Science and Technical Advisory Group:

Defining sediment reference states of New Zealand catchments

Paper Author	NIWA/Cawthron	Classification	**Working draft In confidence**

Meeting date	24 February 2019	Agenda item (number)	6

Paper summary:

This working draft paper presents analyses to achieve the following objectives:

1. Develop “sediment state classification” (SSC) systems for New Zealand rivers
2. Determine reference states for deposited fine and suspended sediment environmental state variables (ESVs; proportion of streambed coverage with sediment <2mm; clarity, turbidy) across segments

The researchers used the following principles to guide the approach with the intention of developing frameworks that are meaningful and achievable for NOF attributes in that they reflect real differences in reference in-stream sediment characteristics and the ecological response to them:

- Balance between generality (simplicity) and sensitivity to changes in sediment status.
- Build on existing river classification systems
- Based on drivers of sediment supply and retention and also observed sediment indicators
- Use a spatial scale reflecting changes in geomorphology and climatology

The analysis produced two sets of SSCs (for deposited and suspended sediment) that group streams by their REC climate, topography, and geology (CTG) variables, which are the primary controls on sediment supply and retention.

Defining sediment reference states of New Zealand catchments

Rationale

The general aim of this analysis was to determine reference states for instream deposited fine sediment and suspended sediment across segments of the New Zealand national river network. For the purposes of this investigation, the reference state of a segment was broadly defined as the levels of deposited and suspended sediment within that segment, on the average through time, assuming minimal urban, agricultural and forestry development within the catchment upstream. The levels of deposited and suspended sediment a segment would experience in its reference state depends on its climatic, topographic and geological context. These factors interact to influence both supply and retention of sediment.

It follows, therefore, that any sediment management objectives—in our case, values delineating the A, B, C and D management bands of the NOF—must take into consideration landscape-scale variability. We require a classification of New Zealand streams such that segments within each class can be assigned a sediment reference state. We require reference states throughout New Zealand for three environmental state variables (ESVs): deposited fine sediment (sediment < 2mm diameter; proportion of streambed covered); turbidity (NTUs); and clarity (m). The specific objectives of this analysis were:

Objective 1. Develop a sediment state classification (SSC) for New Zealand rivers. The SSC will sort New Zealand river segments into groups or 'sediment classes' that have different sediment supply and retention characteristics. As such, the SSC will subdivide the catchments of New Zealand into regions with different sediment supply and retention characteristics.

Objective 2. Within each sediment class, estimate the reference state for each ESV.

Our approach to meeting these two objectives was guided by the following five principles:

1. **The reference state classification should achieve the right balance between generality, hence ease of use, and sensitivity to any change in the sediment status of steams**. If we have too few classes, then streams that naturally have different sediment characteristics are combined in the one class, leading to the situation where reference conditions are biased. Biased reference conditions, in turn, result in management bands that may either (a) not provide the protective and/or restorative direction required, or (b) result in management objectives that are not achievable. By contrast, if we have too many classes then we may yield a classification system that is complicated and impractical to use, with managers having to frequently refer to new sediment management bands as they move among streams within regions. Moreover, the classification system developed herein will be based on data, and so the number of classes will be constrained by the amount of data available to define each class.
2. **The classification should build on existing river classification systems used in New Zealand, particularly those that have been used to inform catchment policy and management**. There is value in building on a classification system that already exists within scientific literature, and that managers and policy makers are already familiar with. By using a familiar classification system we aim to streamline both adoption and use.
3. **The classification should be (a) based on the key geomorphological and climatological variables that drive sediment supply and retention; and (b) also be based on observed deposited and suspended sediment data, hence capture real differences in the sediment**

characteristics of rivers. If key climatological and geomorphological variables that drive sediment supply and retention are used as a basis for our classification, then the classification system will be intuitive to the user. That is, streams in different sediment classes will also have different climatological and/or geomorphological settings. We aim to avoid a classification whereby streams within the one sediment class have obviously contrasting geomorphological or climatological settings—such properties in a classification system may erode confidence in the classification. Equally, we do not wish to have a classification that subdivides streams in an intuitive way yet results in different classes that have indistinguishable sediment characteristics. Therefore, the classification must also be based on real sediment data.

4. **The classification should group stream segments at a spatial grain reflecting likely changes in the geomorphological and climatological variables driving sediment supply and retention**. Spatial grain refers to the spatial resolution of analysis. If we select too fine a grain for analysis we generate a risk of yielding a classification whereby sediment management bands switch back and forth frequently as one moves up- or down-stream. A classification with too coarse a grain would result in whole regions/catchments being grouped together despite the presence of different landscape settings.
5. **Estimates of reference state within all regions of New Zealand should result in NOF management bands—hence management targets—that are achievable**. Reference states of ESVs within each sediment class should not be so stringent that management bands are not achievable. As such, reference state estimates within each SS class need to be representative of the streams within that class as a whole.

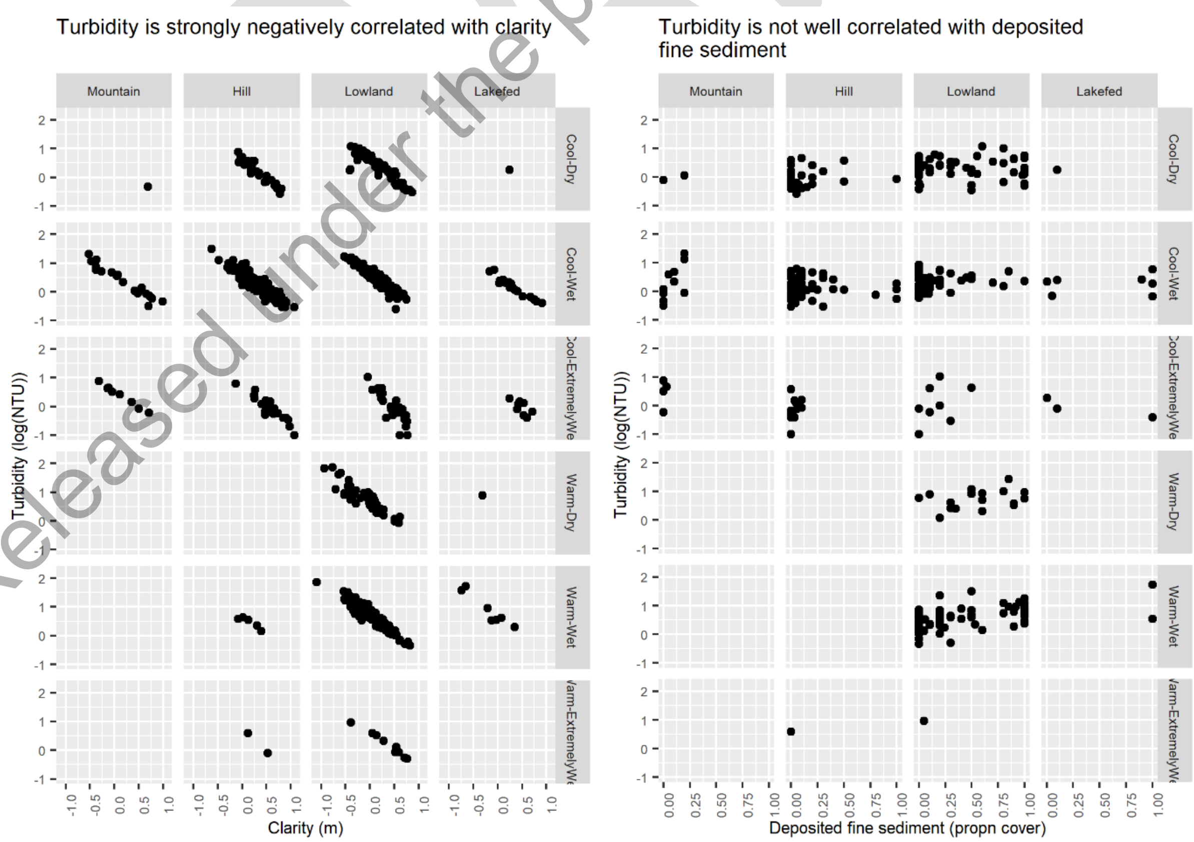

Figure AX.1. Relationships between median turbidity and median water clarity (left plot) and median turbidity and median proportion cover of deposited fine sediment (right plot), within each CTG (Climate-Topography-Geology) class of the New Zealand River Environment Classification (REC). Each point corresponds to a median value from an individual monitoring site (see Methods). Turbidity and clarity data sourced from the National River Water Quality Network, while deposited fine sediment data sourced from the New Zealand Freshwater Fish Database.

Development of the Sediment State Classification

Two SSCs were developed; one for deposited fine sediment (SSC_Dep) and one for suspended sediment (SSC_Sus). A separate SSC for deposited and suspended sediment was deemed necessary since, while turbidity and clarity are strongly correlated within New Zealand river segments, turbidity and deposited fine sediment are not (Fig. AX.1). Given turbidity and clarity are strongly correlated, we used the turbidity data to develop an SSC for suspended sediment.

To satisfy Principles 2, 3a and 4 presented in the Rationale, we used the New Zealand River Environment Classification (REC; Snelder and Biggs 2002) as a basis for our SSCs. Specifically, the first step of developing our SSCs was to group streams by their REC climate, topography and geology values (REC variables: CLIMATE; SRC_OF_FLW; GEOLOGY). These REC variables were selected as three variables likely to drive supply and retention of both deposited and suspended sediment in New Zealand streams (Table AX.1). Combined, these Climate-Topography-Geology (CTG) classes form our 'least aggregated' classification—our starting point—grouping streams that should experience contrasting sediment supply and retention processes.

Table AX.1. Explanation of how REC Climate, Topography and Geology classes were aggregated prior to running the SSC algorithm, which further aggregates the resulting CTG classes based on the similarity of their average sediment states.

REC variable	Values	Aggregation to form new CTG classes, prior to running SSC algorithm
Climate	1. Warm-Wet 2. Warm-Extremely Wet 3. Warm-Dry 4. Cold-Wet 5. Cold-Extremely Wet 6. Cold-Dry	Wet and Extremely Wet were combined given these two climatic classes are both characterised by generally high runoff. Hence six values were aggregated to four: 1. Warm-Wet 2. Warm-Dry 3. Cold-Wet 4. Cold-Dry
Topography (SRC_OF_FLW)	1. Lowland 2. Lakefed 3. Hill 4. Mountain 5. Glacial Mountain	Mountain and Glacial Mountain classes combined on the basis of them both being associated with rivers of high gradient, hence low sediment retention. Yielding four topography classes: 1. Lowland 2. Lakefed 3. Hill 4. Mountain
Geology	1. Soft Sedimentary 2. Hard Sedimentary 3. Alluvium 4. Plutonic Volcanic 5. Miscellaneous 6. Volcanic Basic 7. Volcanic Acidic	Plutonic Volcanic and Miscellaneous were aggregated with Soft Sedimentary, based on exploration of the frequency histograms of sediment values within CTG classes, and consultation with expert geologists. Volcanic Basic and Volcanic Acidic combined to form Volcanic – geology resistant to erosion. This aggregation yielded four geological classes: 1. Soft Sedimentary 2. Hard Sedimentary 3. Alluvium 4. Volcanic

Using the CTG classification as a basis, we then implemented the following steps towards meeting Objectives 1 and 2:

Step 1. Characterise each CTG class as a vector defining the frequency distribution of ESV values within that CTG class; if there are insufficient data in initial CTG classes, aggregate in a logical fashion until the frequency distribution of the ESV can be defined in each CTG class;

Step 2. Based on the frequency distribution of ESV values within each CTG class, use multivariate analysis to determine the dissimilarity of ESV frequency distributions among classes, such that we may aggregate CTG classes into sediment classes based on the similarity of their ESV distributions.

