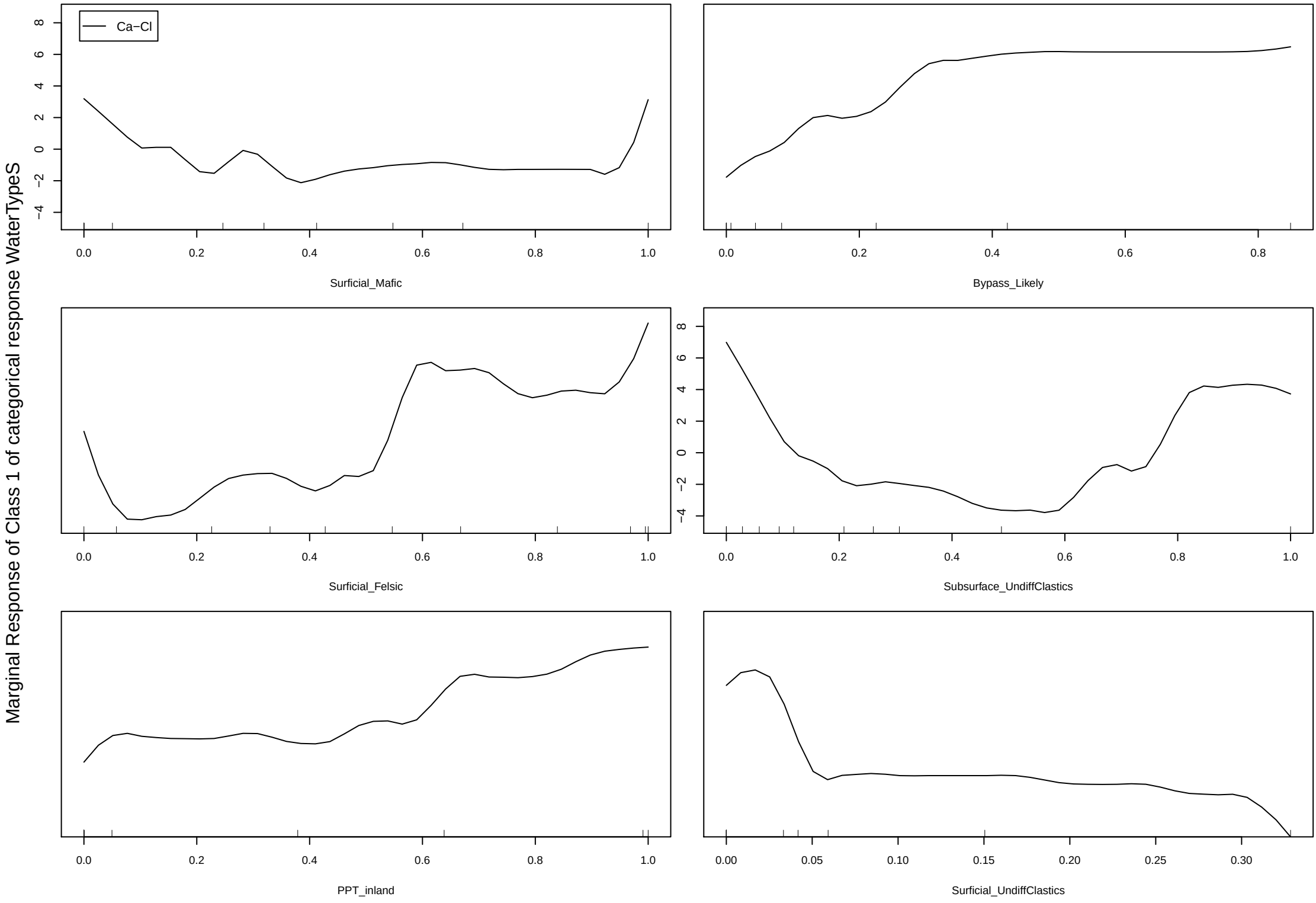

Marginal Response of Class 13 of categorical response Water_Type
Na−Ca−Mg−Cl
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

