## Watershed Characteristics and Sampling Conditions as Driving Forces for Dynamics of Aqueous Natural Organic Matter Across the Leech River Watershed

### Synopsis and introduction

The previous chapter showed that across the Greater Victoria Water Supply Area (GVWSA), the character of natural organic matter (NOM) shifted from predominately aliphatic in the dry season to predominately aromatic during the wet seasons. Additionally, Chapter 3 results confirmed that event-based Rack samples had higher DOC concentrations than standard Grab samples. The seasonal character shift agreed with predictions of the river continuum concept (RCC), as did an observed spatial reduction in NOM molecular diversity from lower-order to higher-order streams (Vannote et al. [1980](#ref-Vannote1980); Abbott et al. [2018](#ref-Abbott2018); Mosher et al. [2015](#ref-Mosher2015); Creed et al. [2015](#ref-Creed2015)). Elevated NOM in stormflow was expected (e.g. Vidon, Wagner, and Soyeux ([2008](#ref-Vidon2008)); Raymond et al. ([2010](#ref-Raymond2010))) and is a key aspect of the pulse-shunt concept (Raymond et al. [2016](#ref-Raymond2016)).

Of the main tributaries to the Leech River mainstem, Cragg crk (Fig X, draining from the east) had the lowest average DOC concentration and least aromatic NOM character, while West Leech was higher in DOC concentration and more aromatic. It was expected that that there would be differences in NOM concentration and character among monitoring sites, but it wasn’t immediately obvious why event-based NOM variances were greater at the West Leech site (20.85 km2 sub-basin) compared to other streams draining catchments of similar size (i.e. Leech-head, 20.59 km2 sub-basin; and Cragg crk 28.06 km2 sub-basin). As water quality is intrinsically linked to watershed characteristics, there may be a key feature (or combination of features) that operate in tandem with weather conditions to drive NOM concentration and character dynamics. This chapter evaluates watershed characteristics (e.g. land cover, parent material, slope) in concert with hydroclimatic conditions (e.g. antecedent rain, sampling stage) to identify possible explanatory variables for differences and drivers of change in NOM concentration and character across the six monitoring sites in the Leech WSA.

#### Random Forests

A Random Forest is a collection of decision trees that collectively comprise a statistical tool for non-parametric regression, prediction, classification and assessment of variable importance (Strobl, Malley, and Tutz [2009](#ref-Strobl2009)). Breiman’s Random Forests (RF) is a machine learning algorithm for practical applications, which is popular for its accuracy in real-world systems (Tyralis, Papacharalampous, and Langousis [2019](#ref-Tyralis2019); Biau and Scornet [2016](#ref-Biau2016)) and does not require independence among samples. As such, it is appropriate for analysis of nested catchments which by definition are not independent. RF has been widely published in hydrologic and water resource research in recent years, particularly in streamflow and water quality studies (see Tyralis, Papacharalampous, and Langousis [2019](#ref-Tyralis2019) and references therein). In RF, a set of predictor variables (features) are used to predict the outcome of another variable (the predictant) through supervised learning algorithms which are grouped as either regression or classification, depending on whether the variables are quantitative (i.e. numeric) or qualitative (Breiman [2001](#ref-Breiman2001); Tyralis, Papacharalampous, and Langousis [2019](#ref-Tyralis2019)).

The RF algorithm learns from many independent Classification And Regression Trees which undergo bootstrap-aggregating (“bagging”) with randomization; meaning that no single tree includes all the data, which reduces over-fitting and improves prediction performance (Breiman [2001](#ref-Breiman2001); Tyralis, Papacharalampous, and Langousis [2019](#ref-Tyralis2019); Biau and Scornet [2016](#ref-Biau2016)). In RF, many decision trees are created, each one generated by randomly sampling from the full data set (with replacement) and each tree is trained until only one possible solution remains (i.e. leaf nodes contain one deciding variable). For problems of regression (i.e. quantitative numeric variables, such as the data in this thesis), the final forest prediction is equal to the average of individual tree predictions (Tyralis, Papacharalampous, and Langousis [2019](#ref-Tyralis2019); Biau and Scornet [2016](#ref-Biau2016)).

The capacity to evaluate variable importance metrics sets RF apart from other data-driven models that focus on prediction (Tyralis, Papacharalampous, and Langousis [2019](#ref-Tyralis2019)). The relative importance of each feature can be determined by assessing the accuracy of how well the response variable (predictant) is anticipated in the absence or presence of each predictor feature. Through the RF algorithm, the variable importance measure (VIM) VIM is assessed by removing predictor variables one at a time and measuring the decrease in prediction accuracy in their absence (Breiman [2001](#ref-Breiman2001)). This is achieved based on either the increase in mean square error (MSE, type 1) or the increase in sum of square errors (SSE, type 2). Type 1 of the VIM is calculated by permutation, where “the prediction error on the out-of-bag portion of the data is recorded (error rate for classification, MSE for regression). Then the same is done after permuting each predictor variable. The difference between the two are then averaged over all trees, and normalized by the standard deviation of the differences” (Liaw and Wiener [2018](#ref-Liaw2018)). VIM type 2 “is the total decrease in node impurities from splitting on the variable, averaged over all trees. For classification, the node impurity is measured by the Gini index. For regression, it is measured by residual sum of squares” (Liaw and Wiener [2018](#ref-Liaw2018)). The SSE method may be more appropriate for categorical variables than quantitative variables, for which MSE method appears to be more appropriate (citation). Despite the ability of RF to handle non-independent real world data, bias can result from cross-correlations among predictor variables, and mixing categorical and quantitative data may result in bias towards variables with many categories or many missing values (C. Strobl et al. [2008](#ref-Strobl2008)). Therefore, some predictor refinement is important.

#### Hysteresis

Storm events are understood to be a key driver for solute and particle mobilization in streams (e.g. Aguilera and Melack ([2018](#ref-Aguilera2018)); …\***add more refs**). Material transport (be it in solution or suspension) tends to show a non-linear relationship to discharge, in that concentrations do not necessarily peak in congruence with flow. This hysteretic behaviour, where the relationship between concentration and discharge (i.e., C-Q) is different on the rising limb versus the falling limb of the hydrograph, creates a hysteresis loop that can reveal information about origins and transport of aquatic material, relative to the point of measurement (Vaughan et al. [2019](#ref-Vaughan2019); Evans and Davies [1998](#ref-Evans1998); Aguilera and Melack [2018](#ref-Aguilera2018)). Concentration-discharge hysteresis may be due to episodic flushing of material or from component mixing processes (Evans and Davies [1998](#ref-Evans1998)). Information about flow-paths and source pools can be obtained from four elements of the C-Q relationship plot: the shape (linear, circular, figure-eight), the direction of a hysteresis loop (clockwise vs. counterclockwise), the magnitude, and the direction of slope (Aguilera and Melack [2018](#ref-Aguilera2018)).

Concentration-discharge relationships can elucidate flow-paths and aid in estimating proportional streamflow contributions (i.e. subsurface vs. overland, or near-stream vs. upland) (Evans, Davies, and Murdoch [1999](#ref-Evans1999)). C-Q can be used to determine whether a catchment’s solute export regime is chemostatic or chemodynamic (Musolff et al. [2015](#ref-Musolff2015)). For example, in an assessment of ~400 stream events in coastal California, Aguilera and Melack ([2018](#ref-Aguilera2018)) found that sediment-associated concentrations peaked during high flows but dissolved nitrogen species had hysteretic behaviour with flow, the direction of which differed depending on land-use.

Dominant sources and flow paths during events can be inferred from the slope (positive or negative) and the direction of the hysteresis loop rotation (Evans and Davies [1998](#ref-Evans1998); Aguilera and Melack [2018](#ref-Aguilera2018)). A clockwise C-Q hysteresis loop indicates rapid flushing and dilution (higher concentration on the rising limb than falling limb) while a counter-clockwise loop indicates enrichment via delayed material delivery from distant (upstream) or deeper subsurface source pools (Aguilera and Melack [2018](#ref-Aguilera2018)). Evans et al. ([1999](#ref-Evans1999)) determined that C-Q analysis requires a minimum of five samples to determine the rotational direction of hysteresis: pre and post episode low-flow, rising and falling limbs and near to peak Q.

Using end-member mixing, with hydrometric monitoring and carbon isotope (δ13CDOC) data, Lambert et al. ([2014](#ref-Lambert2014)) determined that approximately 80% of stream DOC flowed through the riparian domain’s most superficial soil horizons, and that riparian soil DOC was sourced from both riparian and upland areas. Upland NOM source pools were determined to contribute significant amounts of carbon to streams but were supply-limited; whereas riparian NOM was found to be a near-finite carbon source. Lambert et al. ([2014](#ref-Lambert2014)) also found that upland sources contributed relatively more early in the wet season (~30% of stream DOC) and less as high-flows continued (decreased to ~10% of stream DOC later in the wet period). In that vein, antecedent weather conditions are important for nutrient and sediment export and C-Q relationships, as the first precipitation events after a prolonged dry period can result in rapid flushing of accumulated material (Aguilera and Melack [2018](#ref-Aguilera2018)).

### Methods

#### Site details

This chapter focuses on the six Leech WSA monitoring sites (Figure 13), which were equipped with Vertical Rack samplers. Nested catchments of the Leech WSA included two headwater streams, Weeks and Chris crk sub-basins (site 1, 11.5 km2 & site 2 , 5.9 km2); the head of Leech River (‘Leech-head’), below the confluence of those two headwater streams (site 3, 20.6 km2 sub-basin); two major tributaries that feed the Leech, Cragg crk and West Leech sub-basins (site 4, 28.1 km2 & site 5, 20.9 km2), and Leech Tunnel (site 6) which encompassed drainage of the entire Leech WSA above the point of diversion (95.3 km2). Sample collection and analysis follow the methods detailed in Chapter 2.

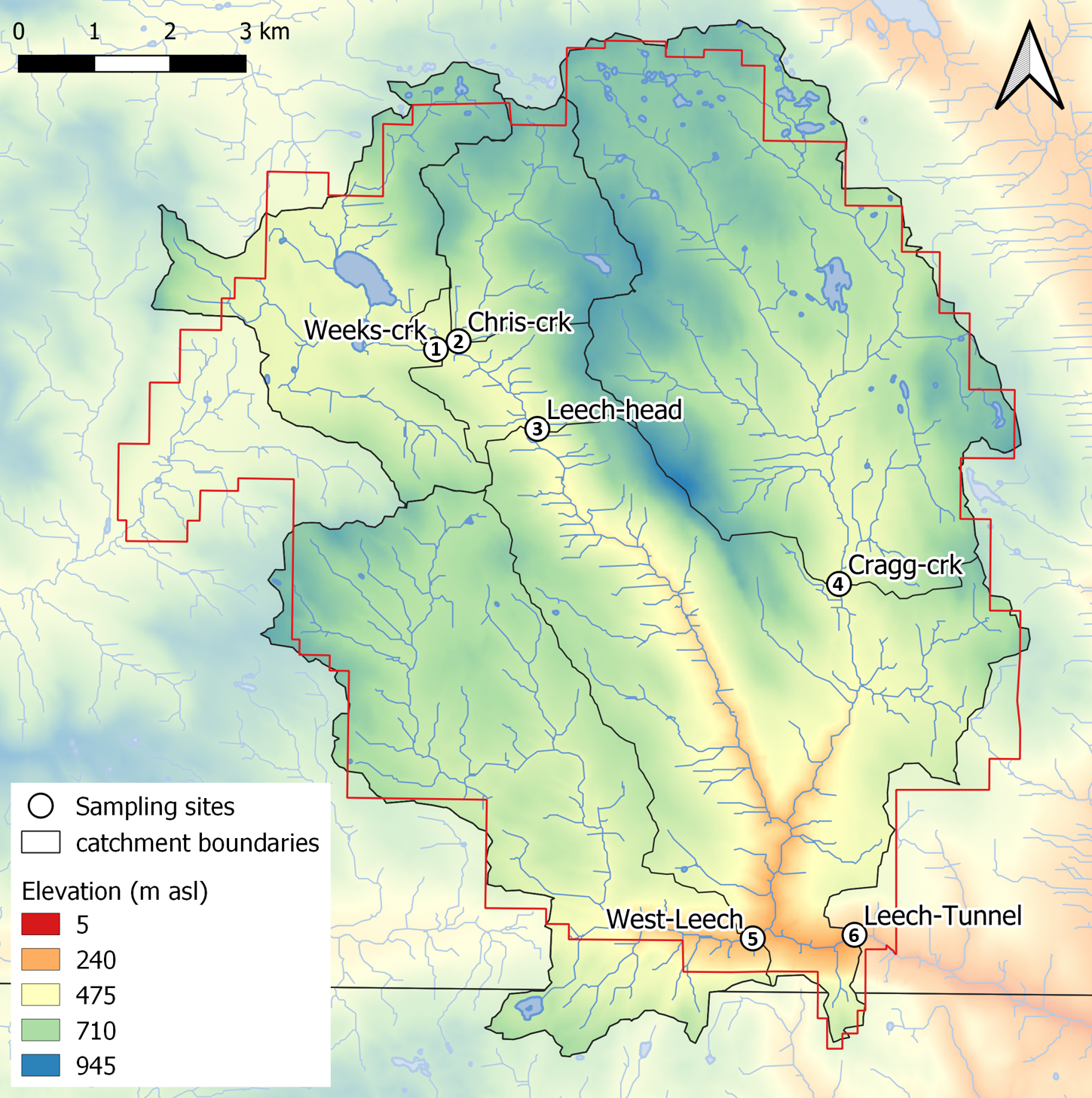


Figure 13: Monitoring sites in the Leech Water Supply Area (WSA). Sites are named and numbered, where the number indicates both the sequence of vertical rack installations as well as the relative progression from headwaters to mouth. Each sub-basin boundary (defined using each sampling site as the drainage outlet) is outlined in black, and the Leech WSA boundary is delineated in red.