Step 3. Within each sediment class, estimate the ESV reference states, determine which level of aggregation provides the most parsimonious description of reference states for each ESV, and map the spatial distribution of sediment classes, hence reference states, to all river reaches of the New Zealand river network.

Step 1: Characterising CTG classes by their ESV characteristics

In accordance with Principle 3 we wished to develop a SSC that was based on observed sediment data, so we needed to define what a 'sample' was. Herein, an individual sample was the median of all ESV values recorded at a monitoring site. What constituted a monitoring site varied between ESVs. For deposited fine sediment a 'site' was an individual reach (NZReach) within the New Zealand Freshwater Fish Database (NZFFD). For the two suspended sediment ESVs, a monitoring site was an individual monitoring station within the regional council monitoring network, which comprises the State of the Environment (SOE) monitoring. Specifically, the suspended sediment data used herein was the data collated by the 2018 MfE State and Trends Projects.

If we were to do no aggregation of the CTG classes in Table AX.1 we have up to 6 (climate classes) x 5 (topography classes) x 7 (geological classes) = 210 possible CTG classes. If we undertook no further aggregation then our sediment classes would be our 210 CTG classes (assuming all CTG classes are represented within New Zealand).

We defined the ESV characteristics of each CTG class as the frequency distribution of ESV values within that CTG class. To estimate a frequency distribution we obviously require a 'reasonable' number of samples. We selected N = 20 samples as the minimum sample size for histogram estimation. This value is somewhat arbitrary, but its selection was based on exploration of the data and seeking a balance between a minimum N that was too stringent (too high, resulting in too many CTGs being excluded the SSC) and too lenient (too low, resulting in an imprecise characterisation of CTGs).

The routine for defining the histogram bins was common to each ESV: 11 bins were established for each ESV, with 10 breakpoints defined as a sequence from the minimum value, to the maximum value, with a step size of range/10. The minimum, maximum and range were estimated using the global dataset for each ESV.

Prior to moving onto Step 2, we aggregated certain CTG classes if (a) one of a pair of CTG classes had N < 20; (b) the two CTG classes were likely to experience similar sediment supply and retention characteristics. The CTG classes resulting from this first step of aggregation are presented in Table AX.1.

Within the REC there are a total of 52 CTG classes represented (the classes in column 3 of Table AX.1). Of these, we had sufficient data to include 34 CTG classes for deposited sediment, and 18 CTG classes for turbidity. As we will see below (Step 3), although we had insufficient data to include a large proportion of the total CTG classes, the CTG classes included comprise a majority of the New

Zealand stream network. The CTG classes for which we had sufficient deposited and suspended sediment data are presented in Fig. AX.2 and Fig. AX.3 respectively. Although variation in ESV composition among CTG classes is evident for both deposited and suspended sediment in these figures, many CTG classes exhibit similar ESV composition, justifying further aggregation of the SSC (Fig. AX.2 and AX.3).

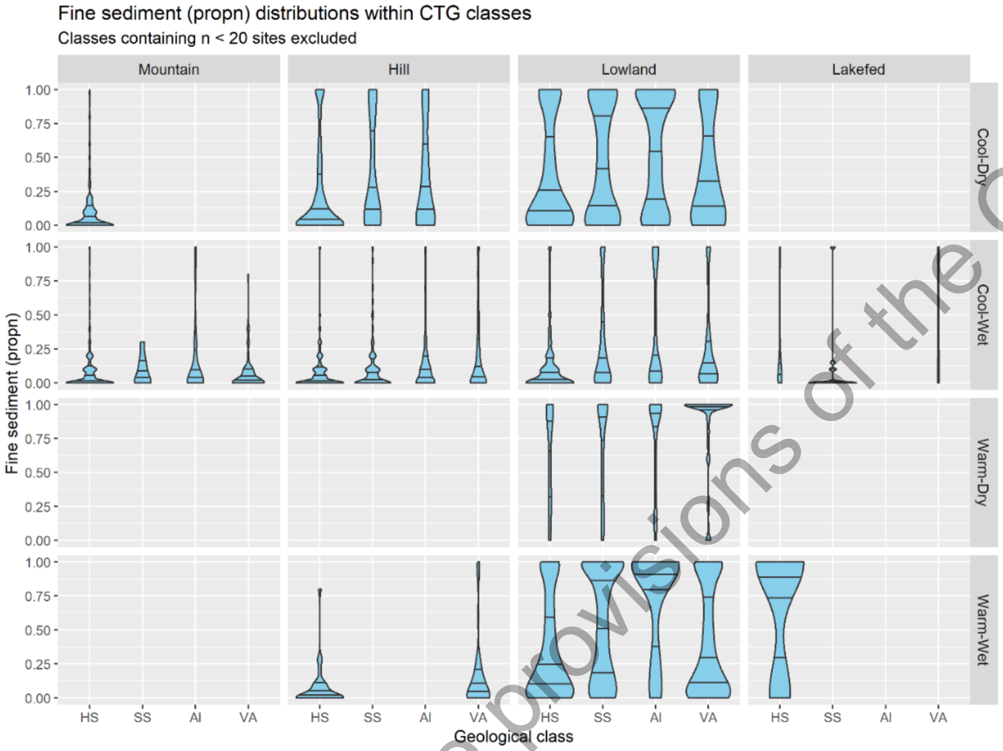

Figure AX.2. Violin plots describing the frequency distributions of deposited fine sediment within all REC CTG (Climate-Topography-Geology) classes considered.

Figure AX.3. Violin plots describing the frequency distributions of suspended fine sediment (turbidity) within all REC CTG (Climate-Topography-Geology) classes considered.

Within Table AX.2 the CTG classes for which we had sufficient data (n ≥ 20) and insufficient data are listed, for both the deposited and suspended SSC.

Table AX.2. Climate-Topography-Geology (CTG) classes that were either mapped or unmapped to the SSC for both deposited and suspended sediment. CTG classes unmapped contained less than 20 sites/samples to use for defining the ESV histogram.

	Deposited SSC		**Suspended SSC**	
	mapped	**unmapped**	**mapped**	**unmapped**
1	CD_Hill_Al	CD_Hill_VA	CD_Hill_HS	CD_Hill_Al
2	CD_Hill_HS	CD_Lake_Al	CD_Low_Al	CD_Hill_SS
3	CD_Hill_SS	CD_Lake_HS	CD_Low_HS	CD_Hill_VA
4	CD_Low_Al	CD_Lake_SS	CD_Low_SS	CD_Lake_Al
5	CD_Low_HS	CD_Lake_VA	CW_Hill_HS	CD_Lake_HS
6	CD_Low_SS	CD_Mount_Al	CW_Hill_SS	CD_Lake_SS
7	CD_Low_VA	CD_Mount_SS	CW_Hill_VA	CD_Lake_VA
8	CD_Mount_HS	CD_Mount_VA	CW_Lake_VA	CD_Low_VA
9	CW_Hill_Al	CW_Lake_Al	CW_Low_Al	CD_Mount_Al
10	CW_Hill_HS	WD_Hill_VA	CW_Low_HS	CD_Mount_HS
11	CW_Hill_SS	WD_Lake_Al	CW_Low_SS	CD_Mount_SS
12	CW_Hill_VA	WD_Lake_HS	CW_Low_VA	CD_Mount_VA
13	CW_Lake_HS	WD_Lake_SS	CW_Mount_HS	CW_Hill_Al
14	CW_Lake_SS	WD_Lake_VA	WD_Low_Al	CW_Lake_Al
15	CW_Lake_VA	WW_Hill_SS	WD_Low_SS	CW_Lake_HS
16	CW_Low_Al	WW_Lake_Al	WW_Low_HS	CW_Lake_SS
17	CW_Low_HS	WW_Lake_SS	WW_Low_SS	CW_Mount_Al
18	CW_Low_SS	WW_Lake_VA	WW_Low_VA	CW_Mount_SS
19	CW_Low_VA			CW_Mount_VA
20	CW_Mount_Al			WD_Hill_VA
21	CW_Mount_HS			WD_Lake_Al
22	CW_Mount_SS			WD_Lake_HS
23	CW_Mount_VA			WD_Lake_SS
24	WD_Low_Al			WD_Lake_VA
25	WD_Low_HS			WD_Low_HS
26	WD_Low_SS			WD_Low_VA
27	WD_Low_VA			WW_Hill_HS
28	WW_Hill_HS			WW_Hill_SS
29	WW_Hill_VA			WW_Hill_VA
30	WW_Lake_HS			WW_Lake_Al
31	WW_Low_Al			WW_Lake_HS
32	WW_Low_HS			WW_Lake_SS
33	WW_Low_SS			WW_Lake_VA
34	WW_Low_VA			WW_Low_Al

Step 2: Aggregation of CTG classes using cluster analysis

Following Step 1 the set of CTG classes for both deposited and fine sediment was characterised as a 'CTG class' x 'histogram-bin' matrix, thus permitting the estimation of multivariate similarity in the frequency distribution of sediment values among CTG classes. Bray-Curtis similarity between CTG classes was estimated prior to classification analysis using hierarchic clustering. Clustering was performed using average linkage, which tends to preserve the structure of dissimilarity among samples better than complete and single linkage algorithms (Oksanen 2015). R package vegan was used for all multivariate analysis (Oksanen et al. 2018).

In hierarchical clustering there are fewer clusters at a higher level of dissimilarity, while at a lower level of dissimilarity more clusters are produced. Thus the classification method used herein yields SSCs containing different numbers of sediment classes, depending on the level of dissimilarity selected to aggregate CTG classes into sediment classes. We generated four SSCs for both the deposited and suspended ESVs; one each for sediment classes grouped at (1) 50%; (2) 30%; (3) 20%; and (4) 15% dissimilarity. For both deposited and suspended sediment, these dissimilarities yielded 2, 4, 8 and 12 sediment classes. For ease of communication we hereafter refer to these different critical dissimilarities as 'levels of aggregation', with Aggregation Levels 1, 2, 3 and 4 corresponding to sediment classes aggregated at 50%, 30%, 20% and 15% dissimilarity respectively. Individual sediment classes within each level are referred to in a manner such that the level of aggregation is explicit; for example sediment classes L1.1 and L4.3 are, respectively, sediment classes 1 at Aggregation Level 1, and 3 at Aggregation Level 4.