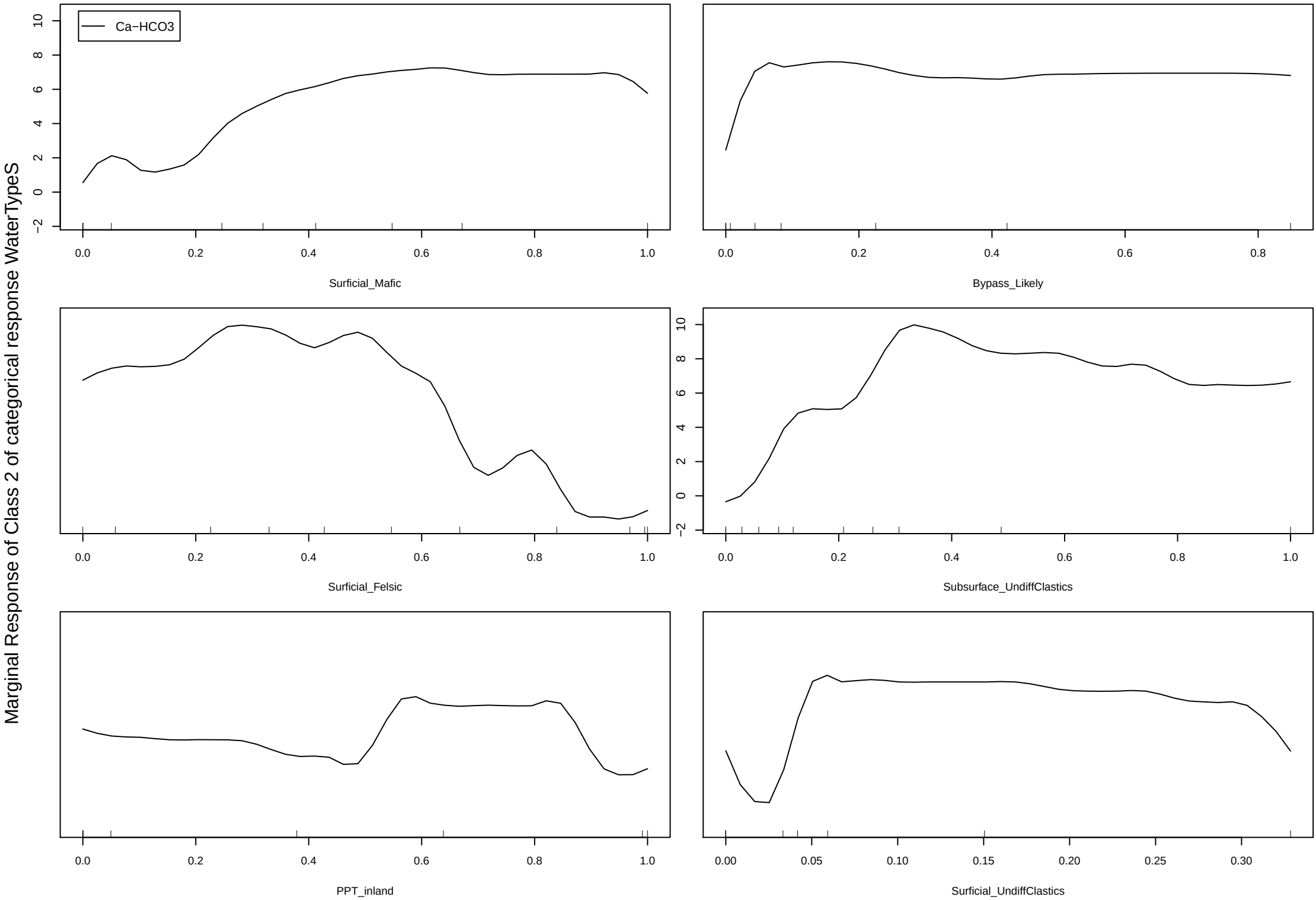

Marginal Response of Class 14 of categorical response Water_Type
Na−Ca−Mg−Cl−HCO3
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