Surficial materials and soils in the Leech WSA are predominantly podzolic with approximately 50% of the parent material in each sub-basin composed of colluvial deposits with O-HFP development (orthic humo-ferric podzol). The second most dominant parent material for soil was till (morainal) with duric humo-ferric podzol (DU-HFP) soil development in all sub-basins except for West Leech, which had till-based O-HFP soils (Government of Canada [2018](#ref-SoilsCanada2018)). DU-HFP differs from O-HFP by the presence of a strongly cemented layer (duric horizon) that is unlike the B horizon in O-HFP (Canadian Society of Soil Science [2020](#ref-SoilScience2020)). Both DU-HFP and O-HFP are well drained, or moderately so, and O-HFP tends to be acidic (pH < 5.6) while DU-HFP is generally moderately acidic to neutral (pH 5.6 - 7.4) (Government of Canada [2018](#ref-SoilsCanada2018)). There was a small amount of poorly drained terric mesisol (T.M., undifferentiated organic soil) present only in the Weeks crk sub-basin (Table 14).

There were a few notable differences in subsurface geology between sub-basins (Table 14). West Leech was characterized by the absence of wark gneiss and chert argillite volcanic parent materials (0%), which were present in all other sub-basins. Both Weeks crk and West Leech sub-basins were dominated by argillite metagreywacke parent material (64% and 77%, respectively). West Leech was also underlain by metagreywacke (7%) and Metchosin volcanics (16%), which were absent in each of the other monitoring sub-basins. Cragg crk sub-basin (site 4) was underlain predominantly by metamorphic wark gneiss (78%) with no argillite metagreywacke parent material (0%). Metagreywacke parent materials are part of the metasedimentary Leech River Formation, where metagreywacke is meta-sandstone and argillite metagreywacke is more of a meta-mudstone (Groome et al. [2003](#ref-Groome2003)). The Leech River Formation tends to be more erodible than wark-gneiss (Ussery and AECOM [2015](#ref-Ussery2015)). Chris crk and Cragg crk sub-basins had no metasedimentary parent material of the Leech River Formation.

At the surface, Weeks crk sub-basin had the greatest proportion of wetland and open water compared to the other sub-basin sites (Table 14). Cragg crk sub-basin had the oldest average tree age (59 years). By area, the most heavily harvested basin since 1980 was the Chris crk sub-basin (63% harvested), then Leech-head and Cragg crk sub-basins (43% and 41% harvested, respectively). West Leech was the least harvested between 1980 and 2011 (26%), followed by Weeks crk sub-basin (28% harvested, but is also 7% open water and wetland), followed by the entire Leech Tunnel catchment area (34%).

Table 14: Summary of Leech Watershed Characteristics for Monitoring Site Sub-basins. Variables Preceded by Square Brackets Indicate Sub-surface Features of Parent Material ([PM]) and Soil Groups ([S])

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable by site | **Weeks crk** | **Chris crk** | **Leech-head** | **Cragg crk** | **West Leech** | **Leech Tunnel** |
| Latitude | 48.57592 | 48.57691 | 48.5666 | 48.54856 | 48.50635 | 48.5069 |
| Longitude | -123.84397 | -123.83995 | -123.82569 | -123.77141 | -123.78618 | -123.768 |
| Interest / characteristic | Headwater | Headwater | Leech Rv. | Major trib. | Major trib. | Mainstem, future diversion |
| Strahler Order | 3 | 3 | 4 | 4 | 4 | 5 |
| Drainage Area (km2) | 11.5 | 5.9 | 20.6 | 28.1 | 20.9 | 95.3 |
| Elevation (m a.s.l) | 521 | 522 | 476 | 509 | 248 | 207 |
| Percent Forest | 94.5 | 99 | 96.6 | 97.8 | 98.5 | 97.6 |
| Percent Wetland | 4.2 | 0.8 | 2.6 | 1.6 | 0.4 | 1.1 |
| Percent open water | 2.9 | 0.6 | 1.8 | 0.9 | 0.3 | 0.7 |
| slope, mean (degrees) | 9.1 | 10.5 | 10.2 | 9.5 | 11.3 | 11.8 |
| slope, max (degrees) | 44 | 38 | 50 | 40 | 56 | 61 |
| Logging history, 1980-2011 (% area) | 28.1 | 62.6 | 42.7 | 41.3 | 25.8 | 33.9 |
| Tree age (average) | 52 | 40 | 48 | 59 | 49 | 53 |
| [PM] Wark-Gneiss (%) | 13.6 | 44.9 | 20.5 | 77.6 | 0 | 30.6 |
| [PM] Argillite-Metagreywacke (%) | 64.2 | 0 | 42.1 | 0 | 76.8 | 45.1 |
| [PM] Chert-Argillite-Volcanic (%) | 22.2 | 55.1 | 37.4 | 22.4 | 0 | 17.9 |
| [PM] Metagreywacke (%) | 0 | 0 | 0 | 0 | 7.2 | 1.6 |
| [PM] Metchosin Volcanics (%) | 0 | 0 | 0 | 0 | 16 | 4.2 |
| [PM] Sooke Gabbro (%) | 0 | 0 | 0 | 0 | 0 | 0.7 |
| [S] Colluvial (O-HFP, %) | 42.2 | 48.1 | 44.3 | 48.1 | 50 | 47.9 |
| [S] Morainal / Till (DU-HFP, %) | 26.6 | 51.9 | 36.1 | 31.8 | 1.8 | 19.2 |
| [S] Morainal / Till (O-HFP, %) | 17.8 | 0 | 12.9 | 1 | 48.2 | 25.6 |
| [S] Undifferentiated organic (T.M, %) | 6.4 | 0 | 3.6 | 0 | 0 | 0.8 |
| [S] Glaciofluvial (DU-HFP, %) | 0.5 | 0 | 0.6 | 6.9 | 0 | 2.2 |
| [S] Fluvial (DU-HFP, %) | 3.9 | 0 | 2.2 | 0 | 0 | 0.5 |

The watershed characteristics included in Table 14 were combined with Leech WSA weather data (Chapter 2) and sample results (Chapter 3) to identify possible explanatory variables that were driving changes in NOM concentration and character across the Leech WSA. Watershed characteristics and conditions were evaluated as possible predictors for NOM change using Random Forest algorithms.

#### Random Forests predictor variable refinement & quality control

RF variable importance measures (VIMs) were used to assess the relative importance of sub-basin characteristics and conditions as predictors of NOM concentration and character using the MSE method (type 1). Possible predictor variables for NOM concentration (i.e. DOC) and character (i.e. SAC254 and E2:E3) included sub-basin characteristics of each of the six monitoring sites (Table 14) as well as sampling stage and antecedent weather conditions. To avoid undue bias (C. Strobl et al. [2008](#ref-Strobl2008)) all predictor variables were numeric values (quantitative) and missing values were removed. As cross-correlated predictors can cause biased VIM in RF, variables with correlation coefficients greater than 0.65 were omitted (though there is no standard definition for levels of correlation coefficient strengths, 0.65 seemed a reasonable threshold to separate moderate and more strongly correlated variables).

Values for median basin slopes were the same across three sites, so mean slopes were selected instead, as means were unique to each sampling basin. Sub-basin percent wetland cover was correlated with slope (-0.73), as well as forest cover (-0.97). Thus, wetland and forest percent cover were not included in VIM assessments, and slope was. Interestingly, the percent of sub-basins underlain by the meta-sedimentary Leech River formation (including argillite-metagreywacke (mudstone) and metagreywacke (meta-sandstone)) was correlated (-0.83) with 30-year logging history (percent of basin harvested 1980-2011). The Leech River geologic formation underlies the Weeks crk and West Leech sub-basins, which were also the two sub-basins least harvested since 1980 (Table 14). These are the types of correlations that could cause spurious VIM bias in RF, because it’s unlikely that the presence of meta-sedimentary parent material was the cause for these basins to be less harvested. Parent material variables were refined to metamorphic (wark gneiss) and igneous (Metchosin volcanics & gabbro stocks), which did not show cross-correlation to other predictor variables. As one might expect, soil materials were correlated with geologic parent materials (as well as with each other and logging history), thus soils were not included with RF VIM predictor variables.

The nine selected predictor variables for RF VIM assessment included:

* metamorphic parent-material (wark gneiss, percent of basin)
* igneous parent material (Metchosin volcanics & gabbro stocks, percent of basin)
* drainage area (km2)
* mean basin slope (degrees)
* tree age (average, years)
* logging history (percent of basin harvested 1980-2011)
* antecedent 7-day air temperatures (°C, mean)
* antecedent 30-day rain (mm, total)
* sampling stage (normalized)

Sampling stages for each site were scaled (min-max-normalized) to account for differences between Vertical Rack installations and to allow for comparison between sites (example calculation in Appendix *XXXX*). For antecedent rain and air temperatures, different periods of time were considered in exploratory data analysis. Intervals of 3, 5, 7, 14, 21 and 30 days prior to sample collection were evaluated for cross-correlations and relative VIM results. The length of antecedent periods of 30 days for antecedent rain and 7 days for antecedent air temperature were chosen because (1) they were not cross-correlated with other predictor variables or each other and (2) each was more easily distinguished in VIM results than other interval options.

Furthermore, the amount of rainfall in the 30 days prior to sample collection could be considered as an indicator of overall antecedent landscape wetness, whereas shorter periods of rain may not have been adequate indicators of wetness due to interception, evaporation, or evapotranspiration losses. A longer period of antecedent rain may be indicative of greater connectivity of streams to distant/deeper terrestrial environments. Using the same VIM parameters tuned for DOC prediction, 30-day antecedent rain was relatively more important than 7-day rain (15.3% vs. 12.8%) for predicting E2:E3, and 7-day antecedent air temperatures were more important than 30-day air temperatures (8.5% vs. 6.5%). This confirmed that relatively short-term temperatures (rolling 7-day mean) and relatively long-term rain (rolling 30-day sum) were appropriate selections for the RF VIM assessments.

Surface and subsurface watershed characteristics were static values (e.g. basin slope, parent material), whereas sampling stage and antecedent weather were dynamic values (different for each sample). Although all variables were numeric, it was possible that RF might treat static values categorically (citation), and RF bias can result from VIM assessment of combined categorical and quantitative predictor variables. So, sub-basin characteristics and conditions were evaluated independently as well as together to see if there was a shift in overall relative VIM ranking results.

RF is a black box; for that reason, a variable of random numbers was included in the VIM assessment as a means for quality assurance (QA). VIM assessment, with the QA variable, was completed for watershed characteristics and conditions to ensure that random numbers were not assigned any predictor importance. The QA variable was removed from final assessment and is not displayed in results (below).

Following RF VIM analyses, the variables with relatively greatest predictor importance were further explored by evaluating their one-on-one relationships with NOM concentration and character. Results of RF VIM assessment were also used to retrospectively infer importance of variables that were excluded from VIM due to their cross correlation(s). Where possible, important predictor variables were assessed for hysteretic behaviour.

#### Hysteresis

Hydrochemical hysteresis is most often evaluated through C-Q relationships, but the concept of hysteresis applies to any time-based predictor and effected values. The principle of hysteresis is based on a lag between a driving force (e.g. discharge, Q) and the analyte that is driven to change (e.g. solute concentration, C). Ideally, streamflow data would accompany all sample data collected for this research, however that was not the case. Stage data was collected, and given the strong relationship between stage and discharge, stage can be evaluated as an indicator for Q. Similarly, because streamflow in the Leech WSA is governed by precipitation, antecedent wetness (rain) could be proxy data in a hysteresis analysis. Conditions that were calculated to be most important for predicting NOM concentration and character were plotted with their predictant to assess possible hysteresis; this was only possible for dynamic conditions with variable measurements (i.e. not static watershed characteristics predictors).

For a C-Q hysteresis loop, clockwise rotation indicates rapid flushing and dilution, while counter-clockwise rotation indicates enrichment via delayed material delivery from distant (upstream) or deeper subsurface source pools (Aguilera and Melack [2018](#ref-Aguilera2018)). Because of the assumed relationship between antecedent rain and streamflow, a similar interpretation was used here despite the absence of discharge data.

#### Evaluating local extrema

Results in the previous chapter showed elevated NOM in event-based samples; to explore that relationship more deeply a peak-to-peak comparison was done to see if DOC concentrations peaked with peaks in stream stage levels. As DOC was quantified from discrete stream samples the temporal synchrony of peaks could not be evaluated in the same manner as stage, which was continuously recorded. However, each river sample was matched to stage and had a corresponding timestamp, thus inferences could be made based on relationships between peak DOC and stage. To determine if local DOC extrema (maxima and minima concentrations) were captured in conjunction with the local extrema of sampled flows (minima and maxima of sample stage), manual synchrony tests were run on sample results. For each site, results of samples were grouped by collection period and rain event to identify samples with extreme DOC concentrations as well as samples collected at extreme stage (max/min). Samples corresponding to maxima and minima DOC and stage were compared to determine how often they overlapped (i.e. was the sample with max DOC also the sample collected at the highest stage?).

### Results

#### Random Forest variable importance

A variable of random numbers was included with watershed characteristic and sampling conditions as a method of quality assurance (QA) in using the RF variable importance measure (VIM). Because random numbers are not a real predictor of NOM concentration or character, the VIM assessment should rank the QA variable as the least important predictor and assign it negligible importance. Recall that there are two methods for evaluating variable importance in randomForest (RF): type 1 measures the mean decrease in accuracy and type 2 measures the decrease in node impurity (Liaw and Wiener [2018](#ref-Liaw2018)). For the Leech WSA data, both VIM methods were checked and type 2 (total decrease in node impurities from splitting on the variable) assigned the random number variable relatively high importance (ranked more important than a real variable) while type 1 (permutation on out-of-bag data) ranked the random number variable as least important with insignificant weight (Liaw and Wiener [2018](#ref-Liaw2018)). These results clarified that type 1 VIM was the best choice and suggests that quantitative data is best assessed for VIM by mean decrease in accuracy (MSE, type 1) rather than mean decrease in node impurity (SSE, type 2).