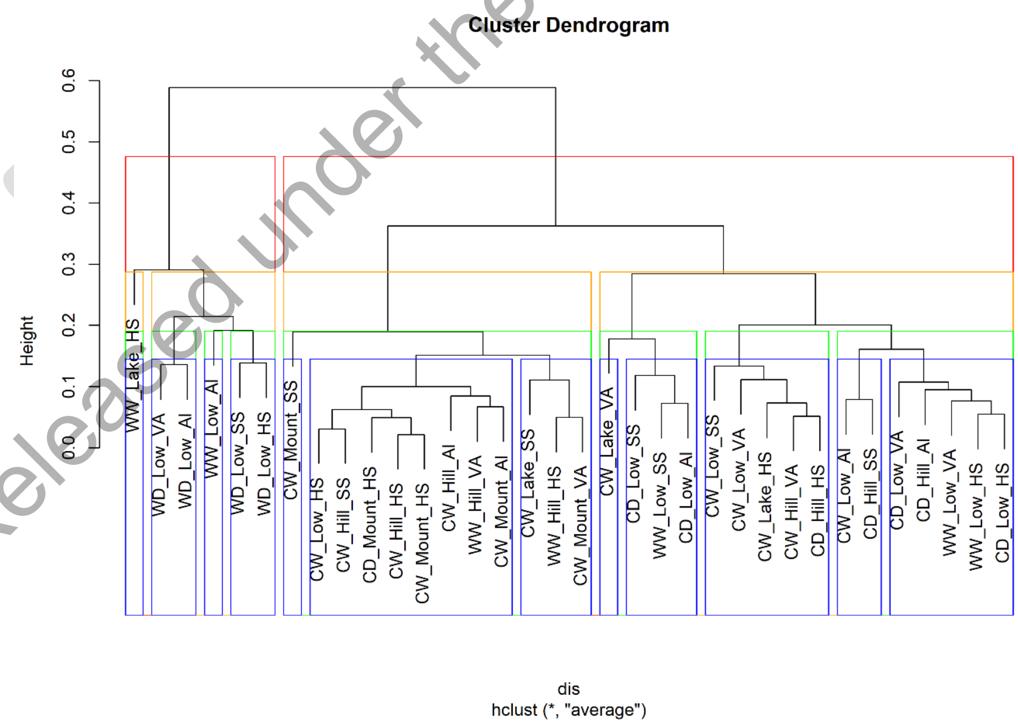

Figure AX.4. Dendrogram showing four levels of aggregation of CTG classes based on the (Bray-Curtis) similarity of their frequency distributions of **deposited fine sediment**. Boxes outline sediment classes at Aggregation Levels 1 (red; two groups); 2 (orange; four groups); 3 (green; eight groups); and 4 (blue; 12 groups).

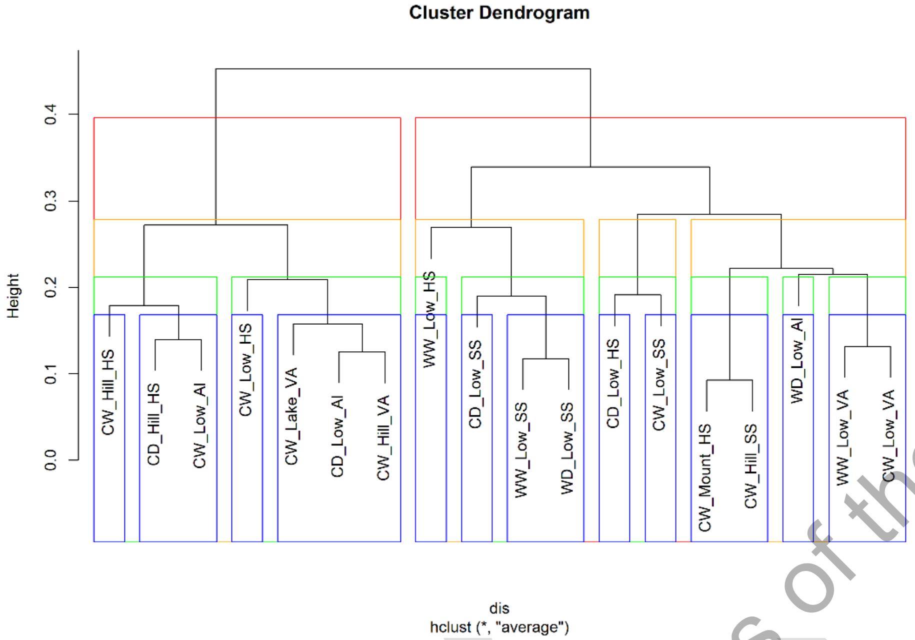

Figure AX.5. Dendrogram showing four levels of aggregation of CTG classes based on the (Bray-Curtis) similarity of their frequency distributions of **turbidity** values (NTUs). Boxes outline sediment classes at Aggregation Levels 1 (red; two groups); 2 (orange; four groups); 3 (green; eight groups); and 4 (blue; 12 groups).

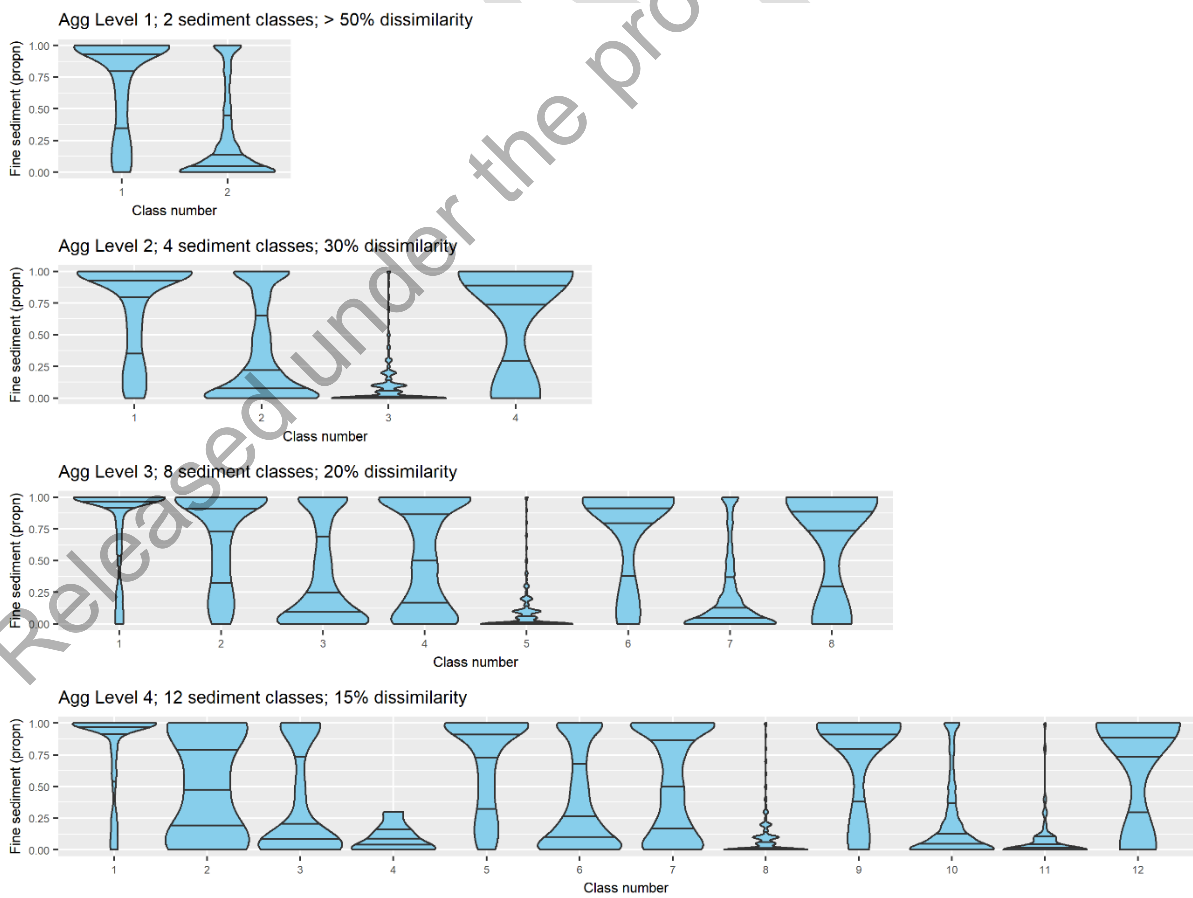

Figure AX.6. Violin plots describing the frequency distributions of **deposited fine sediment** within sediment classes at different levels of aggregation.

For both deposited and suspended fine sediment the cluster analysis yielded sediment classes that clearly had different climatic, topographical and geological characteristics (Fig. AX.4; Fig. AX.5; Table AX.3). Examination of the frequency distributions of values within sediment classes showed very strong differences in distributions at Aggregation Level 1 for both deposited (Fig. AX.6) and suspended sediment (Fig. AX.7). Differences in the distributions of sediment values among classes became more nuanced through Levels 2 – 4 for both deposited and suspended sediment (Fig. AX.6 and Fig. AX.7). Results of the cluster analysis are summarised in Table AX.3.

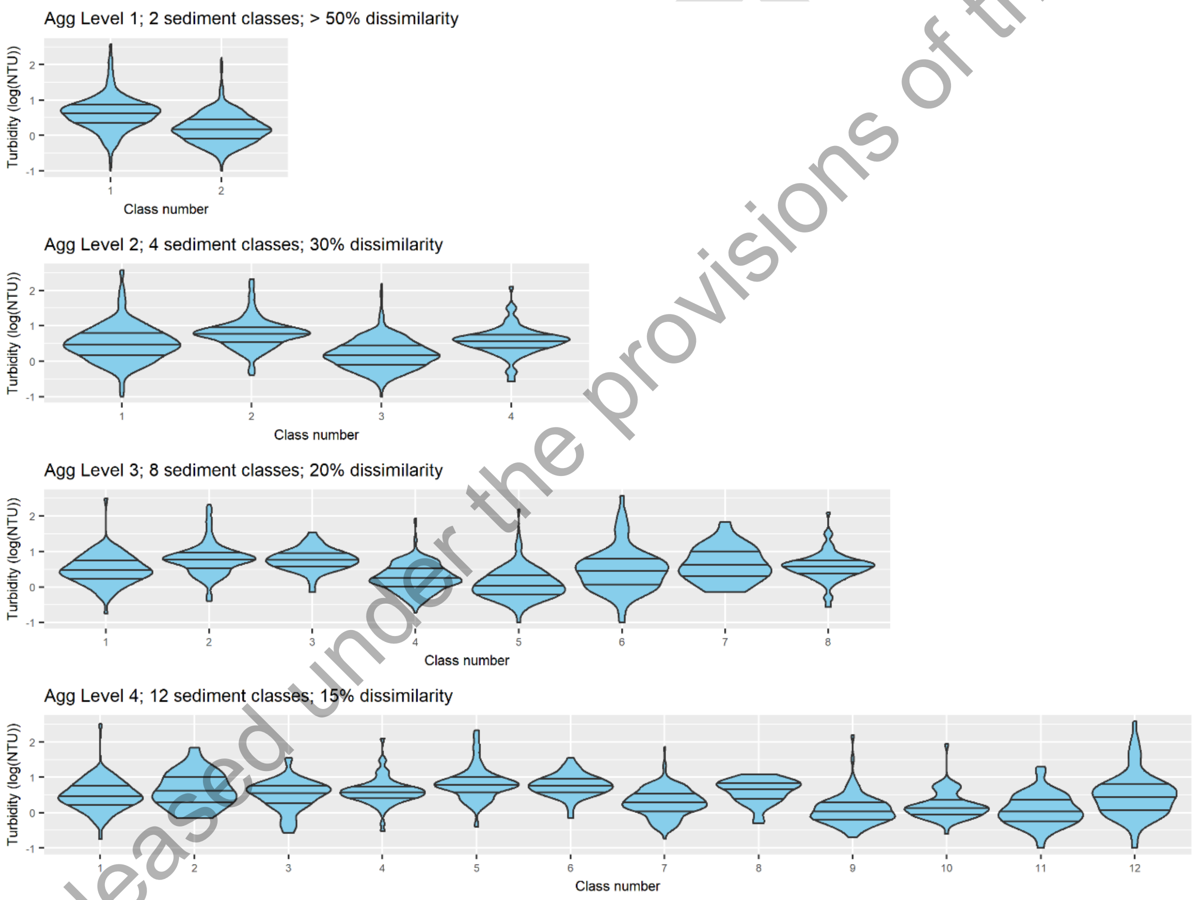

Figure AX.7. Violin plots describing the frequency distributions of **deposited fine sediment** within sediment classes at different levels of aggregation.