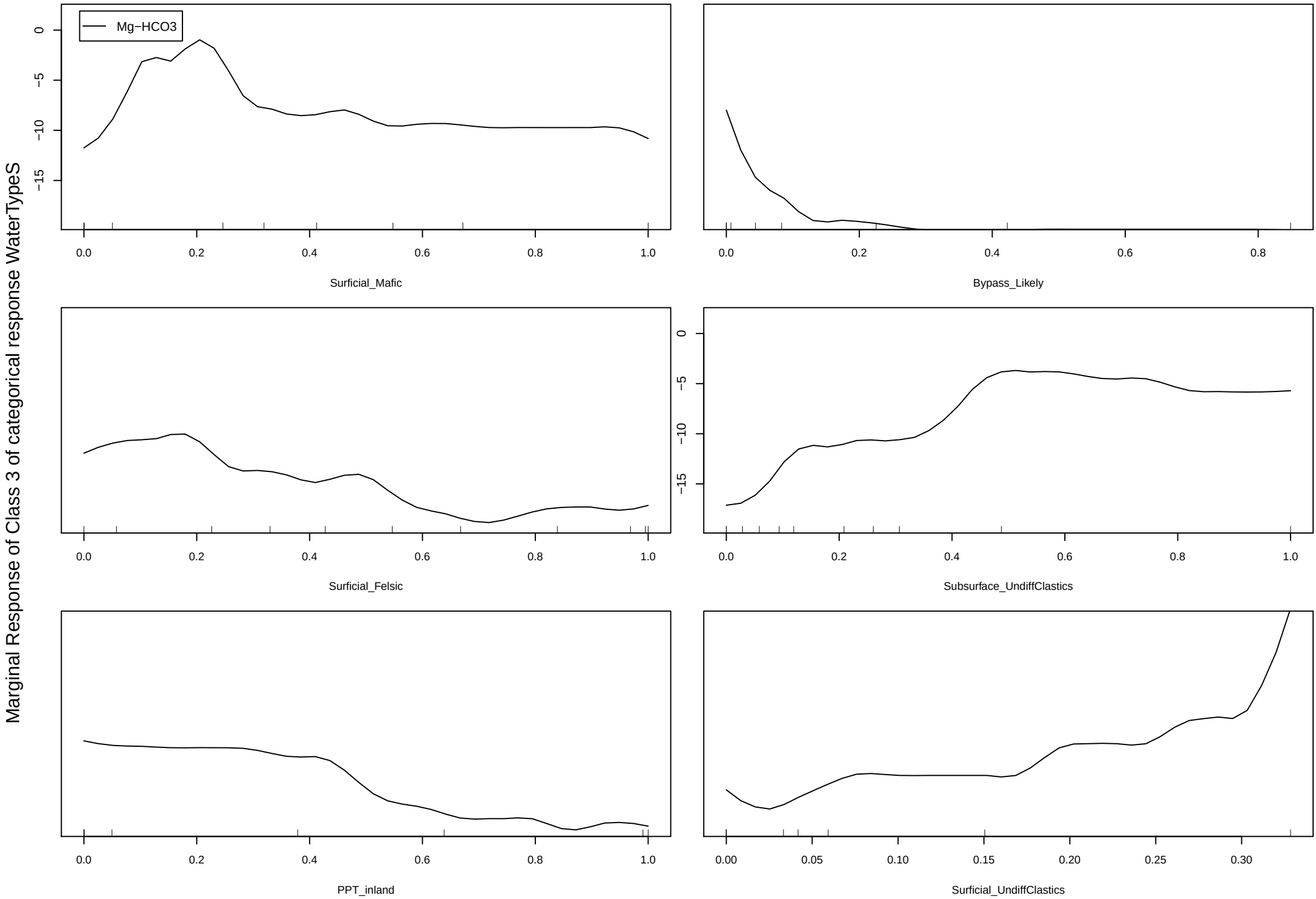

Marginal Response of Class 15 of categorical response Water_Type
Na−Ca−Mg−HCO3−Cl
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