Sub-basin characteristics and conditions were evaluated independently as well as together to see if there was a shift in VIM relative rankings due to the combination of static and dynamic values. While it was expected that the relative importance (as a percent) would shift when all predictant variables were combined, it was not expected that the relative ranking would change – but it did. When dynamic conditions and static characteristics were combined for RF VIM, the relative ranking of several predictants were changed compared to when conditions and characteristics were analyzed separately (see Appendix *XXXX*, Figure 40, for plots of combined predictors). These results suggest that static and dynamic predictor variables should be separated for RF VIM. Here, variables were separated based on whether they were static or dynamic with predictant variables. The combination of dynamic and static predictor variables appeared to create bias relative VIM ranking. The key was not whether the input was numeric, but rather if the predictant changed (or not) with the predictor. It seems that dynamic variables (ones that changed with the predictant) were treated for regression while static variables were treated categorically. VIM results are presented for the three predictants (DOC, SAC254, E2:E3) with conditions and characteristics grouped separately.

##### DOC: concentration

Across the Leech WSA monitoring sites, the hydroclimatic condition with greatest influence on NOM concentration (DOC) was found to be sampling stage, followed by antecedent 7-day air temperature and then antecedent wetness (30-day rain). The top three watershed characteristic predictors for NOM concentrations were slope, percent of the basin underlain by metamorphic parent material and average tree age (Figure 14).

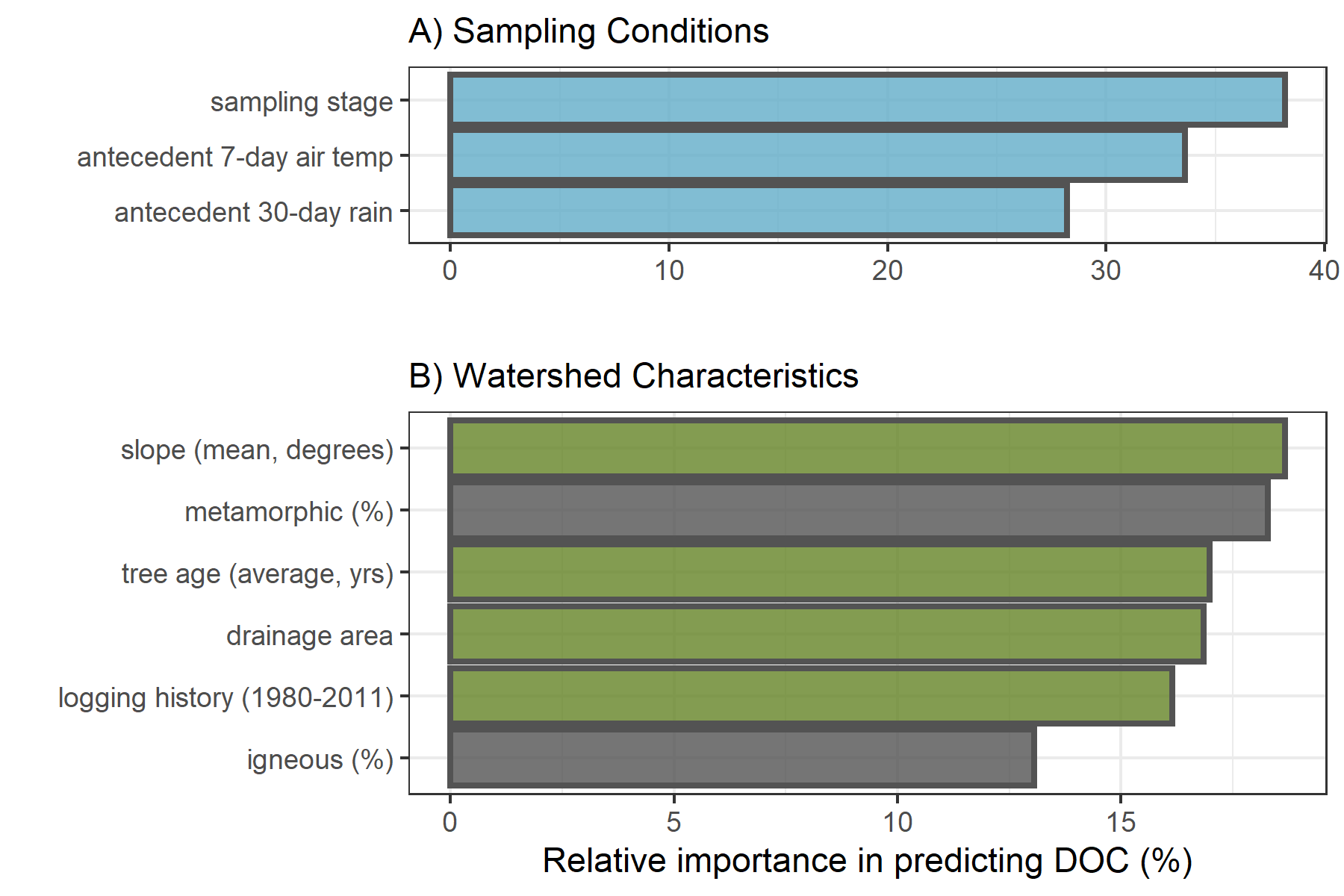


Figure 14: Variable importance for predicting NOM concentration (DOC) across six stream monitoring sites in the Leech River watershed. Variable importance measures were extracted from *randomForest* (in R) and determined by mean decrease in prediction accuracy (increase in the mean square error) in the absence of the predictor variable (i.e. type 1 importance measure).

##### SAC254: reactivity & aromaticity

Like concentration, SAC254 was best predicted by sampling stage. Unlike DOC, SAC254 was relatively more sensitive to antecedent rain than air temperatures. The top three watershed characteristic predictors for NOM aromaticity and reactivity (SAC254) were slope, percent of the basin underlain by metamorphic parent material (wark gneiss), and logging history (Figure 15).

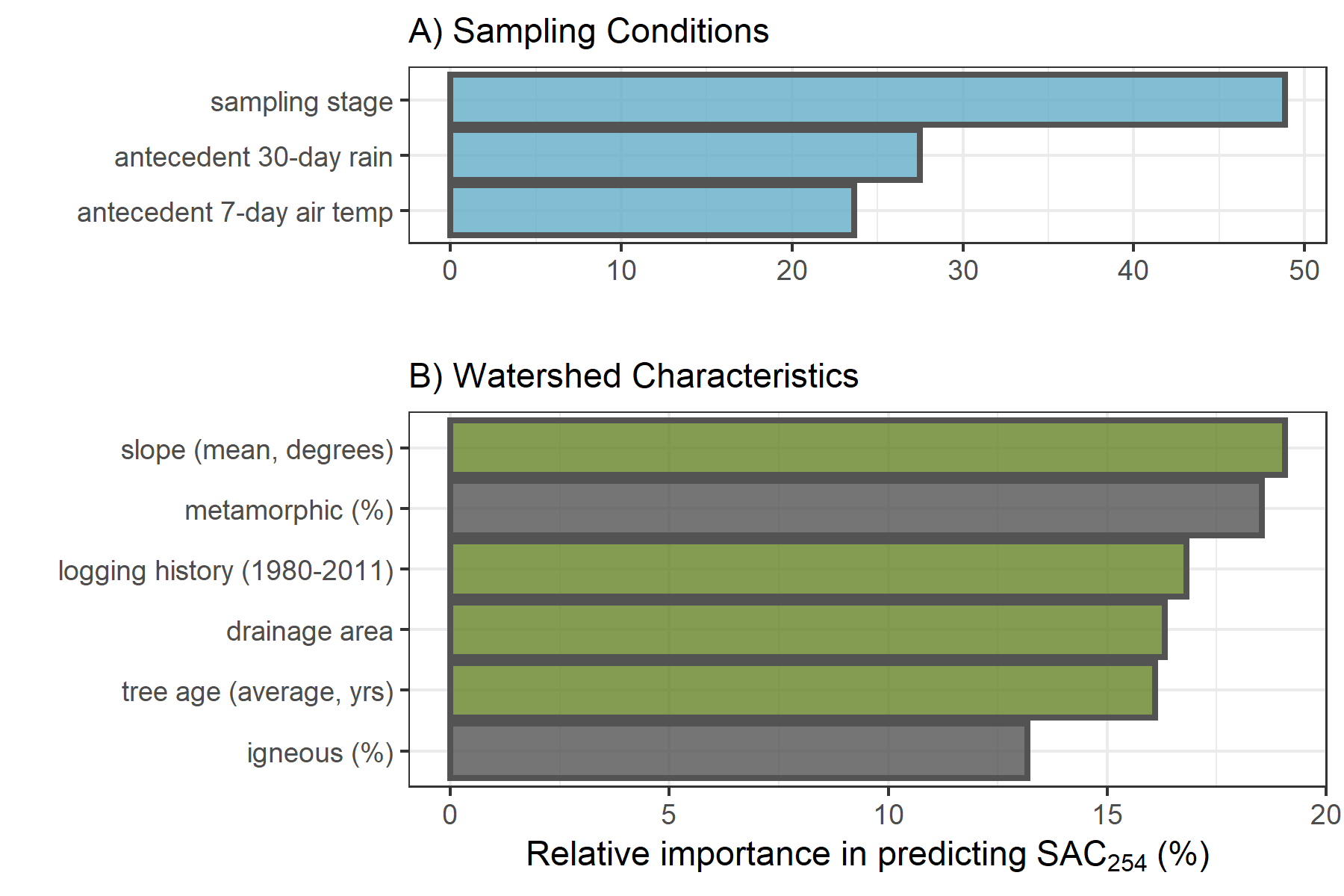


Figure 15: Variable importance for predicting NOM molecular aromaticity (SAC254) across six stream monitoring sites in the Leech River watershed. Variable importance measures were extracted from *randomForest* (in R) and determined by mean decrease in prediction accuracy (increase in the mean square error) in the absence of the predictor variable (i.e. type 1 importance measure).

##### E2:E3: molecular size & aromaticity

Sampling conditions VIM ranking for E2:E3 were the reverse of those for DOC. NOM molecular size and aromaticity (E2:E3) was best predicted by antecedent 30-day rain, followed by 7-day air temperatures and then sampling stage. The top three watershed characteristic predictors for E2:E3 were percent of the basin logged from 1980-2011, followed by percent underlain by metamorphic parent material (wark gneiss) which was nearly tied with average tree age (Figure 16).

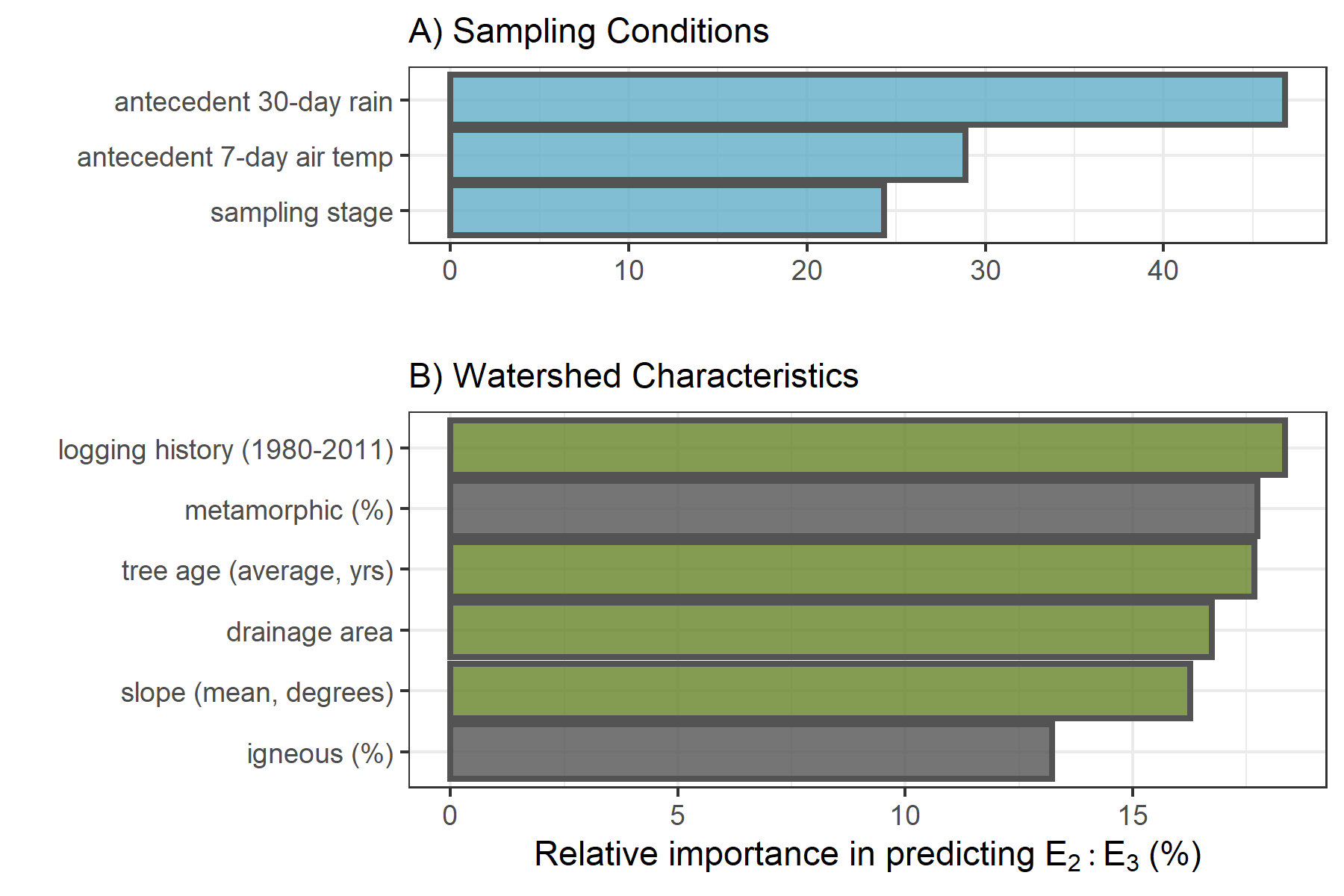


Figure 16: Variable importance for predicting NOM aromaticity & molecular size (E2:E3) across six stream monitoring sites in the Leech River watershed. Variable importance measures were extracted from *randomForest* (in R) and determined by mean decrease in prediction accuracy (increase in the mean square error) in the absence of the predictor variable (i.e. type 1 importance measure).