Table AX.3. Class membership hierarchy for both the deposited and suspended sediment classes at different levels of aggregation (Aggregation Levels 1-4). CTG = Climate-Topography-Geology classes. Rates of sediment supply and retention have been included: very high; high; medium; low; very low.

Deposited fine sediment class hierarchy						Suspended sediment class hierarchy					
Agg L1	**Agg L2**	**Agg L3**	**Agg L4**	**CTG Classes**	**Supply, retention**	**Agg L1**	**Agg L2**	**Agg L3**	**Agg L4**	**CTG Classes**	**Supply**
1	1	1	1	WD_Low_VA; WD_Low_Al	Very high	1	1	1	1	WW_Low_VA; CW_Low_VA	Med
		2	5	WD_Low_SS; WD_Low_HS	High			6	12	CW_Mount_HS; CW_Hill_SS	High
		6	9	WW_Low_Al	High			7	2	WD_Low_Al	High
	4	8	12	WW_Lake_HS	High		2	2	5	WW_Low_SS; WD_Low_SS	Very high
2	2	3	6	WW_Low_VA; WW_Low_HS; CD_Low_VA; CD_Hill_Al; CD_Low_HS	Low				8	CD_Low_SS	Very high
			3	CW_Low_Al; CD_Hill_SS	Low			3	6	WW_Low_HS	Very high
		4	7	WW_Low_SS; CD_Low_SS; CD_Low_Al	Med		4	8	3	CD_Low_HS	Med
			2	CW_Lake_VA	Med				4	CW_Low_SS	High
		7	10	CW_Hill_VA; CW_Low_VA; CW_Low_SS; CD_Hill_HS; CW_Lake_HS	Low	2	3	4	7	CD_Low_Al; CW_Lake_VA; CW_Hill_VA	Med
	3	5	8	WW_Hill_VA; CW_Hill_HS; CW_Low_HS; CW_Mount_HS; CW_Hill_SS; CW_Hill_Al; CD_Mount_HS; CW_Mount_Al	Very low				10	CW_Low_HS	Low
			11	WW_Hill_HS; CW_Mount_VA; CW_Lake_SS	Very low			5	9	CW_Hill_HS	Very low
			4	CW_Mount_SS	Low				11	CD_Hill_HS; CW_Low_Al	Very low

Step 3: Estimating ESV reference states

Two broad approaches to estimating reference state were considered: The first approach involves estimating the state of an ESV within river segments that have no history of significant anthropogenic disturbance upstream—the 'reference site' approach. Under this approach, the reference state is often referred to as the 'minimally disturbed condition' (Lewis et al. 1999, Stoddard et al. 2006). An advantage of the reference site approach is its simplicity; the definition of reference state is intuitive and its calculation requires little to no statistical sophistication and so is easy to explain.

However, minimally-disturbed river segments are usually rare, resulting in very few replicate reference sites per sediment class, which may in turn lead to biased estimates of reference state (McDowell et al. 2013). That is, the lower the number of replicate reference sites the greater the risk of having reference states that are not representative of the broader region we wish to manage.

The second approach we considered for estimating reference states of ESVs was that of Dodds and Oakes (2004). This approach involves (a) selecting a model that describes ESV state as a function of covariates that describe the magnitude of anthropogenic disturbance within a region; and (b) using that model to estimate predicted ESV state at zero anthropogenic disturbance. We refer to this approach as the 'model-based' approach. The model-based approach involves using all the data available within a region, and so it follows that (a) if the data from which sites are obtained are randomly distributed throughout the region we wish to manage; and (b) if our model is a good fit to the data, then we obtain a least biased estimate of reference state.

A disadvantage of the model-based approach is that it is more complex than the reference site approach, and so may be more difficult for various stakeholders to understand.

In the present study we used the model-based approach, due to the small number and restricted distribution of reference sites for deposited and suspended sediment. We sought parsimonious models of reference states within sediment classes. Towards that end the following set of candidate models was fitted to each ESV, at each aggregation level:

$$ESV = \beta_0 + \beta_1 P + \beta_2 C + \beta_3 PC + \varepsilon \qquad \text{Model 1}$$

$$ESV = \beta_0 + \beta_1 P + \beta_2 C + \beta_3 PC + \beta_4 E + \beta_5 EC + \varepsilon \qquad \text{Model 2}$$

$$ESV = \beta_0 + \beta_1 P + \beta_2 C + \beta_3 PC + \beta_4 U + \beta_5 UC + \varepsilon \qquad \text{Model 3}$$

$$ESV = \beta_0 + \beta_1 P + \beta_2 C + \beta_3 PC + \beta_4 E + \beta_5 EC + \beta_6 U + \beta_7 UC + \varepsilon \qquad \text{Model 4}$$

In the above equations the β values are parameters and ε is error. The covariates *P*, *E* and *U* are continuous covariates with domain [0,1] describing the proportions of the catchment upstream comprised of heavy pasture, exotic vegetation (mostly pine forests) and urban development, respectively. *C* is a categorical, fixed covariate referring to the sediment class. The number of values of *C* is dependent on the aggregation level: at Level 1, *C* has two values (one for each of two sediment classes); at Level 2, *C* has 4 values; at Level 3, *C* has 8 values; at Level 4, *C* has 12 values. When the ESV was deposited fine sediment (proportion) we used binomial linear models, but when the ESV was either turbidity or clarity, Gaussian linear models were fitted.

For each ESV, the Akaike Information Criterion (AIC; Burnham and Anderson 2002) was used to select the most parsimonious candidate from Models 1-4, *within each aggregation level*. Consequently, for each ESV we generated four possible models of reference state; one at each level of aggregation. To obtain the reference state within each sediment class, within each aggregation level, we obtained the predicted value within each level of *C*, with other covariates set to zero. Occasionally the slope of the fitted model within a certain sediment class was approximately zero and in the direction opposite to that expected (e.g. turbidity actually decreasing as anthropogenic pressure increases). When this occurred the reference state for this class was estimated as the median ESV value.

It follows that the final step of selecting an appropriate model of reference state was to choose the level of aggregation of sediment classes.

Within this project, sediment reference states are passed to models of biological response to ESV, which in turn are used to estimate NOF management bands. Accordingly, the decisive factor determining which aggregation level to use may be the availability of either ESV or biological data in sediment classes. In any case, to assist decisions concerning the level of aggregation to use, we provided three further outputs:

First, for each ESV the optimal models (from Models 1-4) across each of the four levels of aggregation differed considerably in their complexity. Suppose, for example that Model 4 is the most likely model of reference state for Aggregation Levels 1 (average dissimilarity between classes = 50%) and 4 (average dissimilarity = 15%). Then at Level 1 the most likely model of reference states has 8 parameters while at Level 4 the most likely model has 48 parameters. The Level 4 model is likely to yield less biased estimates of reference states, because the reference state estimation is allowed to vary across a finer-grained decomposition of sediment classes throughout New Zealand. But any reduction in bias comes at the cost of many more parameters. Thus we have a standard model selection problem of the need to find an appropriate balance between model complexity and simplicity. We employed information-theoretic statistics to help find that balance. Specifically, for each ESV, we estimated the following statistics for the most likely models at each of the four levels of aggregation: (a) AIC; (b) the AIC model rank: $\Delta_i = AIC_i - \min(AIC)$; and (c) w_i, the Akaike weight of model *i*, interpreted as the approximate probability that Model *i* is the best model in the candidate set, given the data (Burnham and Anderson 2002).

Second, for each ESV we generated plots to compare and contrast estimates of reference state with:

a) The median and interquartile range of the ESV, within the subset of the data where heavy pasture values were less than the lowest decile of all heavy pasture values (ESV_HPd1). This statistic provides a 'check' on the alignment between our modelled reference estimate and the distribution of observed ESV values under minimal anthropogenic disturbance, within each sediment class.
b) The median and interquartile range of the ESV as measured at reference sites within each sediment class (Reference). In this study a reference site was a site with the following catchment characteristics upstream, as estimated within the NZ REC (Snelder and Biggs 2002): > 90% cover of native vegetation; 0% coverage of urban development; < 5% exotic vegetation (hence < 5% commercial forestry).
c) The median and interquartile range of all ESV data within each sediment class (ESV_allData), such that we may see how our modelled reference states contrast with the contemporary, observed state of that ESV throughout regions defined by each sediment class. One could suggest that, if our modelled reference states are useful, then within sediment classes associated with agricultural development we would ideally see (i) reference states below the median of ESV_allData; but (ii) refrence estimates that are not so far below the IQR of ESV_allData that entire regions of NZ are set unachievable management objectives.

Third, for each ESV we generated plots showing how biased our estimates of reference state might be when we use a higher level of aggregation, when we group together more streams that may have different sediment states. These plots were designed to demonstrate the direction and magnitude of change in estimated reference state—hence the magnitude and direction of bias—as we move from a lower level of aggregation (e.g. Level 2; 30% dissimilarity between sediment classes) to a higher level of aggregation (e.g. Level 1; 50% average dissimilarity between sediment classes). Our SSCs are hierarchical, so multiple reference states within a lower level of aggregation may correspond to a single reference state at the next highest level of aggregation. In these plots we will see just how much several estimates of reference states at lower levels of aggregation are pulled towards the 'average' reference states at the higher levels of aggregation.

Deposited fine sediment

For deposited fine sediment, Model 4 provided the most parsimonious description of the data at all levels of aggregation. Hence, given the data and candidate models, we found variation in deposited fine sediment throughout New Zealand is best explained by the additive effects of heavy pasture, urbanisation and forestry, and how those three drivers interact with sediment classes of the New Zealand landscape. Using Nagelkerke's R^2 for generalised linear models, the R^2 values for the fit of Model 4 to the deposited sediment data were: 0.25 (Level 1); 0.30 (Level 2); 0.33 (Level 3); 0.34 (Level 4).

The fitted optimal models for deposited fine sediment are presented in Figure AX.8. Table AX.4 presents the reference states (intercepts) for, and the proportion of the NZ REC covered by, each sediment class, at each level of aggregation. In one instance (Class L3.6, which is also Class L4.9; Table AX.4) a counterintuitive slope was returned (Fig. AX.8, Agg. Level 3 and 4), resulting in the reference state for that class being estimated as the median deposited sediment value in that class. For most classes we had a good range of heavy pasture values for regression, irrespective of level of aggregation (Fig. AX.8).