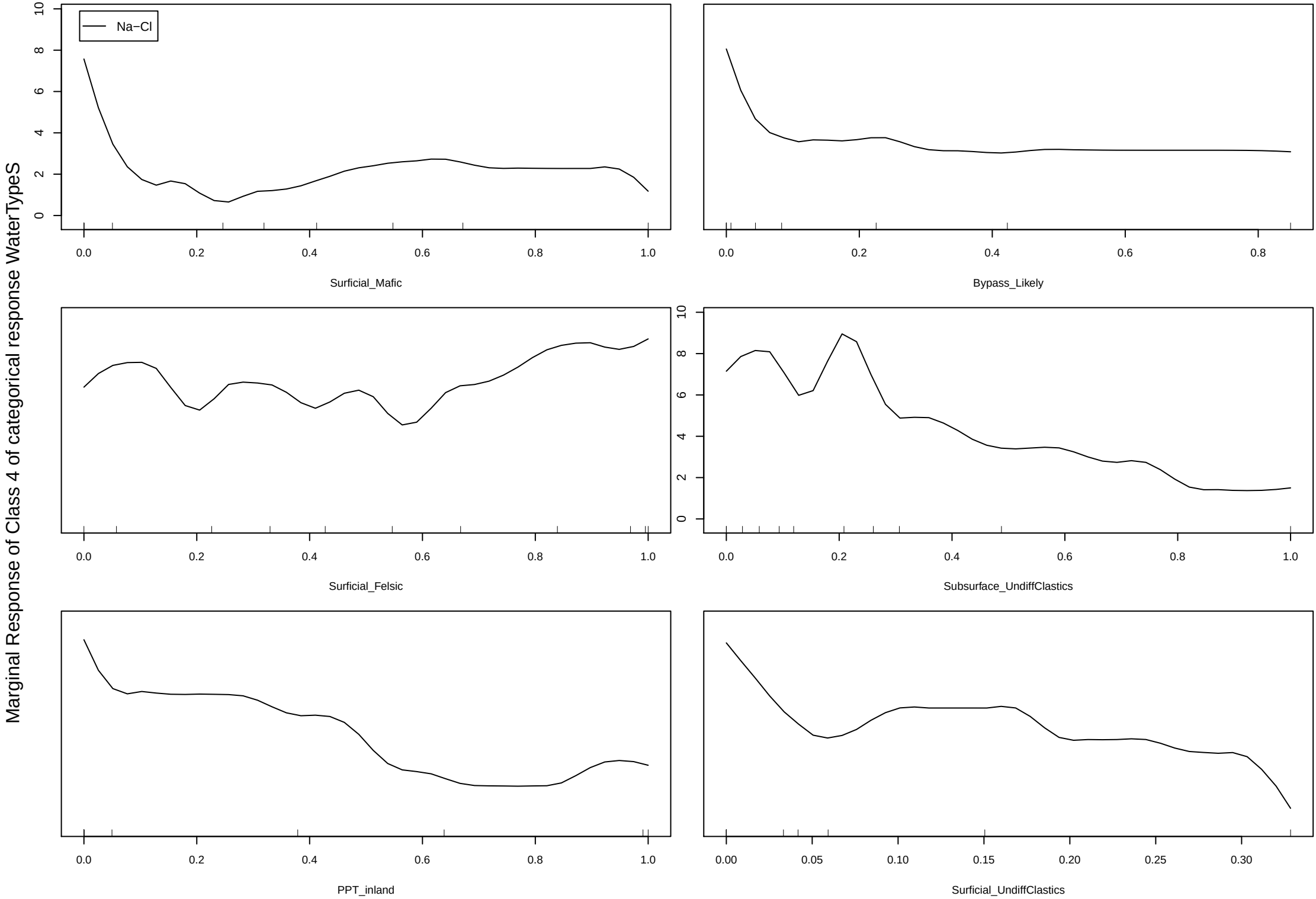

Marginal Response of Class 16 of categorical response Water_Type
Na−Cl
5
0
−5
−10
0.0
0.2
0.4
0.6
0.8
1.0
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional

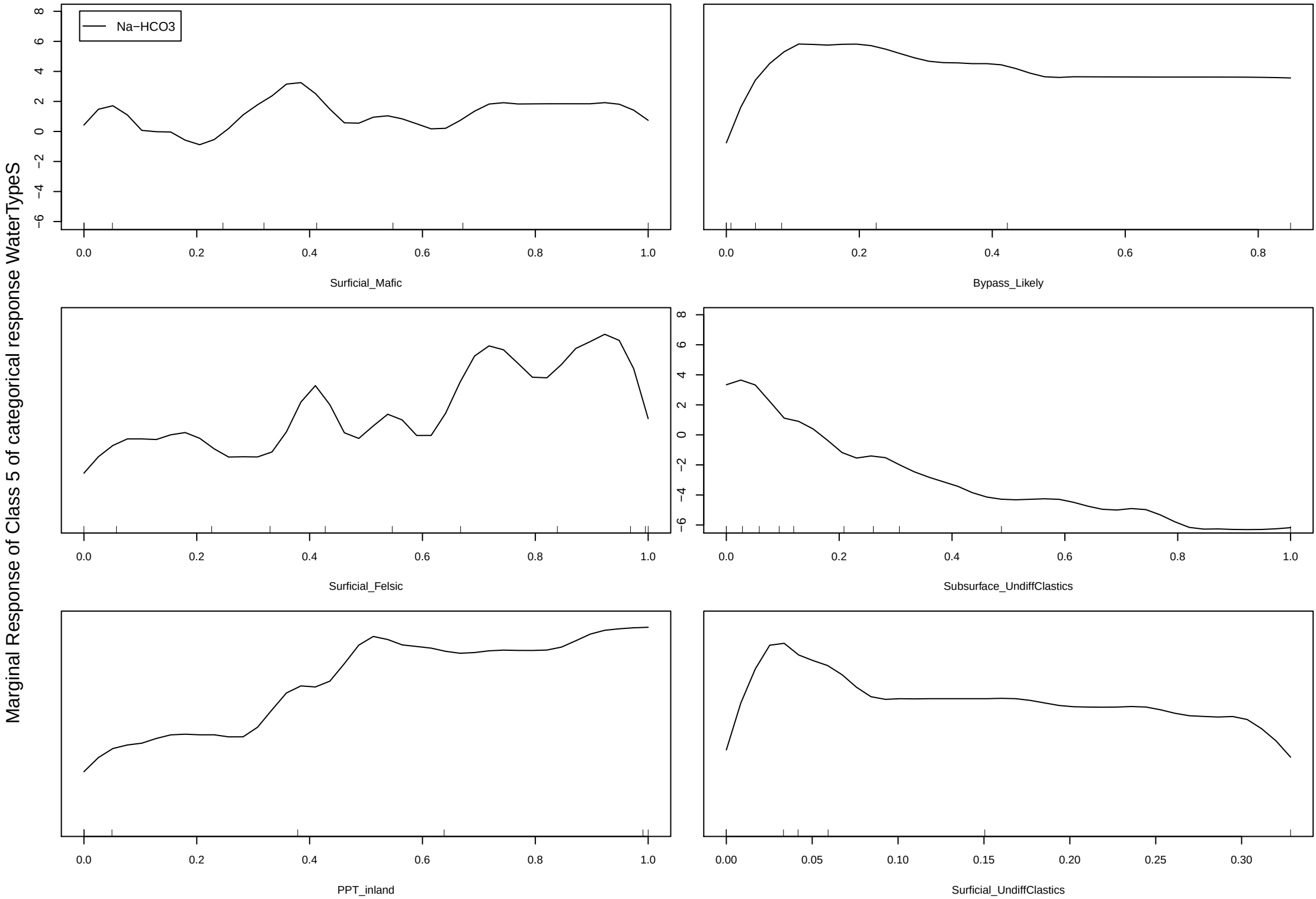

Marginal Response of Class 17 of categorical response Water_Type
Na−Mg−Ca−Cl
Subsurface_UndiffClastics
CDNPNEW_Low.Low
Bypass_Likely
Surficial_Felsic
Surficial_Mafic
Age_Erosional


Marginal Response of Class 1 of categorical response WaterTypeS
Ca−Cl
Surficial_Mafic
Bypass_Likely
Surficial_Felsic
Subsurface_UndiffClastics
PPT_inland
Surficial_UndiffClastics