#### Predictors in relation to NOM concentration and character

When all variables were assessed together, sampling conditions were relatively more important than watershed characteristics for predicting NOM concentration and character across the Leech WSA monitoring sites (Appendix, Figure 40).

##### Sampling conditions

The ranking of relative importance for sampling conditions were in reverse order for DOC (stage, temp, rain) and E2:E3 (rain, temp, stage). Like DOC, SAC254 was influence most by sampling stage; and like E2:E3, SAC254 was more sensitive to antecedent wetness than antecedent air temperatures. The relative sensitivity of NOM concentration to stream stage points to hydrologic mass transport, whereas the sensitivity to antecedent wetness points to the importance of hydrologic connectivity between the terrestrial landscape and stream system for NOM molecular character.

###### Sampling stage

Stream stage at the time of sample collection was the most important condition predictor variable for DOC and SAC254. Concentration of DOC increased with increasing stage up to approximately 60% of maximum sampling stage at each site. The same relationship was seen (albeit more dramatically) for SAC254 (Figure 17). This indicates that as stream levels increased, so too did NOM concentration and aromaticity. This was expected based on results in Chapter 3 that showed higher concentrations of more aromatic NOM in Vertical Rack samples compared to (inter-event) Grab samples. The threshold level (~0.6 of peak stage) points to a change where increasing stage led to dilution rather than mobilization.

Stream stage was the least important sampling condition for predicting E2:E3 (below antecedent rain and air temperature), and the two variables were inversely related. This pattern supports results of stage in relation to SAC254 and shows that as stream stage increased across the Leech WSA, so too did NOM molecular size & aromaticity. Like SAC254, E2:E3 and stage had a weaker relationship around 60-75% of maximum stage (indicated by the greater uncertainty in the loess trend line, Figure 17) which indicates a point at which aromaticity no longer increased with stage. While there were fewer high-flow samples collected compared to mid-range and low-flow samples, this apparent sampling stage threshold may indicate depletion of a terrestrial NOM aromatic pool, or could suggest that the most aromatic sources had reached peak connectivity to the streams and NOM could get no more humic in nature.

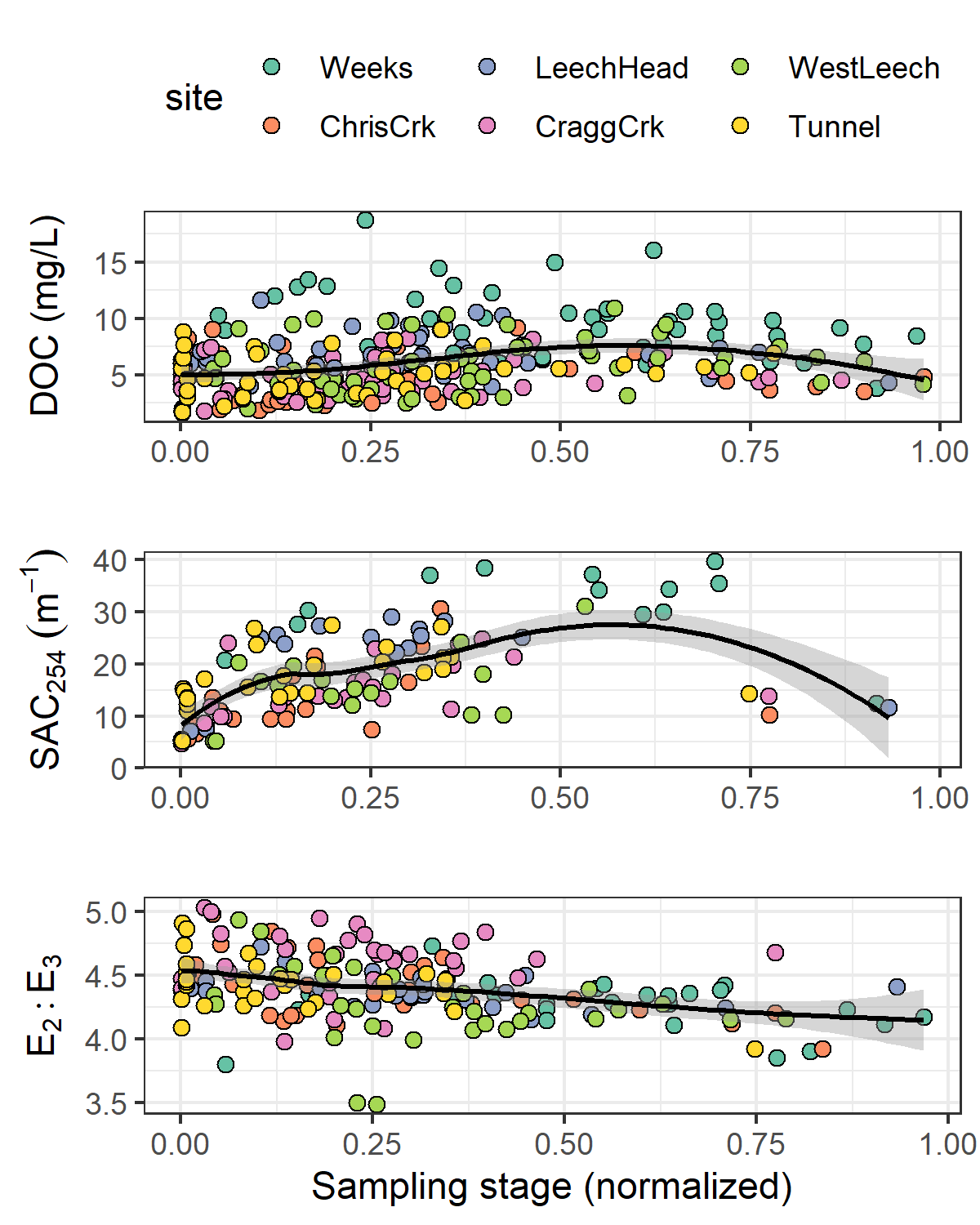


Figure 17: Sampling stage as a predictor for NOM concentration (DOC) and character (SAC254 & E2:E3). Each plot includes a loess trend line.

###### Antecedent 7-day air temperature

Antecedent 7-day air temperature (mean air temperature for the 7-day period prior to sample collection) ranked second most important for predicting DOC and E2:E3 and was the least important (third) sampling condition for predicting SAC254. The concentration, reactivity, aromaticity and molecular weight of NOM increased with antecedent air temperature up to approximately 10°C (Figure 18). DOC and 7-day temperature were positively related in the wet season with no obvious relationship in the dry season (Appendix *XXX*, Figure 41) and too few spectral samples were collected in the dry season to shed light on seasonal trends in NOM character with antecedent temperatures.

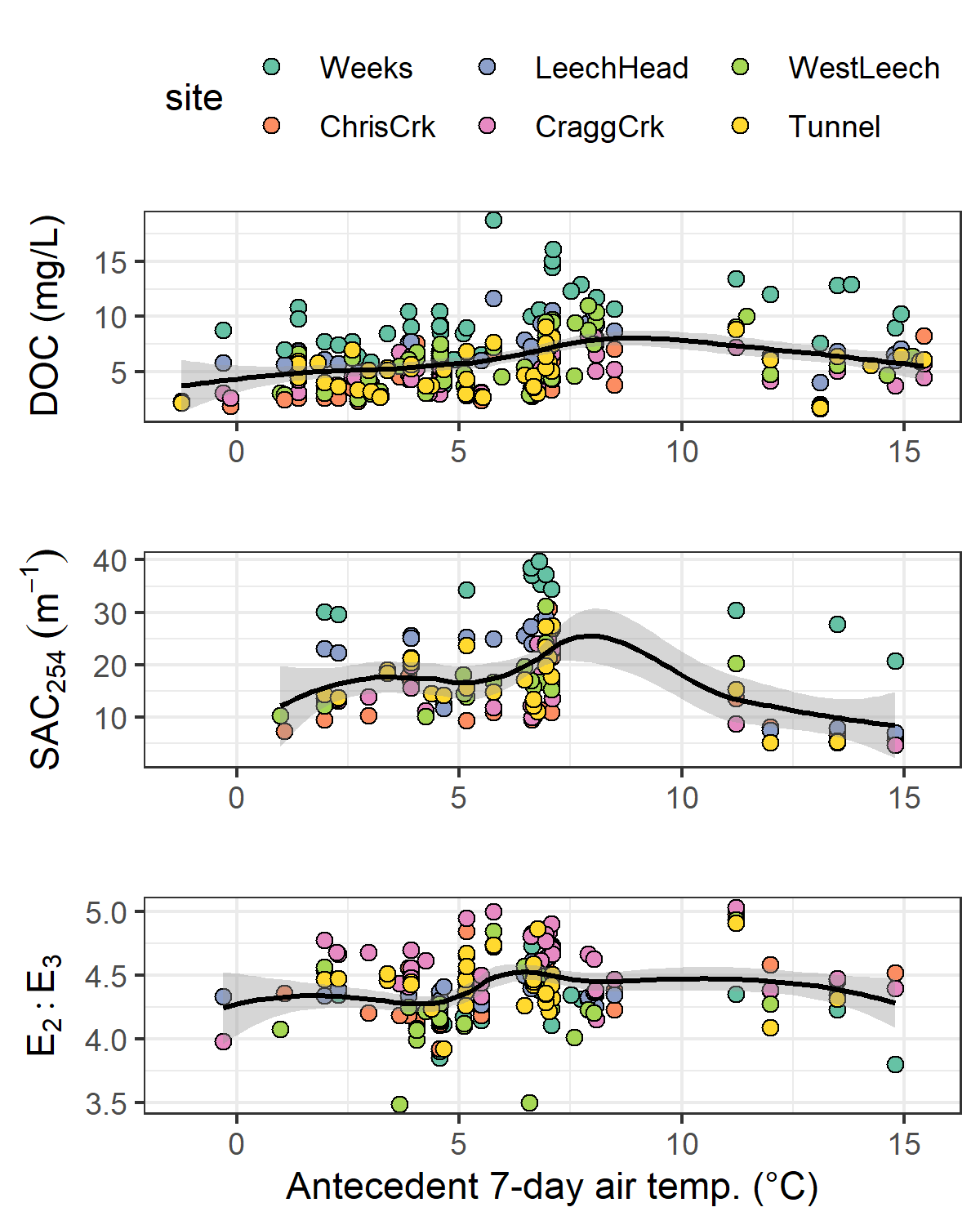


Figure 18: Antecedent 7-day air temperature as a predictor for NOM concentration (DOC) and character (SAC254 & E2:E3) across six monitoring sites in the Leech watershed. Each plot includes a loess trend line.

###### Antecedent 30-day rain

Antecedent 30-day cumulative rain (i.e. antecedent wetness) was the sampling condition calculated to have the greatest relative importance in predicting E2:E3, second-most importance for SAC254 and was the least important sampling condition for predicting DOC concentration. As antecedent 30-day rain increased to approximately 150 mm, there was an apparent decrease in NOM aromaticity and molecular size (E2:E3 increased); above ~150 mm (~25% of maximum antecedent 30-day rain), NOM aromaticity and molecular size increased. At approximately 300 mm of 30-day antecedent rain, NOM aromaticity and molecular size appeared to become less variable (Figure 19).

Fewer sample results were obtained for SAC254, with a data gap for samples collected between 200 mm and 400 mm of 30-day antecedent rain; therefore, the pattern of SAC254 with antecedent wetness was less precise (large uncertainty, Figure 19). For the available data, rain accumulation in the 30 days prior to sample collection was positively related to SAC254 up to approximately 150 mm and SAC254 was relatively low with antecedent 30-day rain above 400 mm. Though the relationship was non-linear, there was a slight initial increase in DOC with increasing antecedent wetness, followed by a gradual decline in DOC with increasing accumulated rain (Figure 18). Lower order streams (Weeks, Chris crk, Leech-head) showed a steeper decline in DOC with increasing 30-day antecedent rain (site-specific plot in Appendix *XXXXX*, Figure 42).

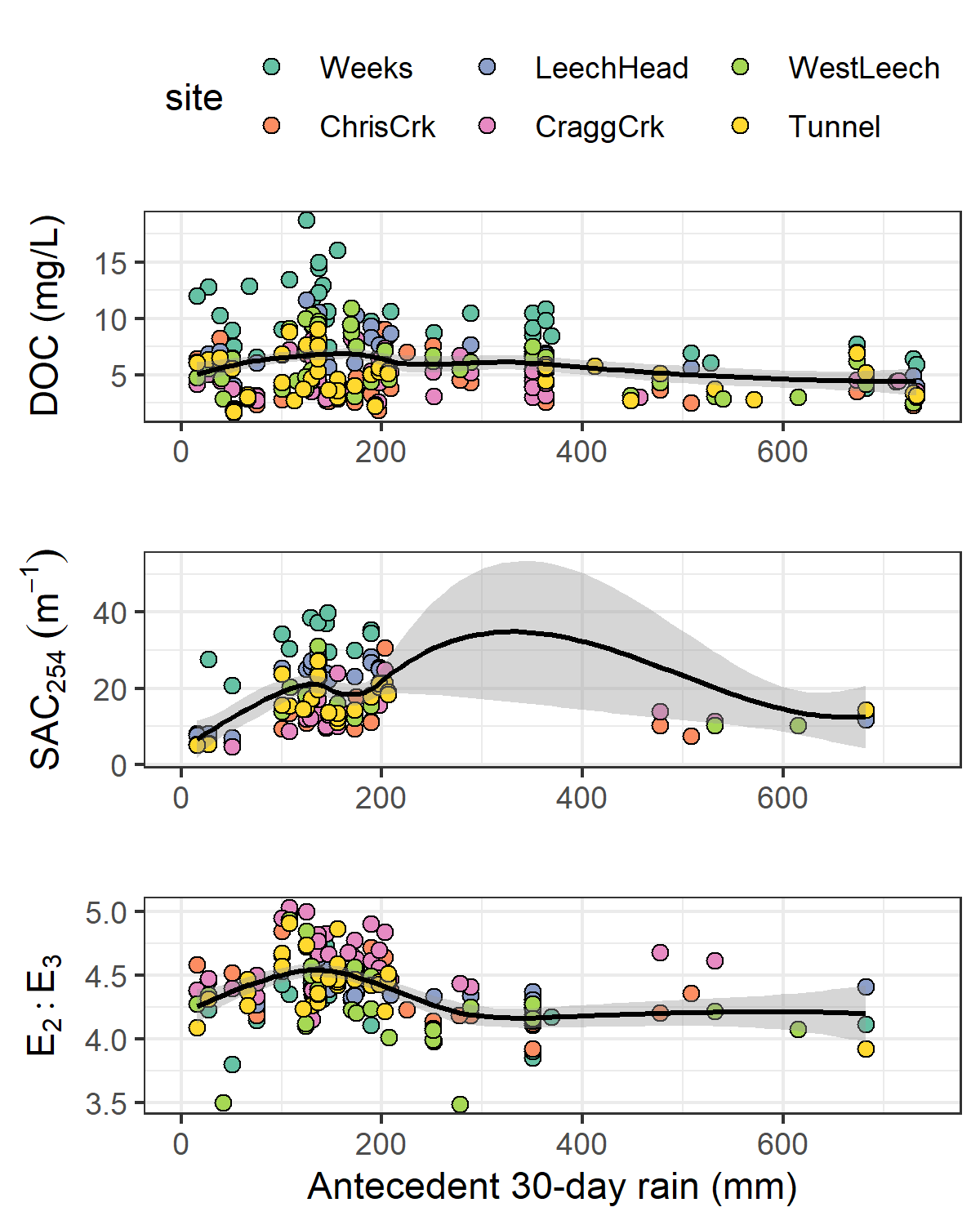


Figure 19: Antecedent 30-day rain as a predictor for NOM concentration (DOC) and character (SAC254 & E2:E3) across six monitoring sites in the Leech watershed. Each plot includes a loess trend line.