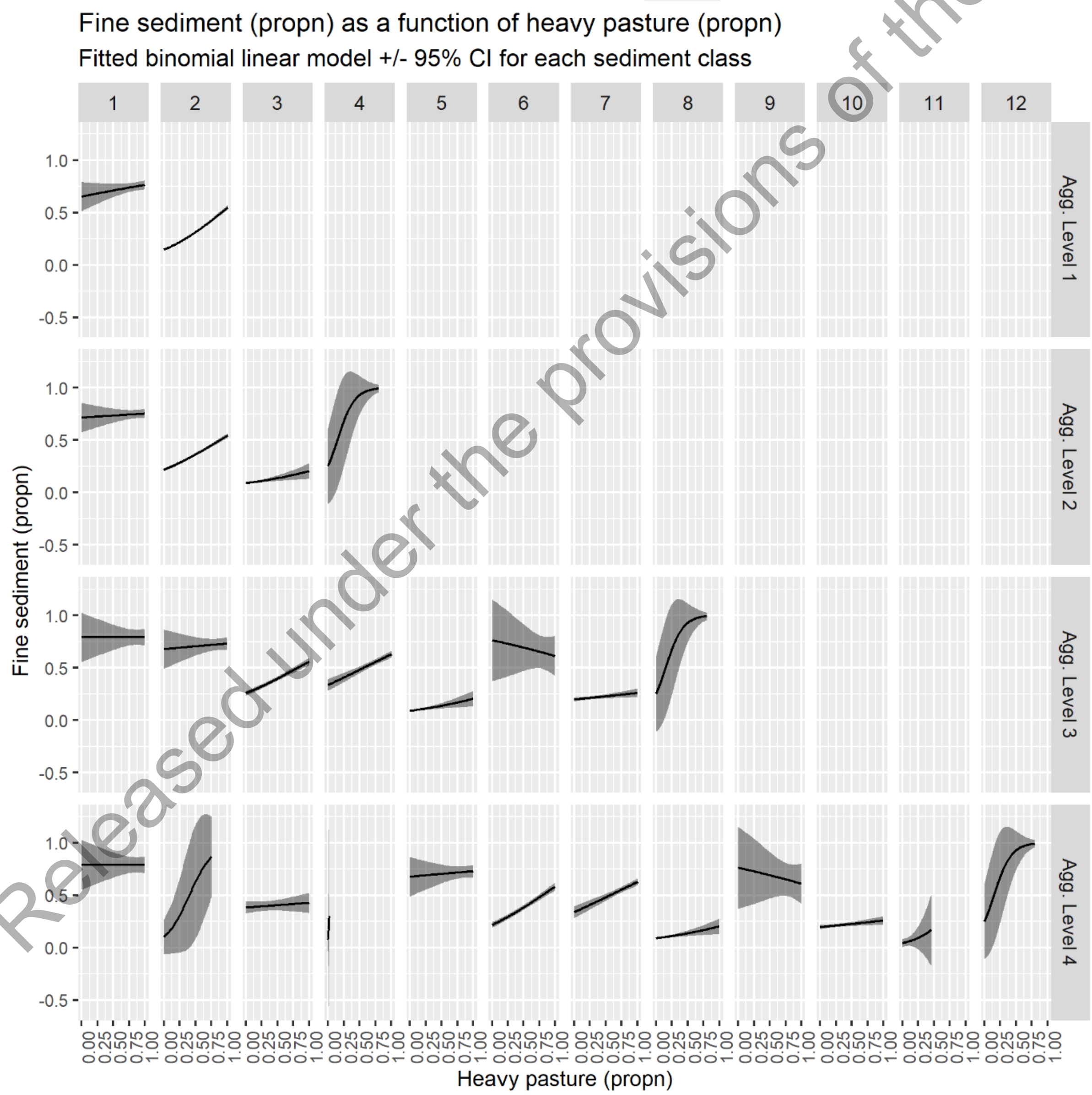

Figure AX.8. Binomial linear regression lines of Model 4, describing proportion of fine sediment as a function of proportion of heavy pasture, within each sediment class, at four different levels of aggregation (dissimilarity between) the REC Climate-Topography-Geology (CTG) classes. These fitted model traces were obtained by setting covariates *U* and *E* to zero, thus focusing on the impact of heavy pasture in a hypothetical catchment with no forestry and urban development.

Table AX.4. Reference states and percentage coverage of the New Zealand REC of deposited fine sediment classes at four levels of aggregation.

Agg. L1	Ref	% cover	Agg. L2	Ref	% cover	Agg. L3	Ref	% cover	Agg. L4	Ref	% cover	SedClasses
1	0.65	5.77	1	0.71	5.74	1	0.79	1.88	1	0.79	1.88	WD_Low_VA; WD_Low_Al
						2	0.36	3.42	5	0.36	3.42	WD_Low_SS; WD_Low_HS
						6	0.46	0.45	9	0.46	0.45	WW_Low_Al
			4	0.12	0.03	8	0.08	0.03	12	0.08	0.03	WW_Lake_HS
2	0.08	92.94	2	0.10	52.54	3	0.08	16.37	6	0.07	13.32	WW_Low_VA; WW_Low_HS; CD_Low_VA; CD_Hill_Al; CD_Low_HS
									3	0.14	3.05	CW_Low_Al; CD_Hill_SS
						4	0.12	15.94	7	0.12	15.51	WW_Low_SS; CD_Low_SS; CD_Low_Al
									2	0.03	0.43	CW_Lake_VA
						7	0.06	20.23	10	0.06	20.23	CW_Hill_VA; CW_Low_VA; CW_Low_SS; CD_Hill_HS; CW_Lake_HS
			3	0.04	40.39	5	0.03	40.39	8	0.03	36.41	WW_Hill_VA; CW_Hill_HS; CW_Low_HS; CW_Mount_HS; CW_Hill_SS; CW_Hill_Al; CD_Mount_HS; CW_Mount_Al
									11	0.01	2.04	WW_Hill_HS; CW_Mount_VA; CW_Lake_SS
									4	0.02	1.95	CW_Mount_SS

It is clear from Fig. AX.8 and Table AX.4 that the variation in reference state across sediment classes increases as we move from Aggregation Level 1 through to Level 4. This can also be seen in Figure AX.9, which presents the comparisons of our reference state estimates for deposited fine sediment with ESV_HPd1, the value of the ESV at reference sites and ESV_allData. The following inferences may be gleaned from Fig. AX.9:

1. Using the method of classification derived here, less than 2% of the New Zealand river network was unclassified.
2. Irrespective of the level of aggregation there was generally good agreement between model-based reference estimates and those based on reference sites alone. When there was discordance between the model-based and reference site estimates, model-based estimates were not necessarily always higher than those based on reference sites. For example, at Aggregation Level 4, the model-based estimate was lower than that based on reference sites for Class L4.9, while the reverse was true for Class L4.1 (Fig. AX.9). Notably, the number of reference sites in each of these classes was very low; 1 and 3, respectively.
3. At lower levels of aggregation, the number of reference sites is often very low (<10; Fig. AX.9 Level 3) and occasionally certain classes are without reference sites (Class L4.2; Fig. AX.9).

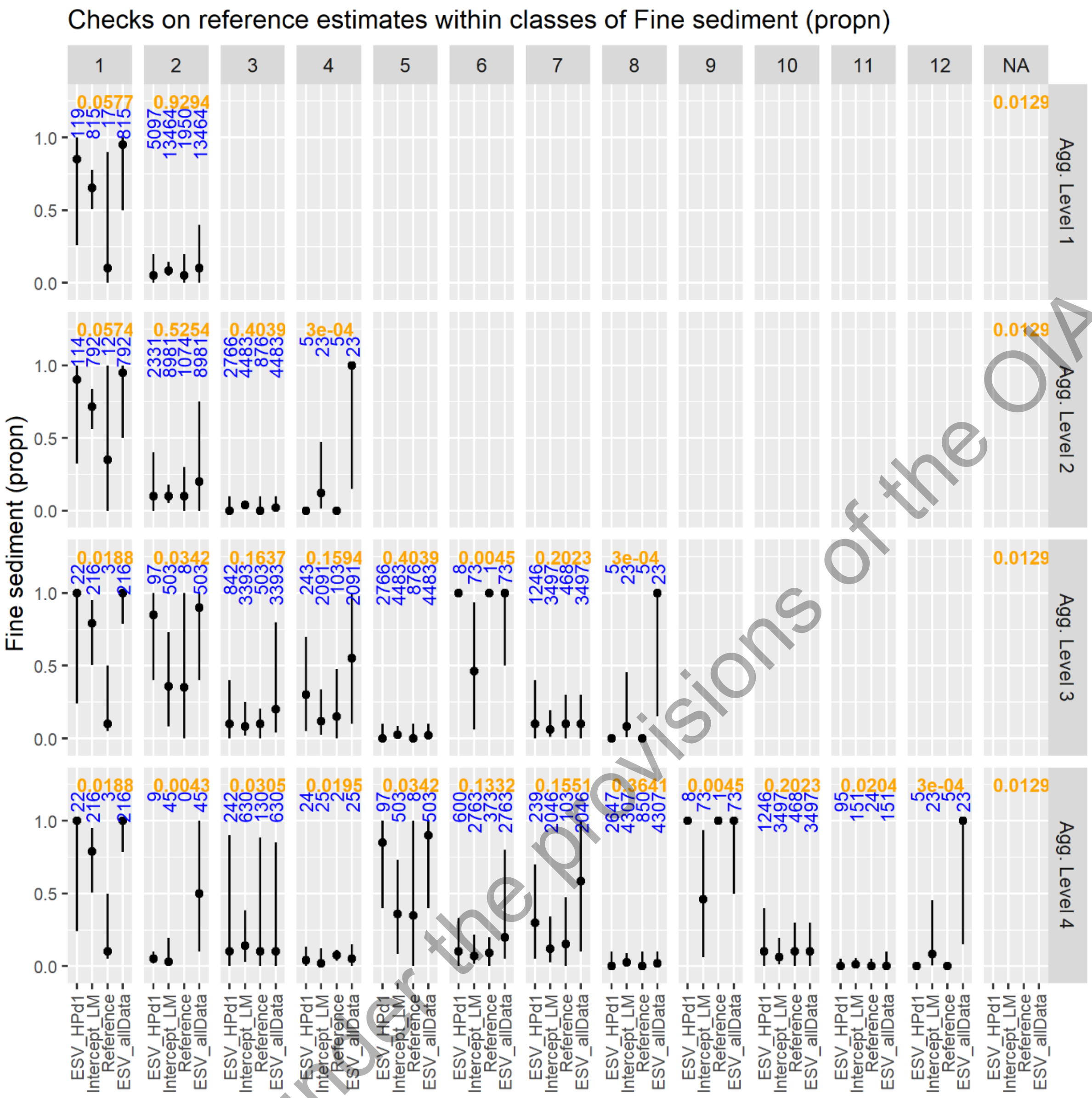

Figure AX.9. Comparison of reference state estimates for deposited fine sediment within each class at four levels of aggregation. ESV_HPd1: The median sediment value within the lowest decile of heavy pasture (error = IQR); Intercept_LM: the estimated y-intercept of Model 4 (error = 95% CI); Reference: The median fine sediment value obtained only from reference sites, within the sediment class (error = IQR). ESV_allData: The median deposited fine sediment value for all data within that class (error = IQR). Blue numbers above each point indicate the number of sites contributing data to each statistic. Orange numbers indicate the proportion of the entire New Zealand REC comprised of each sediment class. NA indicates the 'undefined' class; CTG classes containing insufficient ESV data to enter the classification algorithm.