Marginal Response of Class 2 of categorical response WaterTypeS
Ca−HCO3
Surficial_Mafic
Bypass_Likely
Surficial_Felsic
Subsurface_UndiffClastics
PPT_inland
Surficial_UndiffClastics


Marginal Response of Class 3 of categorical response WaterTypeS
Mg−HCO3
0
−5
−10
−15
0.0
0.2
0.4
0.6
0.8
1.0
Surficial_Mafic
Bypass_Likely
Surficial_Felsic
Subsurface_UndiffClastics
PPT_inland
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Surficial_UndiffClastics


Marginal Response of Class 4 of categorical response WaterTypeS
Na−Cl
0
2
4
6
8
10
0.0
0.2
0.4
0.6
0.8
1.0
Surficial_Mafic
Bypass_Likely
Surficial_Felsic
Subsurface_UndiffClastics
PPT_inland
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Surficial_UndiffClastics


Marginal Response of Class 5 of categorical response WaterTypeS
Na−HCO3
Surficial_Mafic
Bypass_Likely
Surficial_Felsic
Subsurface_UndiffClastics
PPT_inland
Surficial_UndiffClastics

References

Andersen, H.E., 2003. Hydrology, nutrient processes and vegetation in floodplain wetlands. PhD thesis. National Environmental Research Institute, Denmark. http://afhandlinger.dmu.dk

Aislabie, J., Smith, J.J., Fraser, R., McLeod, M., 2001. Leaching of bacterial indicators of faecal contamination through four New Zealand soils. *Aust. J. Soil Res, 39*: 1397-1406.

Bernal, B., 2008. Carbon pools and profiles in wetland soils: The effect of climate and wetland type. The Ohio State University.

Blasch, K.W., Bryson, J.R., 2007. Distinguishing Sources of Ground Water Recharge by Using δ2H and δ18O. *Ground Water 45*, 294-308.

Blaschke, P.M., N.A. Trustrum., R.C. DeRose., 1992. Ecosystem processes and sustainable land use in New Zealand steeplands. *Agriculture, Ecosystems and Environment 41*:153-178.

Bona, F., Falasco, E., Fassina, S., Griselli, B., Badino, G., 2007. Characterization of diatom assemblages in mid-altitude streams of NW Italy. *Hydrobiologia, 583*(1), 265-274.

Borowiak, D., Polkowaska, Z., Przyjazny, A., 2006. The hydrochemistry of high-altitude lakes in selected mountain ranges of Central and Southern Europe. *Limnological Review 6*: 21-30.

Breiman, L., 1996. Bagging Predictors. *Machine Learning, 24*(2), 123-140.

Bundi, U., Peter, A., Robinson, C. T., Schädler, B., Truffer, B., Vollenweider, S., Weingartner, R., 2010. Alpine Waters: *Springer Berlin Heidelberg*.

Burbery, L.; Moore, C., Dumbleton, B., 2013. Towards an Improved Understanding of the Knapdale Aquifer: A Modelling Study of Anomalous Nitrate Levels in the Knapdale Groundwater Zone, report prepared for Environment Southland.

Buscema, M., Mazzetti di Pietralata, M., Salvemini, V., Intraligi, M., Indrimi, M., 1998. Application of artificial neural networks to eating disorders. *Substance Use & Misuse, 33*: 765-792.

Chapelle, F.H., Bradley, P.M., Thomas, M.A. and McMahon, P.B., 2009. Distinguishing iron-reducing from sulfate-reducing conditions. *Ground Water 47*: 300-305.

Clark, I.D., Fritz, P., 1997. Environmental Isotopes in Hydrogeology. *Taylor & Francis*.

Cloutier, V., Lefebvre, R., Therrien, R., Savard, M. M., 2008. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system, *J. Hydrol., 353*, 294–313.

Collins, M.E., Kuehl, R.J., 2000. Organic matter accumulation and organic soils. Chapter 6. Wetland Soils- Genesis, hydrology, landscapes, and classification.

Dansgaard, W., 1964. Stable isotopes in precipitation. Tellus, 16: 436-438.

Daughney, C.J., Reeves, R.R., 2005. Definition of Hydrochemical Facies in the New Zealand National Groundwater Monitoring Programme. *J. Hydrol. (NZ) 44*: 105-130.

Daughney, C.J., 2005. Spreadsheet for Automatic Processing of Water Quality Data: Theory, Use and Implementation in Excel. GNS Science Report 2005/35.

Daughney, C.J., Morgenstern, U., van der Raaij, R., Reeves, R.R., 2010. Discriminant analysis for estimation of groundwater age from hydrochemistry and well construction: application to New Zealand aquifers. *Hydrogeology J. 18*: 417-428.

Daughney, C., Rissmann, C., Friedel, M., Morgenstern, U., Hodson, R., van der Raaij, R., Rodway, E., Martindale, H., Pearson, L., Townsend, D., Kees, L., Moreau, M., Millar, R., Horton, T., 2015. Hydrochemistry of the Southland Region, *GNS Science Report* 2015/24.