These patterns of antecedent wetness with DOC and E2:E3 suggest that smaller, more aliphatic NOM was flushed through the Leech system when antecedent wetness was low (i.e. near the onset of the wet season). As rains continued, streams had greater connection to more aromatic NOM source pools (humic material, possibly from upland sources). The slight decline in aromatic DOC with antecedent wetness (above approximately 350 mm in 30 days) suggests exhaustion of humic NOM source pools and a shift to dilution. These results largely support seasonal characteristic shifts described in Chapter 3, which follow prediction of the River Continuum Concept.

##### Watershed characteristics

Of watershed characteristics, the top predictors for NOM concentration and character were a combination of mean sub-basin slope, metamorphic parent material (percent of basin), logging history (percent of basin harvested 1980-2011), tree age (average) and/or drainage area.

###### Sub-basin slope

Sub-basin slope ranked most important as a predictor for DOC and SAC254. Across the six monitoring sites, the relationship between slope and these two variables was not linear but in general, lower sub-basin slope was related to higher DOC. A pattern appeared, where the six monitoring basins were grouped in pairs with two sites each between mean slopes of 9°- 10°, 10°- 11° and 11°- 12°. Within these pairs, the basin with lower slope had slightly greater NOM aromaticity, reactivity and concentration (Figure 20). Slope was negatively correlated to percent wetland (-0.73), and wetlands are known to be linked with high DOC concentrations and greater NOM aromaticity. Indeed, SAC254 indicated higher aromaticity in sub-basins with lower mean slope (Figure 20). While slope was not ranked as a strong predictor for E2:E3, the relationship between the two largely agreed with results for SAC254.

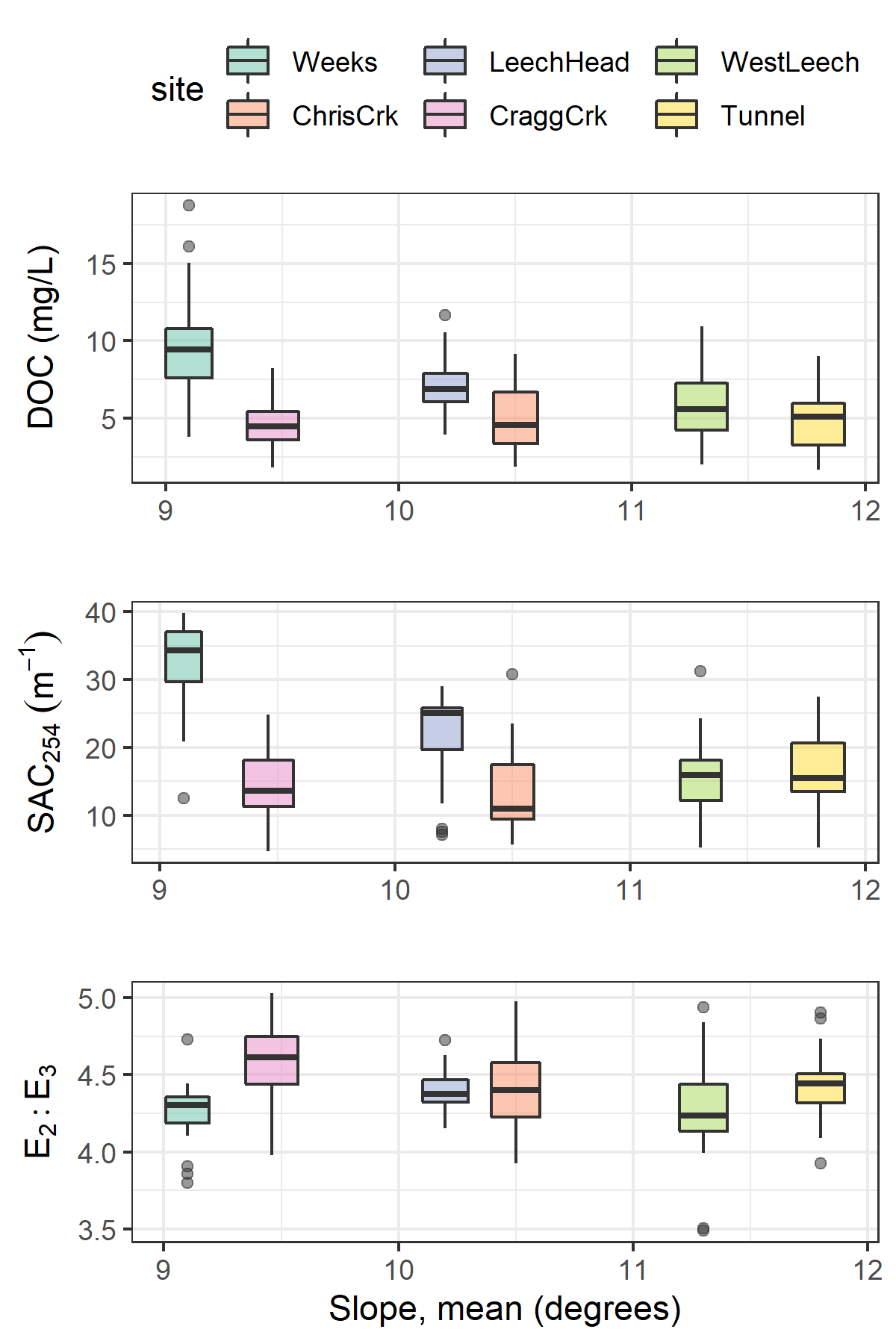


Figure 20: Sub-basin mean slope as a predictor for NOM concentration (DOC) and character (SAC254 & E2:E3) across six monitoring sites in the Leech watershed.

###### Parent material (metamorphic wark gneiss)

The percent of wark gneiss underlying each sub-basin was ranked as the second most important watershed characteristic for predicting each of DOC, SAC254 and E2:E3. Wark gneiss was inversely correlated (-0.92) to the metasedimentary Leech River Formation.

For the five sites that were underlain by some proportion of wark gneiss, stream DOC concentrations decreased with increasing percentage of this metamorphic surficial deposit. While more metamorphic material corresponded to lower DOC, the absence of wark gneiss in West Leech sub-basin did not correspond to the highest DOC concentration (Figure 21). NOM character had a similar pattern to concentration, where greater metamorphic parent material corresponded to lower NOM aromaticity and reactivity (SAC254), but its absence did not correspond to the sub-basin with the lowest SAC254 (Figure 21). E2:E3 showed that greater metamorphic parent material corresponded to lower NOM aromaticity and molecular weight (higher E2:E3 values) and, in contrast to DOC and SAC254, the sub-basin with the greatest molecular weight NOM (lowest mean E2:E3) did correspond with the absence of wark gneiss (Figure 21).

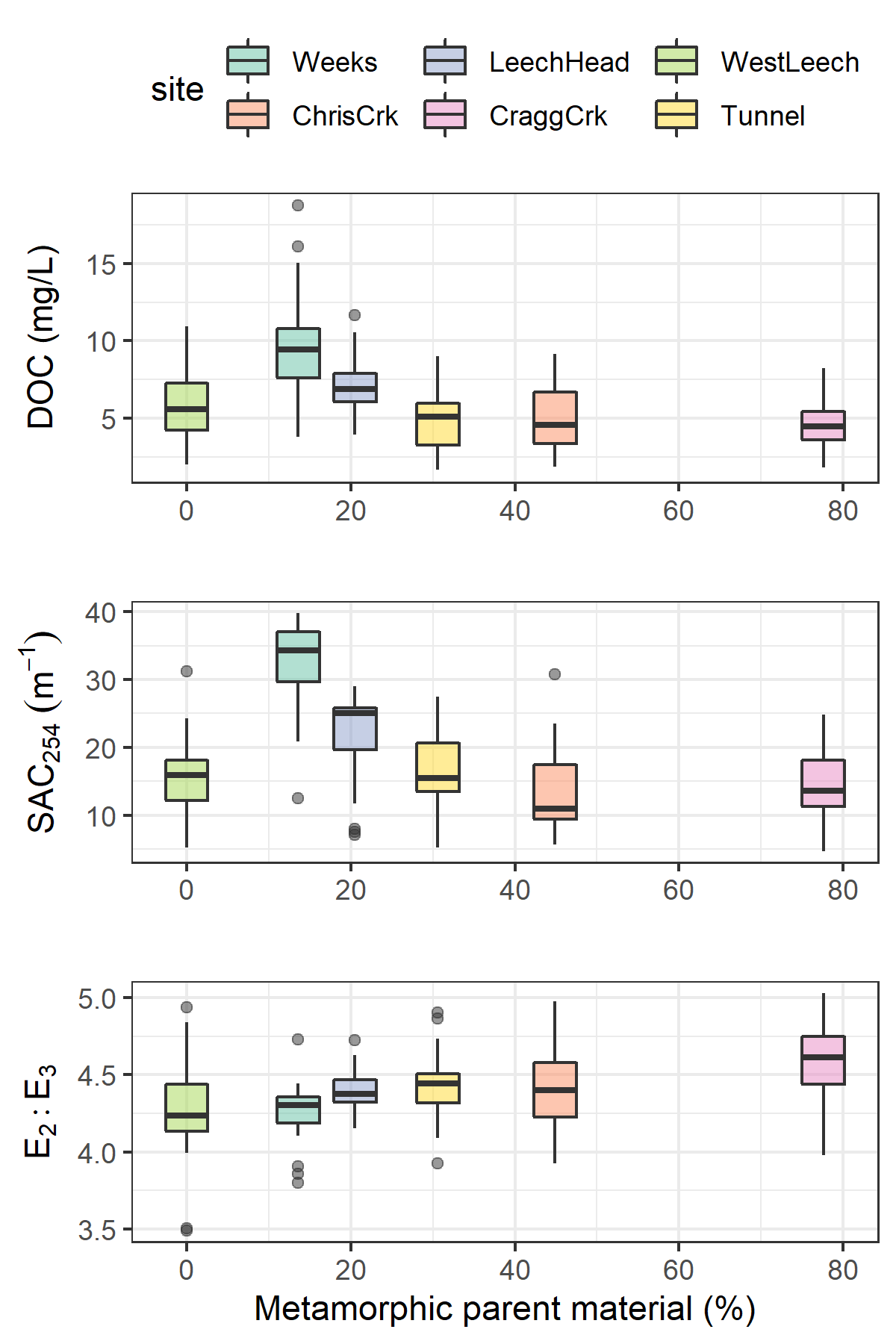


Figure 21: Percent of each sub-basin underlain by wark gneiss (metamorphic parent material) as a predictor for NOM concentration (DOC) and character (SAC254 & E2:E3) across six monitoring sites in the Leech watershed.

###### Logging history and mean tree age

Logging history (percent of basin harvested from 1980-2011) was ranked first among watershed characteristics for predicting E2:E3 and third for predicting SAC254. Average tree age was third in variable importance for predicting E2:E3 and DOC. Tree age and logging history were examined together because average tree age was partially due to the history of forest harvest.

Sub-basin area harvested between 1980 and 2011 was positively related to E2:E3 at four of the six sub-basin sites. West Leech, Weeks, Leech Tunnel, and Cragg crk sub-basins showed decreasing NOM molecular weight and aromaticity (increasing E2:E3) with increasing area harvested. However, the sub-basins of Chris crk and Leech-head (just downstream) had higher harvested area (ergo younger trees) and did not match this pattern of less aromatic, lower molecular weight NOM with greater harvest history (Figure 22). The sites with more aromatic and higher molecular weight NOM were not clearly linked to average tree age (Figure 22).

Chris crk and Cragg crk sub-basins had the youngest and oldest trees, respectively; yet, these two sites had very similar averages for NOM concentration and character, though Chris crk had slightly higher molecular weight NOM. While Cragg crk sub-basin had, on average, the oldest trees (likely due to old growth maintained on Horton Ridge), it was West Leech that was the least-harvested sub-basin of the six monitoring sites. Cragg crk and West Leech had similar mean DOC and SAC254 values but mean E2:E3 was different (***statistically??***); Cragg crk had greater molecular weight NOM than West Leech (but their error bars overlapped entirely). It is possible that the GIS data for forest harvest and tree age were missing some details, or that the average tree age is not the best indicator.