Fig. AX.10 presents the direction and magnitude of change in reference state estimates as we further aggregate sediment classes from one level in our classification hierarchy to the next highest level. It is clear that the higher the level of aggregation we use the more biased our reference estimates. For example, at Aggregation Level 4, Classes L4.6 and L4.3 have, respectively, reference states of 0.07 and 0.14 (proportionate coverage). At Level 3 these two classes are aggregated yielding a reference state of 0.08 (Class L3.3). Thus aggregating from the lowest level to the the next highest level in the classification hierarchy can result in bias of 8% of total ESV range.

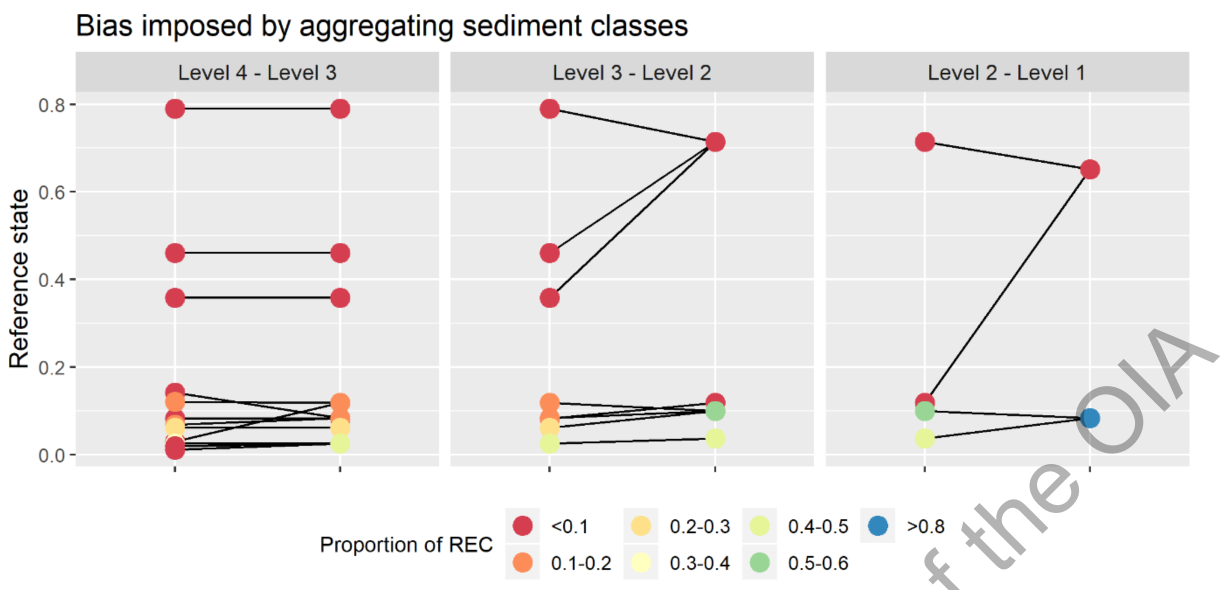

Figure AX.10. Change in reference state of deposited fine sediment as classes at one aggregation level are further aggregated into classes at the next highest level in our classification hierarchy.

The AIC statistics for Model 4 fitted to the deposited fine sediment data at Levels 1 through to 4 are presented in Table AX.5. Despite the large number of parameters, the most parsimonious model of deposited fine sediment as a function of anthropogenic development is the one that includes interactions between covariates and sediment classes at the lowest level of aggregation (12 classes). Indeed, relative to the other three levels of aggregation in the hierarchy, there is a probability of 1 that Level 4 is the most likely model in the candidate set. Thus the data very strongly indicate that the lowest level of aggregation provides the most parsimonious description of deposited fine sediment reference states in New Zealand.

Table AX.5. AIC statistics for Model 4 fitted to deposited fine sediment data at all four levels of aggregation in the classification hierarchy. K is the number of parameters in the regression model; AICc is the corrected AIC statistic; Δ_i = AIC_i – min(AIC) is known as the model rank; w_i, is the Akaike weight of model *i*, interpreted as the approximate probability that Model *i* is the best model in the candidate set, given the data; LL is log-likelihood of each model; Cum.Wt is the cumulative model weight of the ranked models.

Agg. Level	K	AICc	Δ_i	w_i	LL	Cum.Wt
4	47	12742.08	0	1	-6323.88	1
3	31	12779.02	36.94	0	-6358.44	1
2	15	13149.96	407.88	0	-6559.96	1
1	8	13676.5	934.42	0	-6830.25	1

The spatial distribution of deposited fine sediment classes at each of the four levels of aggregation is presented in Fig. AX.11.

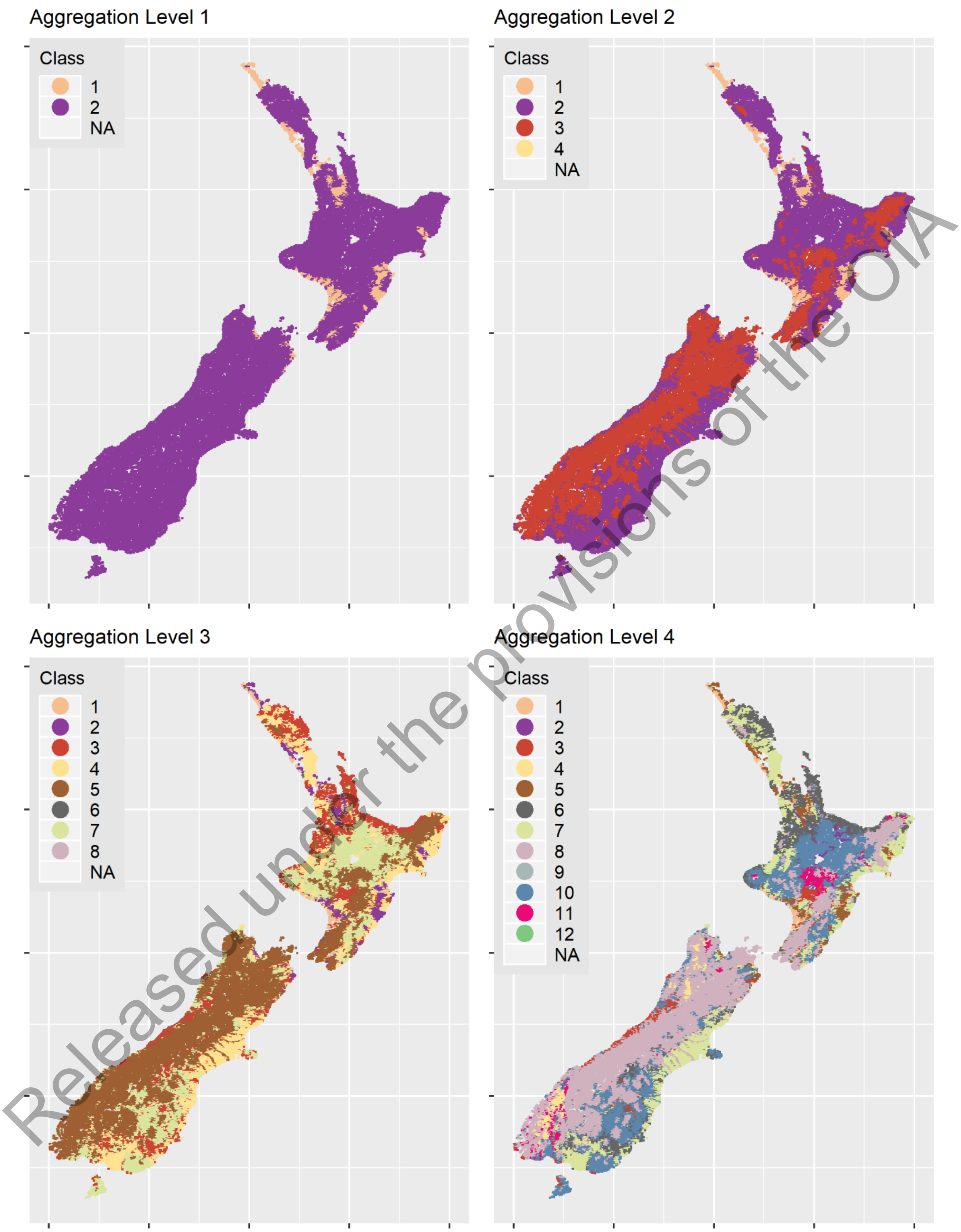

Figure AX.11. Spatial distribution of the deposited fine sediment classes under four different levels of aggregation of the REC Climate-Topography-Geology (CTG) classes. See Table AX.3 for description of sediment classes.

Turbidity

The model that best explained variation in turbidity as a function of our stressor covariates was dependent on the level of aggregation. At Aggregation Level 1 (50% dissimilarity between 2 sediment classes), Model 4 provided the most parsimonious description of the data. By contrast, at Aggregations Levels 3-1 the most parsimonious model of turbidity as a function of our stressor covariates was the most simple model in the set, Model 1. Hence, given the data and candidate models, at the highest level of aggregation we found variation in turbidity throughout New Zealand is best explained by the additive effects of heavy pasture, urbanisation and forestry, and how those three drivers interact with the two sediment classes of the New Zealand landscape. At finer levels of aggregation the model containing interactions between sediment classes and heavy pasture alone was optimal.