Dickson, B.L., Giblin, A.M., 2007. An evaluation of methods for imputation of missing trace element data in groundwaters. *Geochem. Explor. Environ. Anal. 7*: 173-178.

Dinka, M.O., Loiskandl, W., Ndambuki, J.M., 2015. Hydrochemical characterization of various surface water and groundwater resources available in Matahara areas, Fantalle Woreda of Oromiya region. *Journal of Hydrology: Regional Studies 3*, 444-456.

Drever, J.I., 1997. The geochemistry of natural waters, surface and ground water environments third edition. *Prentice Hall*.

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. In: Monographs on Statistics and Applied Probability, vol. 57. Chapman & Hall, London, 436 pp.

Essington, M. E., 2015. *Soil and Water Chemistry: An Integrative Approach, Second Edition*: CRC Press.

Fessant, F., Midenet, S. 2002. Self-organizing map for data imputation and correction in surveys. Neural Computing & Applications 10: 300-310.

Fissore, C., Giardina, C. P., Kolka, R. K., Trettin, C. C., 2009. Soil organic carbon quality in forested mineral wetlands at different mean annual temperature. *Soil Biology and Biochemistry, 41*(3), 458-466.

Fitts, C.R., 2002. Groundwater Science. Academic Press, Massachussets

Fitzsimons, S. J.,Veit, H., 2001. Geology and Geomorphology of the European Alps and the Southern Alps of New Zealand. *Mountain Research and Development, 21*(4), 340-349.

Freeze, R.A., Cherry, J.A., 1979. Groundwater. Prentice-Hall, New Jersey.

Friedel, M.J., Souza, O.F., Iwashita, F., Yoshinaga, S.P., Silva, A.M., 2012. Data-driven modeling for groundwater exploration in fractured crystalline terrain, northeast Brazil. *Hydrogeol. J. 20*: 1061-1080.

Friedel, M.J., 2014. Data-driven modeling of background and mine-related acidity and metals in river basins. *Environmental Pollution, 184*: 530-539.

Friedel, M.J., Esfahani, A., in review. Toward real-time 3D mapping of surficial aquifers using a hybrid modeling approach, submitted to Hydrogeology Journal (10-01-2015).

Frossard, E., Condron, L. M., Oberson, A., Sinaj, S., Fardeau, J., 2000. Processes governing phosphorus availability in temperate soils. *Journal of Environmental Quality, 29*(1), 15-23.

Gat, J.R., 1980. The relationship between surface and subsurface waters: water quality aspects in areas of low precipitation. *Hydrological Sciences Bulletin 25*, 257-267.

Goyal, M. R., Harmsen, E. W., 2013. Evapotranspiration: Principles and Applications for Water Management. *Apple Academic Press*.

Guggenmos, M.R., Daughney, C.J., Jackson, B.M., Morgenstern, U., 2011. Regional-scale identification of groundwater-surface water interaction using hydrochemistry and multivariate statistical methods, Wairarapa Valley, New Zealand. *Hydrol. Earth Syst. Sci. 15*: 3383-3398.

Güler, C., Thyne, G.D., McCray, J.E., Turner, A.K., 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology Journal 10*:455–474.

Gusyev, M., Moreau-Fournier, M., Tschritter, C., 2011. Capture Zone delineation of National Groundwater Monitoring Programme Sites in the Southland region. GNS Science Consultancy Report 2011/31, …pp.

Hastie, T., Tibshirani, R., Friedman, J., 2002. The Elements of Statistical Learning. Springer-Verlag, Berlin.

Hem, J.D., 1985. Study and interpretation of the chemical characteristics of natural waters. United States Geological Survey Water Supply Paper 2254.

Hornung, R. W., Reed, L. D., 1990. Estimation of average concentration in the presence of nondetectable values. *Applied occupational and environmental hygiene, 5*(1), 46-51.

Houlbrooke, D., Horne, D., Hedley, M., Snow, V., Hanly, J., 2008. Land application of farm dairy effluent to a mole and pipe drained soil: implications for nutrient enrichment of winter-spring drainage. *Soil Research, 46*(1), 45-52.

Hughes, B., 2013. Discharge to Land Application – Southland District Council, Te Anau Wastewater, http://es.govt.nz/media/39315/brydon_hughes_11-11-2013.pdf.

Iwashita, F., Friedel, M.J., de Souza, C.R., Filho., Fraser, S.J., 2011. Hillslope chemical weathering across Paraná, Brazil: a data mining-GIS hybrid approach. Geomorphology 132: 167-175.

Jiang, Y., Wu, Y., Groves, C., Yuan, D., Kambesis, P., 2009. Natural and anthropogenic factors affecting the groundwater quality in the Nandong karst underground river system in Yunan, China. *Journal of Contaminant Hydrology 109*, 49-61.

Jurgens, B.C., McMahon, P.B., Chapelle, F.H., Eberts, S.M., 2009. An Excel workbook for identifying redox processes in ground water. USGS Open File Report 2009-1004, USGS, Sacramento, California.