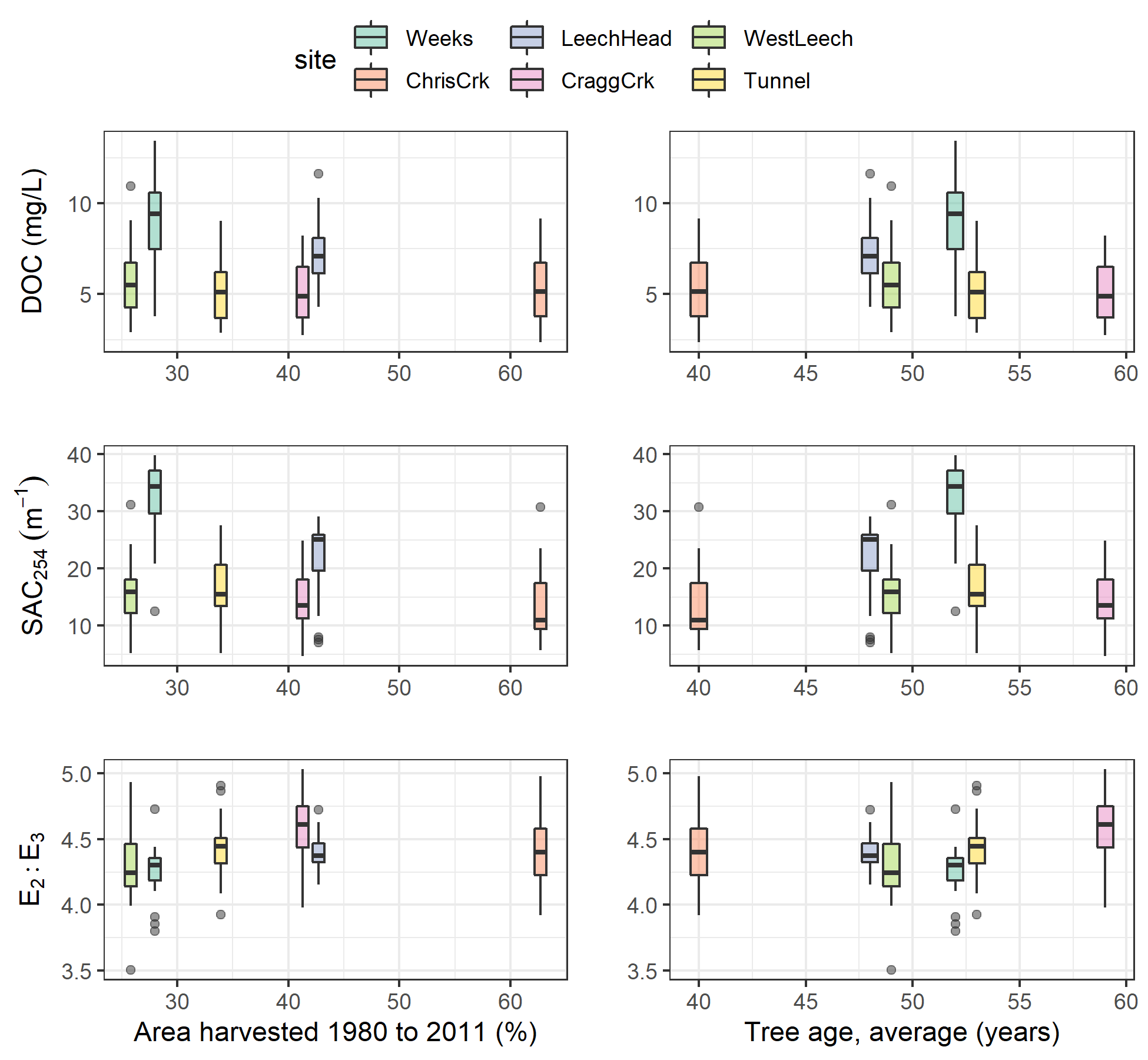


Figure 22: Sub-basin mean tree age as a predictor for NOM concentration (DOC) and character (SAC254 & E2:E3) across six monitoring sites in the Leech watershed.

##### Variable importance Summary

Of the watershed characteristics examined, the percent of each basin with metamorphic wark gneiss parent material appeared to be the most relevant and was a common predictor variable to each of DOC, SAC254 and E2:E3. The absence of wark gneiss was not related to the site with highest DOC, despite DOC decreasing with increasing metamorphic representation. Because subsurface materials were largely cross-correlated, the observed patterns with wark gneiss may point to more nuanced relationships that couldn’t be teased out in this RF analysis because of cross-correlations with other predictor variables. It could be that what is observed as a negative relationship with wark gneiss is a positive relationship between NOM and another parent material, probably the schist-like Leech River Formation.

Across the Leech WSA, warm and wet conditions were an important driver for NOM concentration and character. Antecedent rain was positively related to DOC concentrations and sample aromaticity up to a threshold of approximately 150 mm of 30-day antecedent rain, at which point NOM concentration and aromaticity/reactivity/molecular weight declined with increasing wetness. This suggested a threshold at which dilution overtook NOM transport. In a similar way, NOM aromaticity increased with sampling stage up to approximately 60-75% of maximum stage at each site.

#### Warm and wet: seasonal patterns and rain events

Moisture and temperature, important to each of the three predictant variables, corresponded to different seasons. Given the evidence of strong seasonality for NOM character (Chapter 3), a comparison between wet and dry season sample results was used to assess spatial difference between seasons for DOC, SAC254 and E2:E3.

During the dry-season, SAC254 and E2:E3 indicated that Weeks crk had greater aromaticity compared to the other sites. Aside from the highly aromatic character of NOM observed at Weeks, E2:E3 suggested that aqueous NOM in the dry-season had increasing aromaticity from upstream to downstream sites; while this quotient can also be indicative of molecular size (by an inverse relationship), the simultaneous decrease in DOC concentration indicates that molecular weight was not increasing downstream (Figure 23). SAC254 showed a similar increase in aromatic character from headwaters to mouth during the dry season, albeit the pattern was less pronounced and there was a larger difference between Chris crk and Leech-head in SAC254 than in E2:E3. These results show that dry season aqueous NOM increased in aromaticity from upstream to downstream sites.

The dry season pattern of decreasing DOC from low- to high-order streams was apparent in the wet season also. But the wet season spatial pattern of E2:E3 was almost the reverse of what was seen in the dry season. While dry season samples’ E2:E3 showed increasing aromaticity from upstream to downstream, the wet season E2:E3 showed decreasing aromaticity from head to mouth. These seasonal differences support a shift in NOM source material, and/or a shift from NOM processing in the dry season and transport/dilution effects in the wet season.

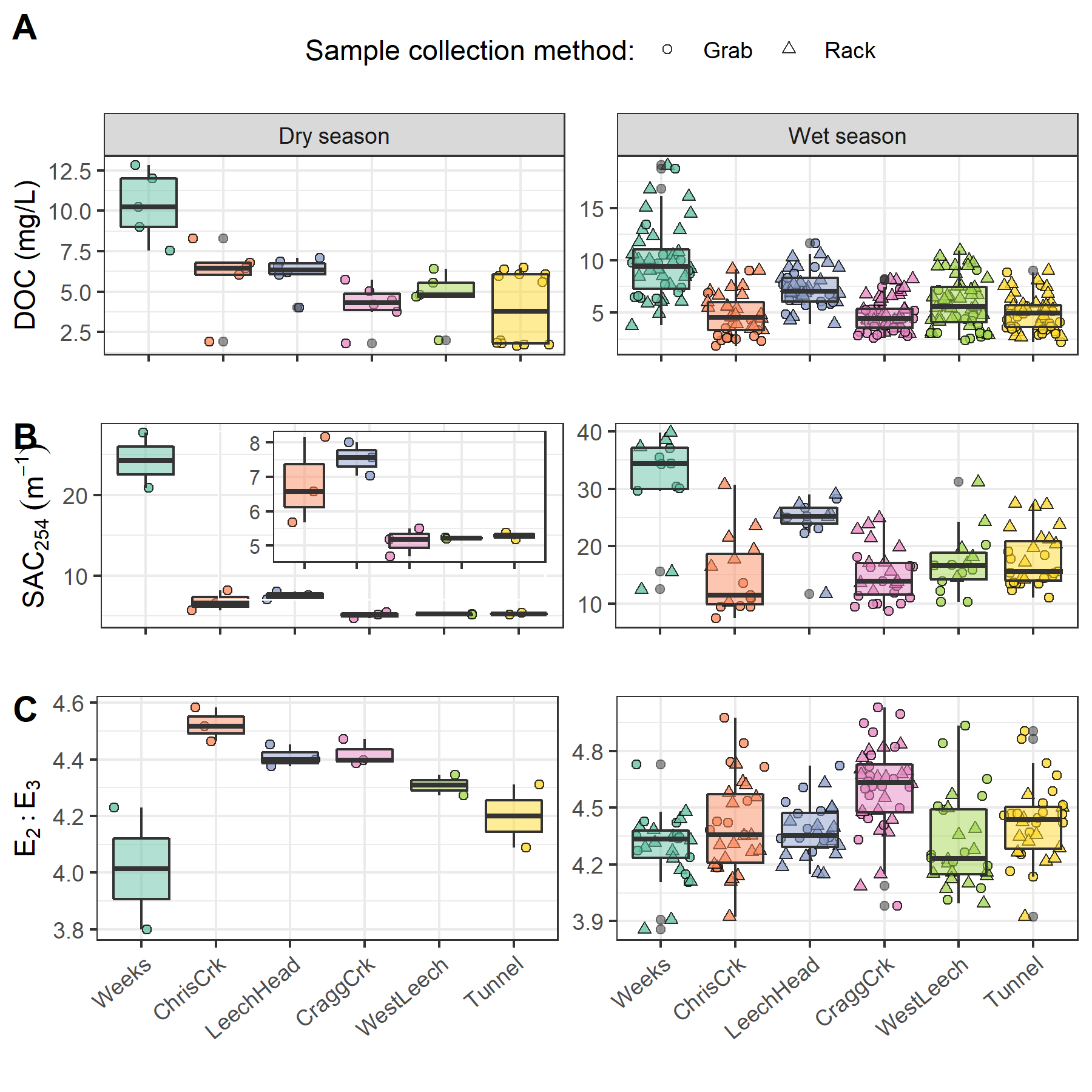


Figure 23: Sample results over the dry and wet seasons for [A] dissolved organic carbon concentrations (DOC), [B] specific absorbance coefficient at 254 nm (SAC254, m-1), and [C] the quotient of SAC250/SAC365 (E2:E3). Note that each season has a different y-scale to clarify patterns. The inset in panel B is an enlargement of SAC254 for all sites except Weeks crk.

##### Rising stage and aqueous NOM dynamics

Antecedent moisture and/or sampling stage was important for each of the three predictant variables evaluated though RF VIM. Across the Leech WSA, the highest DOC occurred in the earliest events of each wet season (events 1 & 9, Table 15), while SAC254 peaked later than DOC during the third major event in the 2019/2020 wet season (event 11, Table 15). E2:E3 indicated maximum aromaticity and molecular size occurred during the last major events, at the end of each wet season (events 3 & 17, Table 15).

Table 15: Rain Events (Threshold of 50 mm and 14-hour Inter-event Period) Defined by Vertical Rack Sample Collection Across the Six monitoring Sites of the Leech WSA

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Major event no. | Start Date | Duration (days) | Rainfall (mm) | Intensity (mm/24-hr) | Samples collected | mean DOC (mg/L) | mean SAC254 | mean E2:E3 |
| 1 | 2018-10-27 | 6.1 | 124.4 | 20.2 | 19 | 9.09 | 29.13 | 4.34 |
| 2 | 2018-11-03 | 0.9 | 54.8 | 60.5 | 18 | 6.40 | 25.15 | 4.20 |
| 3 | 2018-11-25 | 3.6 | 156.1 | 43.9 | 34 | 6.23 | 27.93 | 4.18 |
| 4 | 2018-12-09 | 4.9 | 205.1 | 41.6 | 18 | 5.88 |  |  |
| 7 | 2019-01-02 | 4.2 | 227.6 | 54.2 | 22 | 4.68 |  |  |
| 8 | 2019-01-17 | 3.0 | 68.7 | 22.8 | 3 | 2.97 |  |  |
| 9 | 2019-09-12 | 3.1 | 58.4 | 18.8 | 14 | 9.97 | 16.75 | 4.82 |
| 10 | 2019-10-15 | 6.4 | 136.2 | 21.3 | 44 | 6.73 | 19.74 | 4.53 |
| 11 | 2019-11-15 | 2.3 | 67.6 | 29.1 | 28 | 6.44 | 22.84 | 4.51 |
| 12 | 2019-12-10 | 3.1 | 70.4 | 22.8 | 25 | 5.30 | 18.90 | 4.51 |
| 13 | 2019-12-18 | 4.1 | 112.1 | 27.7 | 1 | 4.46 | 16.55 | 4.61 |
| 15 | 2020-01-02 | 5.7 | 180.0 | 31.7 | 3 | 5.13 | 15.78 | 4.40 |
| 16 | 2020-01-18 | 9.9 | 238.3 | 24.0 | 8 | 4.11 | 11.42 | 4.43 |
| 17 | 2020-01-30 | 1.9 | 208.8 | 111.3 | 5 | 4.46 | 12.82 | 4.15 |
| 18 | 2020-02-05 | 3.4 | 75.9 | 22.4 | 10 | 3.79 | 9.61 | 4.48 |

Streams across the Leech WSA responded harmoniously to precipitation with synchronous changes in stage. Over the full sixteen-month study period, stage changes at the six monitoring sites were synchronized and confirmed to be congruent with high confidence (p-value < 0.001, based on 1050 randomizations) via Kendall’s coefficient of concordance (Kendall’s W = 0.9721) and Spearman’s ranked correlation (ρ = 0.9666). Rates of change in stream response were calculated for each site to determine the fastest and slowest times to peak stage and relative magnitudes of stage change (Table 16). As expected, rate of change in stage was greatest at the highest order stream, the Tunnel site, and smallest at Weeks crk.