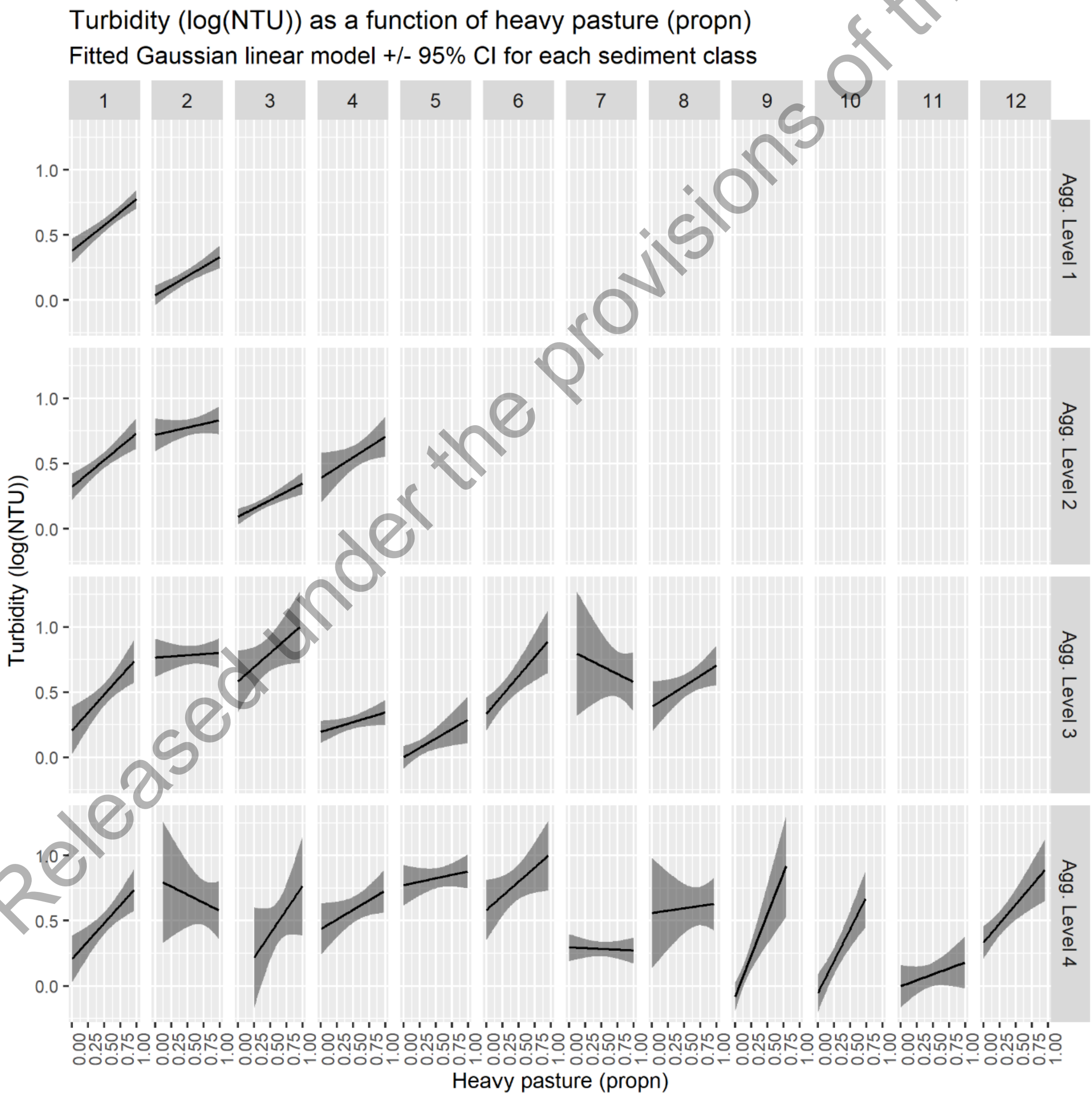

Figure AX.12. Gaussian linear regression lines of either Model 4 (Agg. Level 1) or Model 1 (Agg. Levels 2-4), describing turbidity as a function of proportion of heavy pasture, within each sediment class, at four different levels of aggregation (dissimilarity between) the REC Climate-Topography-Geology (CTG) classes. In the case of Agg. Level 1, these fitted model traces were obtained by setting covariates *U* and *E* to zero, thus focusing on the impact of heavy pasture in a hypothetical catchment with no forestry and urban development. NOTE change in y scale among plots.

Using Nagelkerke's R^2 for generalised linear models, the R^2 values for the fit of the most parsimonious models to the turbidity data were: 0.27 (Level 1; Model 4); 0.28 (Level 2; Model 1); 0.30 (Level 3; Model 1); 0.34 (Level 4; Model 1).

The fitted optimal models for turbidity are presented in Figure AX.12. Table AX.6 presents the reference states (intercepts) for, and the proportion of the NZ REC covered by, each sediment class, at each level of aggregation. At Aggregation Level 3, Class L3.7 (which was also Class L4.2) returned a counterintuitive slope, so the reference state for that sediment class was estimated as the median turbidity value for all data in that class (Table AX.6). At Aggregation Level 4, in addition to L4.2, the reference state of Class L4.4 was estimated as the median turbidity value within that class (Table AX.6). For most classes we had a good range of heavy pasture values for regression, irrespective of level of aggregation (Fig. AX.12).

Table AX.6. Reference states and percentage coverage of the New Zealand REC of turbidity sediment classes at four levels of aggregation.

<table>
<tr><th>Agg. L1</th><th>Ref</th><th>% cover</th><th>Agg. L2</th><th>Ref</th><th>% cover</th><th>Agg. L3</th><th>Ref</th><th>% cover</th><th>Agg. L4</th><th>Ref</th><th>% cover</th><th>SedClasses</th></tr>
<tr><td rowspan="8">1</td><td rowspan="8">0.38</td><td rowspan="8">56.82</td><td rowspan="3">1</td><td rowspan="3">0.32</td><td rowspan="3">30.83</td><td>1</td><td>0.21</td><td>7.05</td><td>1</td><td>0.21</td><td>7.05</td><td>WW_Low_VA; CW_Low_VA</td></tr>
<tr><td>6</td><td>0.13</td><td>22.37</td><td>12</td><td>0.13</td><td>22.37</td><td>CW_Mount_HS; CW_Hill_SS</td></tr>
<tr><td>7</td><td>0.62</td><td>1.42</td><td>2</td><td>0.62</td><td>1.42</td><td>WD_Low_Al</td></tr>
<tr><td rowspan="3">2</td><td rowspan="3">0.40</td><td rowspan="3">17.26</td><td rowspan="2">2</td><td rowspan="2">0.56</td><td rowspan="2">14.42</td><td>5</td><td>0.56</td><td>10.81</td><td>WW_Low_SS; WD_Low_SS</td></tr>
<tr><td>8</td><td>0.35</td><td>3.61</td><td>CD_Low_SS</td></tr>
<tr><td>3</td><td>0.38</td><td>2.84</td><td>6</td><td>0.38</td><td>2.84</td><td>WW_Low_HS</td></tr>
<tr><td rowspan="2">4</td><td rowspan="2">0.07</td><td rowspan="2">8.72</td><td rowspan="2">8</td><td rowspan="2">0.18</td><td rowspan="2">8.72</td><td>3</td><td>-0.17</td><td>2.72</td><td>CD_Low_HS</td></tr>
<tr><td>4</td><td>0.23</td><td>6.01</td><td>CW_Low_SS</td></tr>
<tr><td rowspan="4">2</td><td rowspan="4">-0.34</td><td rowspan="4">30.50</td><td rowspan="4">3</td><td rowspan="4">-0.23</td><td rowspan="4">30.50</td><td rowspan="2">4</td><td rowspan="2">-0.01</td><td rowspan="2">13.38</td><td>7</td><td>0.09</td><td>11.35</td><td>CD_Low_Al; CW_Lake_VA; CW_Hill_VA</td></tr>
<tr><td>10</td><td>-0.26</td><td>2.03</td><td>CW_Low_HS</td></tr>
<tr><td rowspan="2">5</td><td rowspan="2">-0.21</td><td rowspan="2">17.12</td><td>9</td><td>-0.29</td><td>9.25</td><td>CW_Hill_HS</td></tr>
<tr><td>11</td><td>-0.21</td><td>7.87</td><td>CD_Hill_HS; CW_Low_Al</td></tr>
</table>

Based on Fig. AX.12 and Table AX.6, the variation in reference state across sediment classes increases as we move from Aggregation Level 1 through to Level 4. This can also be seen in Figure AX.13, which presents the comparisons of our reference state estimates for turbidity with ESV_HPd1, the value of the ESV at reference sites and ESV_allData. The following inferences may be gleaned from Fig. AX.13:

1. Using the method of classification derived here, less than 13% of the New Zealand river network was unclassified for the turbidity ESV.
2. Irrespective of the level of aggregation there was generally good agreement between model-based reference estimates and those based on reference sites alone. When there was discordance between the model-based and reference site estimates, model-based estimates were not necessarily always higher than those based on reference sites. For example, at

Aggregation Level 4, the model-based estimate was lower than that based on reference sites for Classes L4.6 and L4.10, while the reverse may be true for Class L4.11 (Fig. AX.13).

3. At lower levels of aggregation, the number of reference sites is often very low (<5; Fig. AX.13 Level 4) and often classes are without reference sites at Aggregation Levels 3 and 4 (Fig. AX.13).

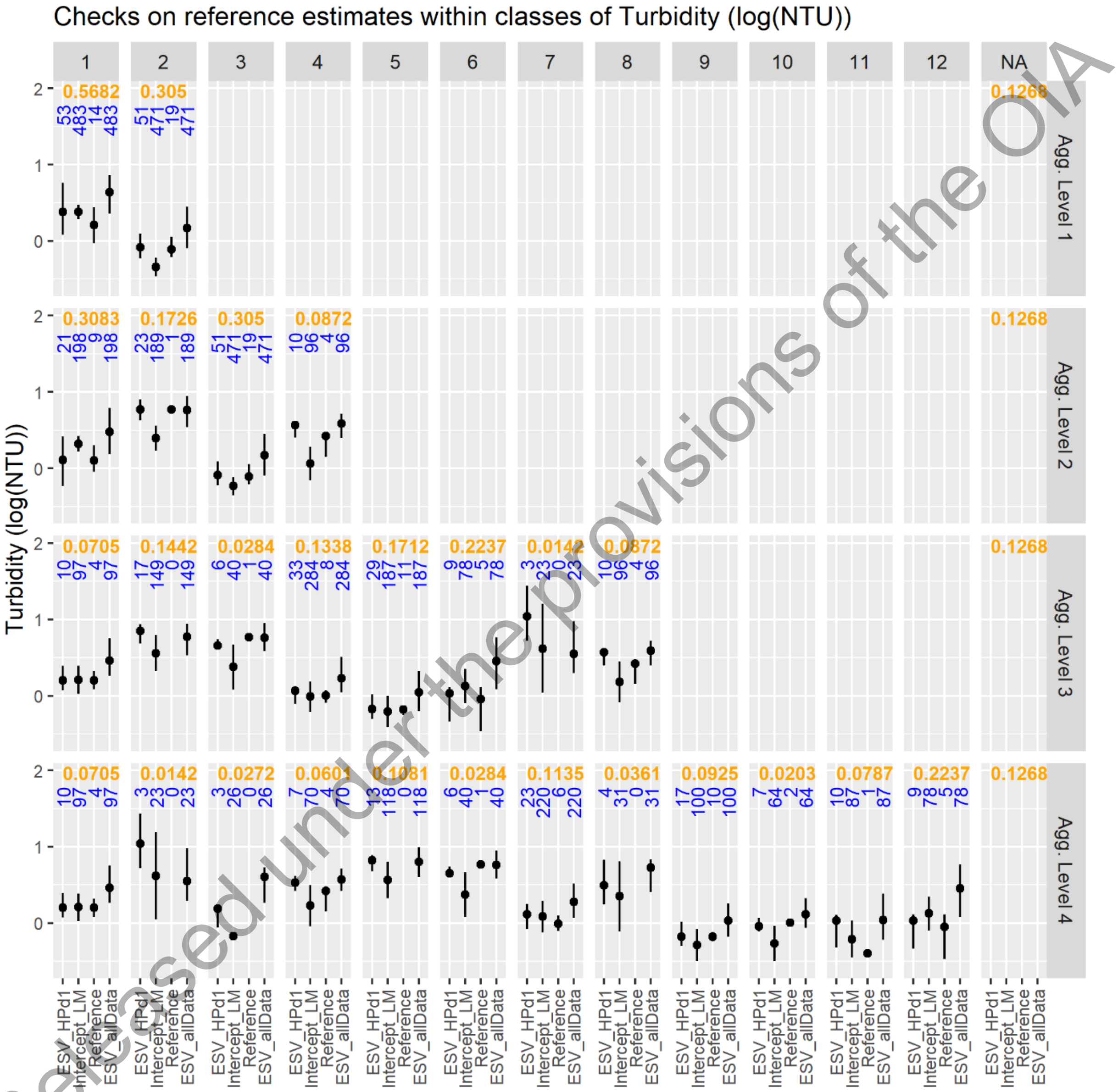

Figure AX.13. Comparison of reference state estimates for turbidity within each class at four levels of aggregation. ESV_HPd1: The median turbidity value within the lowest decile of heavy pasture (error = IQR); Intercept_LM: the estimated y-intercept of the most parsimonious models of turbidity as a function of stressor covariates (error = 95% CI); Reference: The median fine sediment value obtained only from reference sites, within the sediment class (error = IQR). ESV_allData: The median deposited fine sediment value for all data within that class (error = IQR). Blue numbers above each point indicate the number of sites contributing data to each statistic. Orange numbers indicate the proportion of the entire New Zealand REC comprised of each sediment class. NA indicates the 'undefined' class; CTG classes containing insufficient ESV data to enter the classification algorithm.