Kalteh, A.M., Berndtsson, R., 2007. Interpolating monthly precipitation by self-organizing map (SOM) and multilayer perceptron (MLP), *Hydrol. Sci. J. 52*: 305-317.

Kendall, C., Caldwell, E.A., 1998. Fundamentals of Isotope Geoichemistry. In: Kendall c., McDonnell , J.J (eds.), Isotope Tracers in Catchment Hydrology. *Elsevier Science B.V., Amsterdam*. Pp. 41-86.

Killick, M., Stenger., R., Rissmann, C., 2014. Estimating soil zone denitrification potential for Southland, Environment Southland and Lincoln Agritech Ltd, Invercargill. 29p

King, A. C., Raiber, M., Cendón, D. I., Cox, M. E., Hollins, S. E., 2015. Identifying flood recharge and inter-aquifer connectivity using multiple isotopes in subtropical Australia. *Hydrol. Earth Syst. Sci., 19*(5), 2315-2335.

Kohonen, T., 2001. Self-organizing maps. Third Edition. Springer, Berlin.

Krantz, D.E., Powars, D.S., 2002. Hydrogeologic setting and potential for denitrification in groundwater, Coastal Plain of Southern Maryland: U.S. Geological Survey Water-Resources Investigations Report 00-4051.

Lambert, M.G., Devantier, B.P., Nes, P., Penny, P.E., 1985. Losses of nitrogen, phosphorus, and sediment in runoff from hill country under different fertiliser and grazing management regimes. *New Zealand Journal of Agricultural Research 28*:371.

Lambrakis, N., Antonakos, A., Panagopoulis, G., 2004. The use of multicomponent statistical analysis in hydrogeological environmental research. *Water Res 38*: 1862-1872.

Langmuir, D., 1997. Aqueous Environmental Geochemistry. Prentice-Hall, New Jersey.

Lasaga, A., 1984. Chemical Kinetics of Water-Rock Interactions**.** *Journal of Geophysical Research, vol. 89*, no. B6, pages 4009-4025, June 10, 1984.

Legrand, M., Puxbaum, H., 2007. Summary of the CARBOSOL project: Present and retrospective state of organic versus inorganic aerosol over Europe. *Journal of Geophysical Research: Atmospheres, 112*(D23).

Malek, M.A., Harun, S., Shamsuddin, S.M., Mohamad, I. 2008. Imputation of time series data via Kohonen self organizing maps in the presence of missing data. Engineering and Technology 41: 501-506

Martins, J.M.F., Majdalani, S., Vitorge, E., Desaunay, A., Navel, A., Guine, V., Daian, J.F., Vince, E., Denis, H., Gaudet, J.P., 2013. Role of macropore flow in the transport of Escherichia coli cells in undisturbed cores of a brown leached soil. *Environmental Science: Processes & Impacts 15*, 347-356.

McMahon, P.B., Chapelle, F., 2008. Redox processes and water quality of selected principal aquifer systems. *Groundwater 46*: 259-271

Mitsch, W.J., Gosselink, J.G., 2007. Wetlands. New York. John Wiley & Sons, Inc. 582pp.

Monaghan, R. M., Smith, L. C., Muirhead, R. W., 2016. Pathways of contaminant transfers to water from an artificially-drained soil under intensive grazing by dairy cows. *Agriculture, Ecosystems & Environment, 220*, 76-88.

Moreau, M., Nokes, C., Cameron, S., Hadfield, J., Gusyev, M., Tschritter, C., Daughney, C., 2014. Capture zone guidelines for New Zealand, GNS Science Report 201 3/56. 52 p.

Moreau, M., Cameron, S., Daughney, C., Gusyev, M., Tschritter, C., 2014a. Envirolink Tools Project – Capture Zone Delineation – Technical Report, GNS Science Report 2013/57.98 p.

Morgan, M., Evans, C., 2003. Southland Water Resources Study Stage 1-3, report prepared for the Southland Regional Council, report number 4597/1.

Nichol, S. E., Harvey, M.J. Boyd, I.S., 1997. Ten years of rainfall chemistry in New Zealand. pp 30.

Nõges, T., 2009. Relationships between morphometry, geographic location and water quality parameters of European lakes. *Hydrobiologia, 633*(1), 33-43.

Nziguheba, G., Palm, C. A., Buresh, R. J., Smithson, P. C.,1998. Soil phosphorus fractions and adsorption as affected by organic and inorganic sources. *Plant and Soil, 198*(2), 159-168.

Oliver, D. M., Clegg, C. D., Haygarth, P. M., Heathwaite, A. L., 2003. Determining hydrological pathways for the transfer of potential pathogens from grassland soils to surface waters. 36-41.

Oliver, D.M., Heathwaite, L.A., 2013. Pathogen and Nutrient transfer through and across agricultural soils. In: Meyers (ed). Encyclopedia of Sustainability Science and Technology

Olson, J, R., 2012. The Influence of Geology and Other Environmental Factors on Stream Water Chemistry and Benthic Invertebrate Assemblages. *All Graduate Theses and Dissertations.* Paper 1327

Parkhurst, D.L., Appelo, C.A.J., 2013. Description of input and examples for PHREEQC version 3—A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations: U.S. Geological Survey Techniques and Methods, book 6, chap. A43

Peukert, S., Griffith, B.A., Murray, P.J., Macleod, C.J.A., Brazier, R.E., 2014. Intensive Management in Grasslands Causes Diffuse Water Pollution at the Farm Scale. *Journal of Environmental Quality 43*.