Table 16: Summary of Stream Response to Rain Events Across Six Monitoring sites in the Leech Water Supply Area.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| site | shortest time to peak stage (hr) | longest time to peak stage (hr) | smallest change in stage (cm) | largest change in stage (cm) | minimum rate of stage change (cm/hr) | maximum rate of stage change (cm/hr) |
| Weeks | 16.5 | 159.8 | 9.9 | 47.9 | 0.17 | 0.6 |
| ChrisCrk | 15.0 | 192.3 | 5.3 | 44.6 | 0.03 | 2.7 |
| LeechHead | 33.0 | 167.7 | 22.1 | 57.1 | 0.29 | 1.3 |
| CraggCrk | 18.2 | 125.2 | 32.7 | 97.3 | 0.26 | 5.4 |
| WestLeech | 25.8 | 122.8 | 36.7 | 70.2 | 0.33 | 1.4 |
| Tunnel | 18.2 | 123.8 | 29.0 | 135.3 | 0.45 | 7.4 |

To determine if a similar harmony was present for fluctuations in DOC or spectral properties, the proportion of common DOC and stage extrema samples were calculated (Table 17). A proportion of 1 indicates perfect agreement between samples of extreme DOC and sample stage, and zero indicates complete asynchrony between DOC and stage highs and lows. Most samples showed congruence between highs and lows of DOC with stage. West Leech was the only site which had a greater proportion of common maxima compared to minima, all other sites had more frequent occurrence of low DOC occurring at low stage that they did of high DOC occurring at high stage.

Table 17: Proportion of Samples for Which Peak Dissolved Organic Carbon (DOC) was Collected at the Highest Sample Stage.

|  |  |  |
| --- | --- | --- |
| Site | Proportion of common maxima | Proportion of common minima |
| Weeks | 0.5882 | 0.6471 |
| ChrisCrk | 0.8000 | 0.8000 |
| LeechHead | 0.8889 | 0.8889 |
| CraggCrk | 0.8000 | 0.9000 |
| WestLeech | 0.8636 | 0.7727 |
| Tunnel | 0.7647 | 0.9412 |
| all sites | 0.7895 | 0.8246 |

Four of the six site’s samples showed majority, but not absolute, agreement in extremes of DOC-stage relationships; Chris crk and Leech-head each had the same proportion of common maxima and minima. In general DOC concentrations were lowest at the beginning of events and increased with a rise in stage (Figure 24). Early in the wet season, peak stage was more associated with minimum DOC, and later in the wet season peak stage was more associated with peak DOC. These results suggest dilution of NOM early in the wet season, and enrichment later in the wet season.

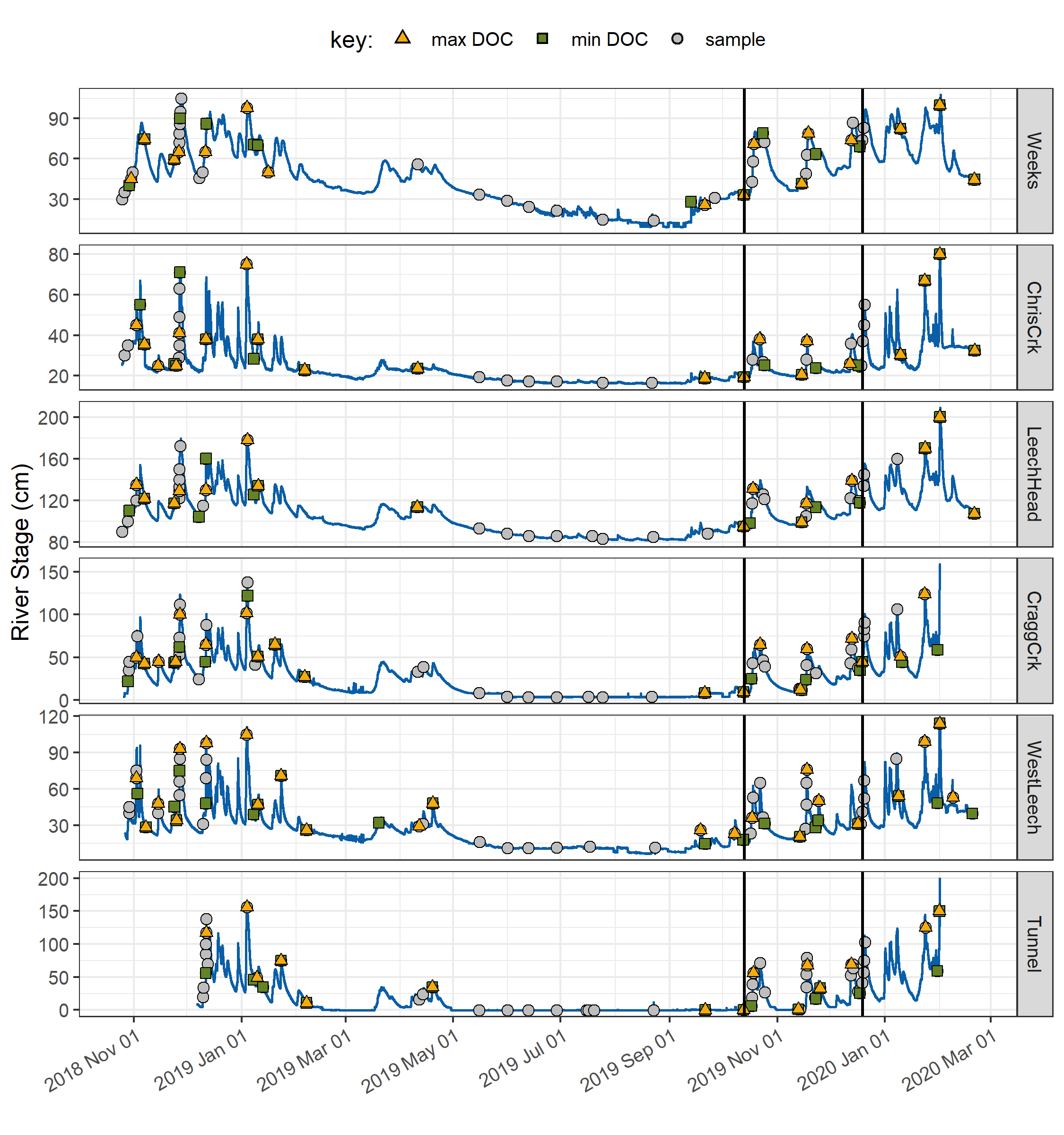


Figure 24: Stage and samples collected, highlighting samples with maximum and minimum DOC concentrations for each rain event and collection period. Black vertical lines indicate a subset of samples that were assessed more closely.

For events where several samples were collected, the change in DOC concentration varied from a little more than 1% to nearly 100% (Table 18). The smallest event-based concentration change occurred at Leech head (1.2%) and the Tunnel (1.4%). The largest event-based change in concentration was observed at West Leech (94.6%), and the second largest change occurred at Chris crk (82% change in DOC).

Table 18: Summary of DOC Changes Within Stormflow Response to Precipitation Events Across the Leech WSA (Samples from Wet Season Only).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| site | lowest DOC in stormflow (mg/L) | highest DOC in stormflow (mg/L) | smallest change in DOC (mg/L) | largest change in DOC (mg/L) | smallest difference in DOC during stormflow (%) | largest difference in DOC during stormflow (%) |
| Weeks | 6.1 | 16.1 | 1.0 | 6.4 | 9.6 | 53.2 |
| ChrisCrk | 2.3 | 9.2 | 1.2 | 3.9 | 26.1 | 82.0 |
| LeechHead | 5.7 | 10.3 | 0.1 | 2.0 | 1.2 | 29.5 |
| CraggCrk | 3.0 | 8.2 | 1.2 | 3.2 | 28.8 | 67.5 |
| WestLeech | 2.5 | 10.9 | 0.1 | 5.8 | 4.5 | 94.6 |
| Tunnel | 3.4 | 5.9 | 0.0 | 2.0 | 1.4 | 42.5 |

At the Tunnel, Cragg crk and Chris crk, NOM concentration and aromaticity generally increased with stage; while at West Leech, aromaticity was inversely related to stage. Weeks crk maintained the most consistent NOM aromaticity and molecular weight (E2:E3) with changing stage (Figure 25).

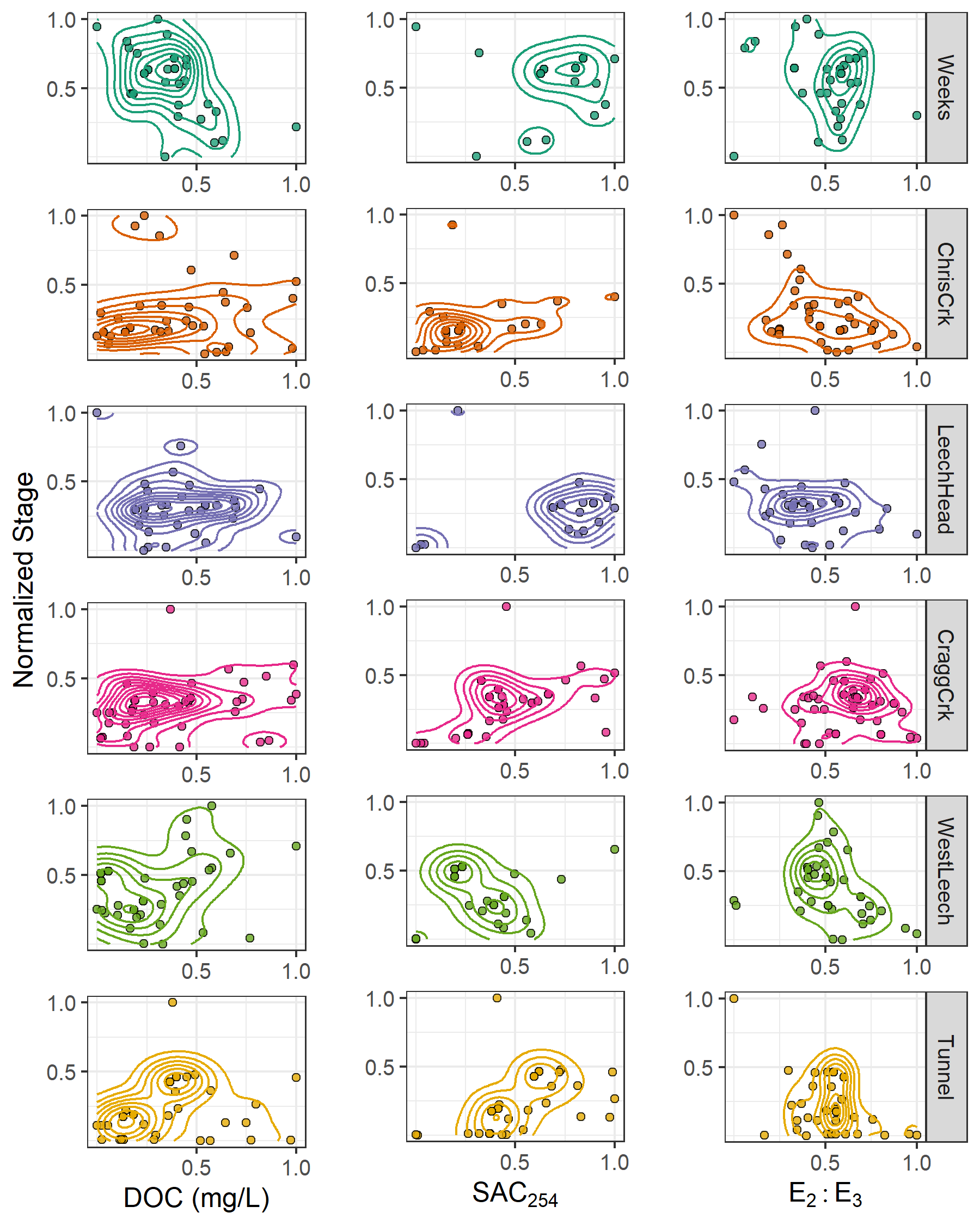


Figure 25: Relationships between river stage and sample NOM concentrations and character, where connected lines show data density. Data for each variable were normalized (min-max normalization) for comparison of relative scales in each relationship between sites.

#### Hysteresis of NOM with antecedent wetness

Events 10 and 11 were relatively well represented by sampling at each of the six sites and were found to be representative of other rain events as well. Wilcoxon tests showed events 10 and 11 were not statistically different from the other sampling events, with the same mean rain durations (p = 0.837), rain amounts (p = 0.732), rain intensity (p = 0.549), DOC (p = 0.512), SAC254 (p= 0.218) and E2:E3 (p = 0.148). Samples were collected across the rising limb, near peak-flows, and at least one sample on the falling limb and at low flow (Figure 26).