Fig. AX.14 presents the direction and magnitude of change in turbidity reference state estimates as we further aggregate sediment classes from one level in our classification hierarchy to the next highest level. As was the case for deposited fine sediment, the higher the level of aggregation we use the more biased our reference estimates. Consider Class L1.1: it can be decomposed into classes at Level 4 whose reference states span almost the entire range of reference states (-0.17 – 0.62; Table AX.6; Fig. AX.14). Indeed, even when aggregating from Level 2 to Level 1 we collapse a range of reference state estimates spanning ca. half the range of all reference estimates (0.07-0.40; L2.4-L2.2; Table AX.6; Fig. AX.14) to a single value (0.38; L1.1; Table AX.6; Fig. AX.14). Of note is the fact that the bias estimates presented here are in log-scale, so the real bias in units of NTUs would be greater.

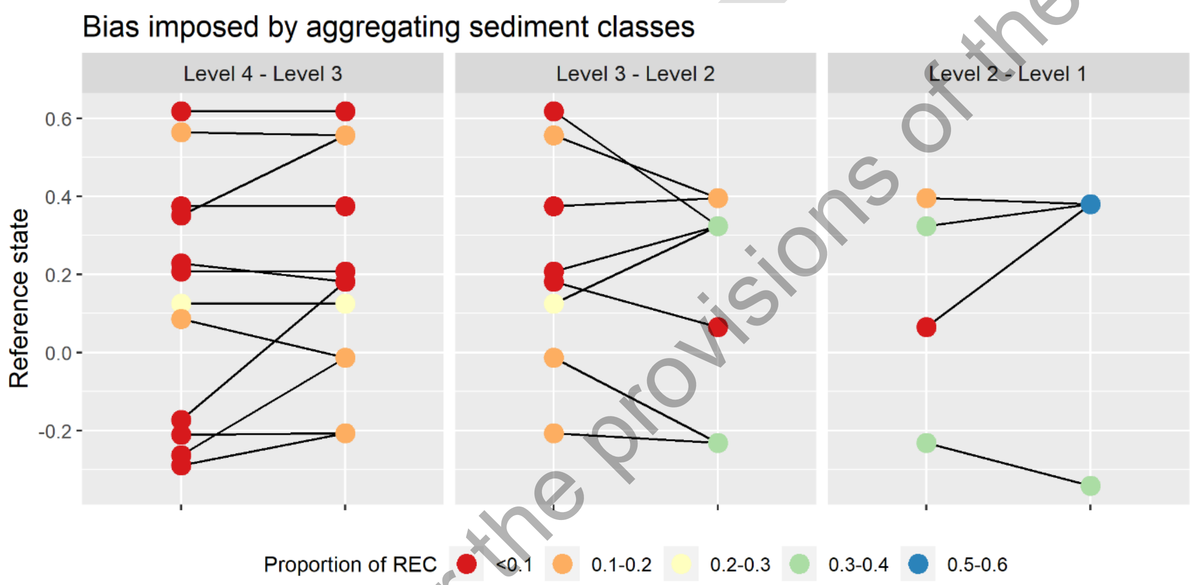

Figure AX.14. Change in reference state of turbidity as classes at one aggregation level are further aggregated into classes at the next highest level in our classification hierarchy.

The AIC statistics for the optimal models of turbidity as a function of stressor covariates Levels 1 through to 4 are presented in Table AX.6. As was the case for deposited fine sediment, and despite the large number of parameters, the most parsimonious model of deposited fine sediment as a function of anthropogenic development is the one that includes interactions between covariates and sediment classes at the lowest level of aggregation (Level 4; 12 classes). Indeed, relative to the other three levels of aggregation in the hierarchy, there is a probability of 1 that Level 4 is the most likely model in the candidate set. Thus the data very strongly indicate that the lowest level of aggregation provides the most parsimonious description of turbidity reference states in New Zealand.

Table AX.6. AIC statistics for optimal models fitted to turbidity data at all four levels of aggregation in the classification hierarchy. K is the number of parameters in the regression model; AICc is the corrected AIC statistic; $\Delta_i = AIC_i - \min(AIC)$ is known as the model rank; w_i, is the Akaike weight of model *i*, interpreted as the approximate probability that Model *i* is the best model in the candidate set, given the data; LL is log-likelihood of each model; Cum.Wt is the cumulative model weight of the ranked models.

Agg. Level	K	AICc	Δ_i	w_i	LL	Cum.Wt
4	25	1069.14	0.00	1.00	-508.87	1.00
3	17	1097.42	28.28	0.00	-531.38	1.00
2	9	1104.84	35.69	0.00	-543.32	1.00
1	9	1123.93	54.79	0.00	-552.87	1.00

The spatial distribution of the turbidity sediment classes at all level of aggregation is presented in Fig. AX.14.

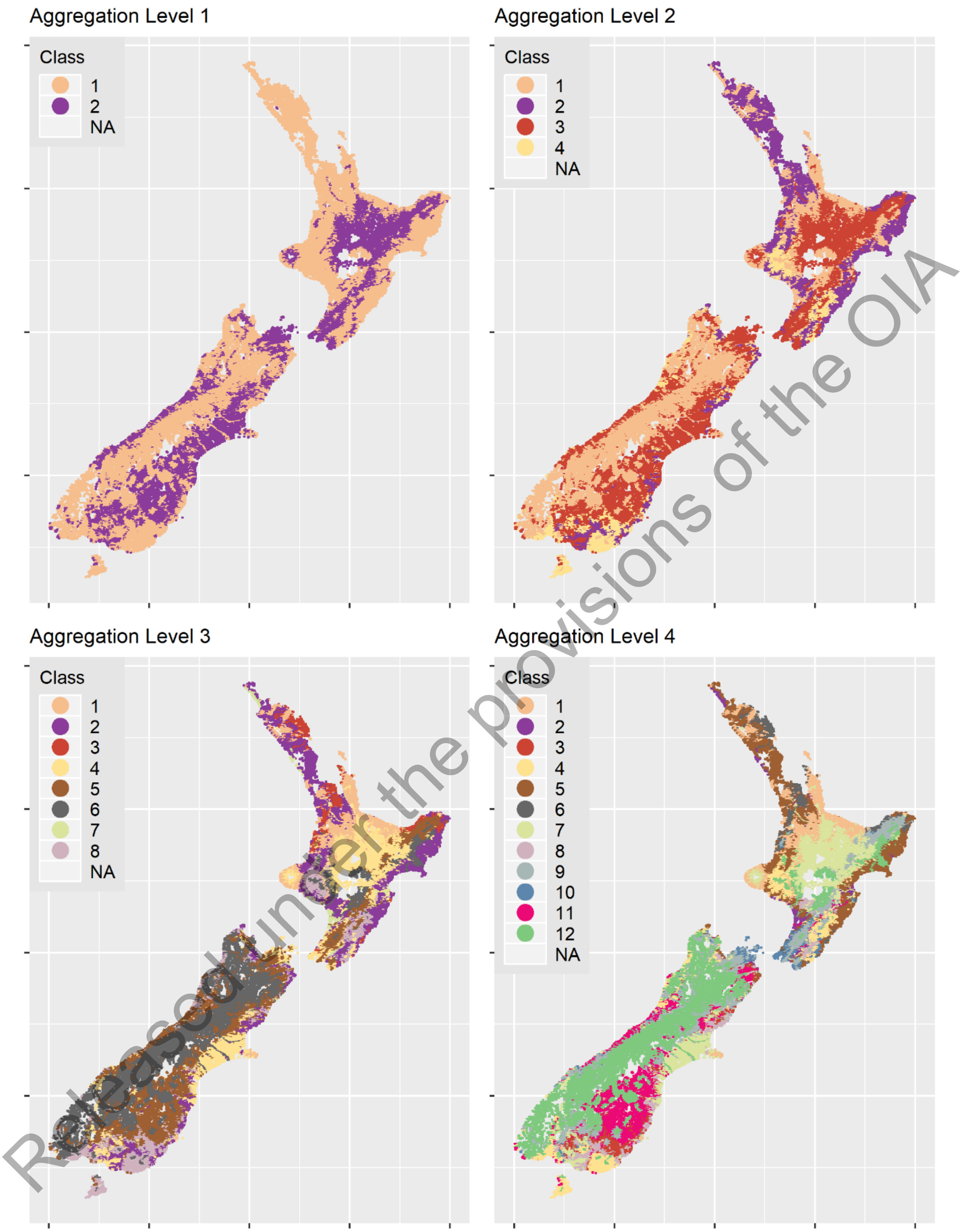

Figure AX.14. Spatial distribution of the deposited fine sediment classes under four different levels of aggregation of the REC Climate-Topography-Geology (CTG) classes. See Table AX.3 for description of sediment classes.

References

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. 2nd edition. Springer, New York.

Dodds, W. K., and R. M. Oakes. 2004. A technique for establishing reference nutrient concentrations across watersheds affected by humans. Limnology and Oceanography: Methods **2**:333-341.

Lewis, W. M., J. M. Melack, W. H. McDowell, M. McClain, and J. E. Richey. 1999. Nitrogen yields from undisturbed watersheds in the Americas. Biogeochemistry **46**:149-162.

McDowell, R. W., T. H. Snelder, N. Cox, D. J. Booker, and R. J. Wilcock. 2013. Establishment of reference or baseline conditions of chemical indicators in New Zealand streams and rivers relative to present conditions. Marine and Freshwater Research **64**:387-400.

Oksanen, J. 2015. Multivariate Analysis of Ecological Communities in R: vegan tutorial. http://cc.oulu.fi/~jarioksa/opetus/metodi/vegantutor.pdf.

Oksanen, J., F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. B. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, S. E., and H. Wagner. 2018. vegan: Community Ecology Package. R package version 2.5-3. https://CRAN.R-project.org/package=vegan.

Snelder, T. H., and B. J. F. Biggs. 2002. Multiscale River Environment Classification for water resources management. Journal of the American Water Resources Association **38**:1225-1239.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications **16**:1267-1276.