Phreatos Limited., 2007. Mid-Mataura Groundwater Model. Report for Environment Southland.

Rademacher, L.K., Clark, J.F., Hudson, G.B., Erman, D.C., Erman, N.A., 2001. Chemical evolution of shallow groundwater as recorded by springs, Sagehen basin; Nevada County, California. *Chem. Geol. 179*: 37-51.

Rallo, R., Ferre-Gine, J., Arenas, A., Giralt, F., 2002. Neural virtual sensor for the inferential prediction of product quality form process variables. *Comput. Chem.Eng. 26*: 1735-1754.

Rissmann, C., 2011. Regional Mapping of Groundwater Denitrification Potential and Aquifer Sensitivity, Technical Report. Environment Southland publication No 2011-12, Invercargill. 40p

Rissmann, C., Wilson, K.L., Hughes, B.N., 2012. Waituna Catchment Groundwater Resources, Technical Report. Environment Southland publication No 2012-04, Invercargill. 92p

Rissmann, C., Leybourne, M., Benn, C., Christenson, B., 2015. The origin of solutes with in the groundwaters of a high Andean aquifer. Chem. Geol., 396: 164-181/

Rosen, M.R., 2001. Hydrochemistry of New Zealand's aquifers. p. 77-110 In: Rosen, M.R.; White, P.A. (eds.) Groundwaters of New Zealand. Wellington: New Zealand Hydrological Society.

Salvador, P., Artíñano, B., Pio, C., Afonso, J., Legrand, M., Puxbaum, H., Hammer, S., 2010. Evaluation of aerosol sources at European high altitude background sites with trajectory statistical methods. *Atmospheric Environment, 44*(19), 2316-2329.

Scott, C.A., Geohring, L.D., Walter, M.F., 1998. Water quality impacts of tile drains in shallow, sloping, structured soils as affected by manure application. *Applied Engineering in Agriculture*, 14 (6): 599-603.

Sinclair, A.J., 1974. Selection of threshold values in geochemical data using probability graphs. *Journal of Geochemical Exploration*: Vol. 3, No. 2, pp. 129–149.

SKM (Sinclair Knight Merz)., 2005. Northern Southland Groundwater Model – Model Development Report, report prepared for Environment Southland.

Smith, L.C., Monaghan, R.M., 2003. Nitrogen and phosphorus losses in overland flow from a cattle-grazed pasture in Southland. *New Zealand Journal of Agricultural Research, 46*(3*)*: 225-237

Stenger. R., Clague, J., Woodward, S., Moorhead. B., Wilson, S., Shokri, A., Wohling, T., Canard, H., 2013. Denitrification – the key component of a groundwater system's assimilative capacity for nitrate. In: *Accurate and efficient use of nutrients on farms*. (Eds L.D. Currie and C L. Christensen). http://flrc.massey.ac.nz/publications.html. Occasional Report No. 26. Fertilizer and Lime Research Centre, Massey University, Palmerston North, New Zealand, 11 pp.

Stenger. R., Clague, J., Woodward, S., Moorhead. B., Wilson, S., Shokri, A., Wohling, T., Canard, H., 2014. Rootzone losses are just the beginning. In: *Nutrient management for the farm, catchment and community.* (Eds L.D. Currie and C.L.Christensen). http://flrc.massey.ac.nz/publications.html. Occasional Report No. 27. Fertilizer and Lime Research Centre, Massey University, Palmerston North, New Zealand, 8 pp.

Strang, D., Aherne, J., Shaw, D.P., 2010. The hydrochemistry of high-elevation lakes in Georgia Basin, British Columbia. *J. Limnol., 69(Suppl. 1): 56-66*

Toews, M.W., Gusyev, M.A., 2013. GIS tools to delineate groundwater capture zones, GNS Science Report 2012/06. 24 p.

Vesanto, J. 1999. SOM-based data visualization methods. Intelligent Data Analysis 3: 111-126.

Vesanto, J., Alhoniemi, E., 2000. Clustering of the self-organizing map. IEEE Transactions on Neural Networks 11: 586-600.

Watson, C.J., Jordan, C., Lennox, S.D., Smith, R.V., Steen, R.W.J., 2000. Inorganic Nitrogen in Drainage Water from Grazed Grassland in Northern Ireland. *Journal of Environmental Quality 29*.

Winter, T. C., Harvey, J. W., Franke, O. L., Alley, W. M., 1998. Ground water and surface water, a single resource, US Geological Survey Circular 1139, US Geological Survey, Denver, USA.

Zhang, W., Tang, X.-Y., Weisbrod, N., Zhao, P., Reid, B.J., 2015. A coupled field study of subsurface fracture flow and colloid transport. *Journal of Hydrology 524*, 476-488