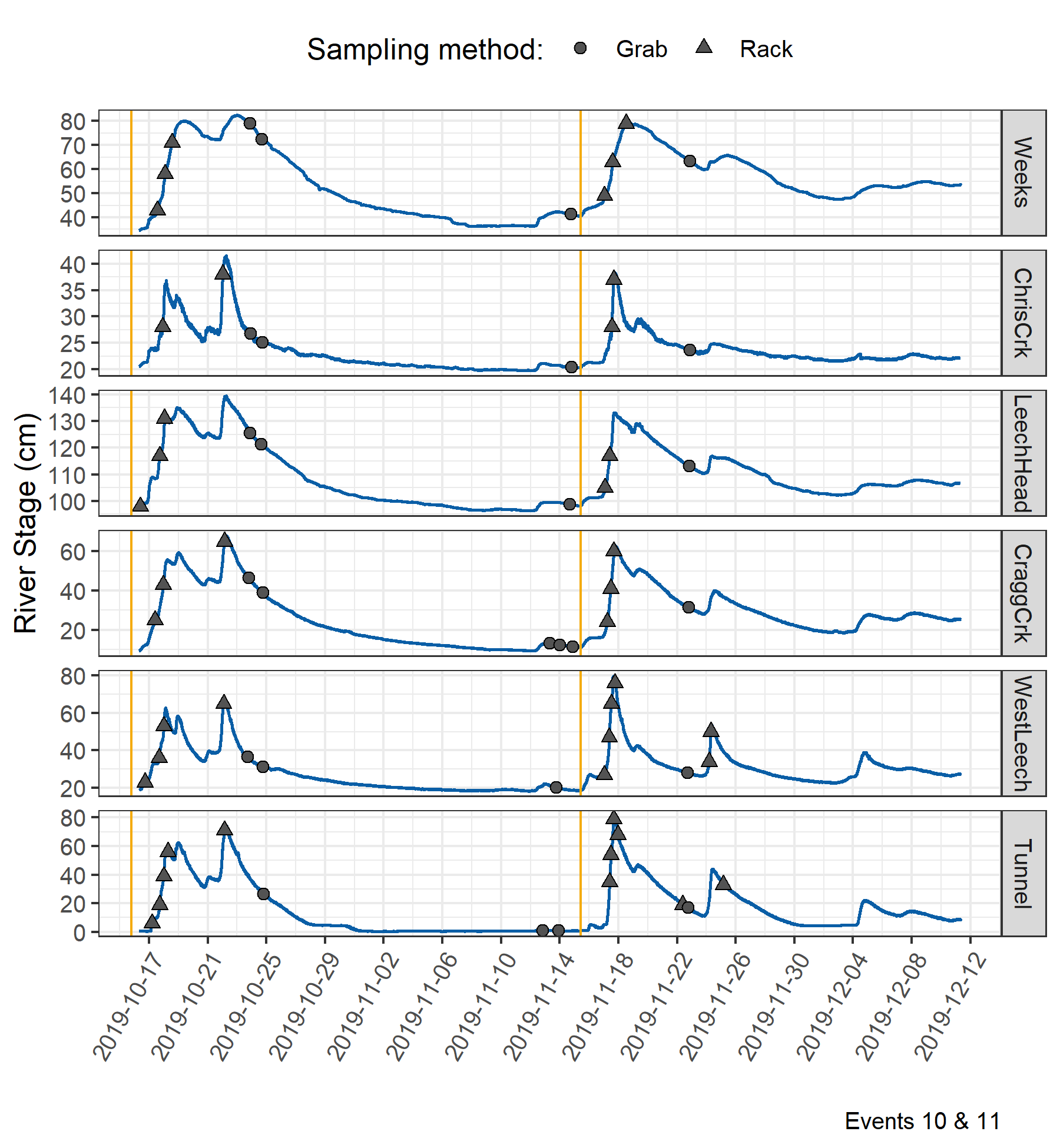


Figure 26: Stage and sample collection points at six Leech watershed monitoring sites during rain events 10 and 11 (early fall 2019). Vertical yellow lines indicate the start of each rain event (event 10 and 11, respectively). Results of samples in each event were compared to antecedent 30-day rain to assess hysteretic behaviour of NOM with catchment wetness.

Sample DOC concentrations from rain events 10 and 11 were plotted with their corresponding antecedent wetness (30-day rain) to explore possible hysteresis (Figure 27). The relationship of DOC to antecedent wetness within a hydrologic event was interpreted in the same manner as a concentration-discharge (C-Q) relationship, where rapid flushing and dilution are indicated by a hysteresis loop with clockwise rotation, and counterclockwise rotation points to delayed material delivery and stream enrichment. Event 10 was the second measured rain event in the 2019/2020 wet season and had twice as much rain fall as event 11 (136 mm in event 10 compared to 68 mm in event 11, Table 15). Event 10 was longer in duration (6.4 days) compared to event 11 (2.3 days) with slightly lower intensity (21 vs 29 mm/24-hr.). Between these two events, mean DOC concentrations were comparable (6.7 mg/L and 6.4 mg/L). There was a clockwise rotation pattern for DOC with antecedent wetness in event 10 and a counterclockwise rotation in event 11 (Figure 27). At the Tunnel, there was a change in loop direction at the end of each event.

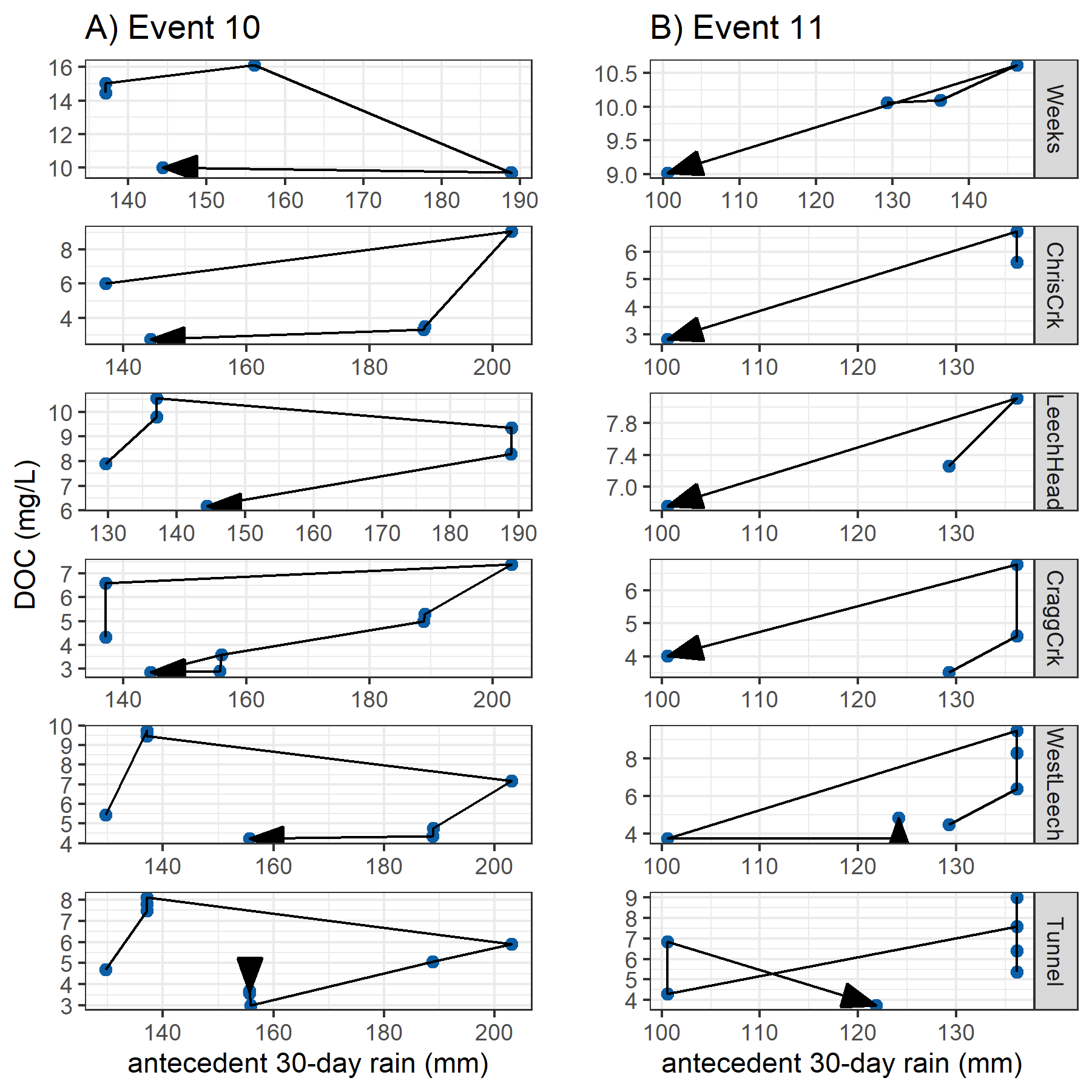


Figure 27: Dissolved organic carbon (DOC) concentrations plotted with antecedent 30-day rain during events 10 and 11 at six monitoring sites in the Leech watershed.

### Discussion

The sampling conditions and watershed characteristics included in Random Forest variable importance measures (RF VIM) showed that across the six monitoring sites in the Leech WSA, stream NOM concentration and character were influenced by warm and wet conditions and the parent material in each basin. No model is perfect, and the relative importance derived from RF VIM in this chapter was limited to the data provided and the assumptions associated with those data. Based on the variables assessed, sampling stage and antecedent wetness ranked among the most important variables for predicting NOM quantity and quality, as did the percent of wark gneiss underlying each basin.

Sub-basin NOM concentration and aromaticity decreased with increasing representation of metamorphic wark gneiss parent material, but its absence did not correspond to the site with highest concentration or aromaticity. The patterns observed between NOM and metamorphic parent material suggests a more complex relationship between watershed characteristics and NOM dynamics. The metasedimentary Leech River formation (argillite metagreywacke and metagreywacke) was inversely correlated (-0.92) to wark gneiss in the monitoring basins and wasn’t included in RF VIM analyses. There may be an effect on aqueous NOM associated with a combination of parent materials and soil development that wasn’t full elucidated in this analysis.

Antecedent wetness is an important watershed driver for NOM export as it influences flow paths and hydrologic connectivity to source material (McMillan et al. [2018](#ref-McMillan2018)). Solute supply and hydrologic connectivity can be interpreted through the magnitude and direction of water quality changes in response to precipitation (Vidon, Wagner, and Soyeux [2008](#ref-Vidon2008); Abbott et al. [2018](#ref-Abbott2018); Creed et al. [2015](#ref-Creed2015); Zarnetske et al. [2018](#ref-Zarnetske2018)). In the Leech WSA, stage changed in response to precipitation and there was an apparent threshold (near 0.75 of peak sampling stage) that suggested a point at which terrestrial aromatic NOM pools had reached peak connectivity to the monitored streams.

A pattern in DOC concentration during rising stage can provide information about NOM source and flux dynamics. Increasing DOC with rising stage suggests NOM transport dynamics were driven by hydrologic connectivity to an unlimited supply of source material. Alternatively, if DOC concentrations decreased with rising stage it’s likely that the NOM source pool was limited (Zarnetske et al. [2018](#ref-Zarnetske2018)). Event-based rising limb changes in DOC concentration varied from a little more than 1% (at Leech-head and the Tunnel) to 95% (at West Leech).

Across the Leech WSA, stream samples with the highest DOC were collected at the highest sampled stage 83% of the time; and there was an 80% co-occurrence of low DOC with low sampling stage. These results indicate an overall pattern of increasing DOC with stage, suggesting that approximately 80% of the time streams were hydrologically connected to a rich supply of NOM. Furthermore, there was a positive relationship between sample aromaticity and stage at the Tunnel, where streamflow from all sub-basins’ is integrated; when coupled with the observed stage-DOC dynamics, this indicates that the supply of NOM was humic in origin approximately 80% of the time.

In three of the six monitoring sites, low DOC was more often collected at low stages than high DOC was at high stages; the opposite was found at West Leech, where concentration and stage had a greater proportion of common maxima compared to common minima. The headwater sub-basin sites of Chris-crk and Leech-head had the same proportions of samples with common maxima and minima of DOC and stage. These patterns changed across the wet season. Peak stage was more often associated with minimal DOC early in the wet season, whereas DOC tended to peak with stage (or close to peak stage) later in the wet season. The observed temporal change in NOM concentration is evidence of early wet-season NOM being limited in supply and subject to rapid flushing, while late wet-season NOM in streams was drawn from a more sustainable pool of source material. The apparent stage and wetness threshold for NOM concentration and character (approximately 75% of maximum stage and about 150 mm of antecedent 30-day rain), suggested a point at which humic NOM sources either reached maximum connectivity with the streams, or that the aromatic source pools were depleted.

The quantity of stream NOM was greatest early in the wet season, with peak concentrations found in the first event-based samples; while NOM aromaticity, reactivity and molecular weight peaked later in the wet season. Coupled with the other results, this indicates that early wet-season rain events exported high concentrations of aliphatic NOM from sources that were likely autochthonous and quickly depleted.

Between the wet and dry season, there were opposite spatial patterns for NOM molecular size and aromatic character. NOM aromaticity increased from upstream to downstream in the dry season, while the wet season showed decreasing aromaticity from head to mouth. The different seasonal patterns in NOM quality support a change in source material and a shift from NOM processing in the dry season to dilution effects in the wet season.

### Summary and future directions

The parent material in each sub-basin was important for NOM dynamics. Although the precise relationship was not completely resolved, it was evident that a greater proportion of metamorphic parent material corresponded to lower concentrations of less aromatic and less reactive NOM in streams.

Approximately 80% of the time during hydrologic pulses in the wet season, monitoring sites were connected to rich supplies of humic NOM. For a shorter time, early in the wet-season, streams exported high concentrations of aliphatic NOM from autochthonous sources that were subject to rapid flushing and were quickly depleted. Streams were hydrologically connected to sustainable pools of allochthonous source material later in the wet season. It appeared that when streams were near 75% of their maximum stage (therefore, approximately 0.75 of peak flow) humic source pools reached maximum connectivity with the streams.

In the future, vertical rack sampling (and matched sample filling-stage with continuous logger stage) could be combined with a rating curve to determine mass transport or loading. Combining methods of vertical sampling racks with standard synoptic grab sampling provided useful time-stamped discrete river samples data that could be adapted and correlated to local rating curves to calculate loads of river material (e.g. nutrients, metals, organics, etc.) exported from monitored catchments. Material transport loads are important for management decisions and comprehensive system understanding.

It was assumed that the arithmetic mean of precipitation measured at Chris Creek and Martin’s Gulch Fire Weather Stations was representative of conditions across the Leech WSA and mean rain data were applied to all monitoring sites (Chapter 2). However, it is entirely likely that the West Leech sub-basin experienced different weather than the rest of the watershed; observations during field work (Appendix *XXXX*) support this possibility. When a weather station is constructed in the West Leech basin, more precise explorations could be conducted with respect to antecedent wetness and water quality dynamics.

With additional weather data (e.g. humidity), it would be interesting to explore whether the variables associated with increased fire risk or mass wasting might also be predictors for NOM dynamics (or any other water quality parameter of interest). The security of drinking water supply quality depends on weather, which is also a major player in other risks to the forested watershed. With greater understanding of the conditions and characteristics that drive water quality dynamics, an index system could be created that incorporates predictors for drinking water supply quality with those for forested watershed threats (e.g. wildfire or windthrow).