Catena 163 (2018) 130–146

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena

Predicting river bed substrate cover proportions across New Zealand

Arman Haddadchi*, Doug J. Booker, Richard J. Measures

National Institute of Water and Atmospheric Research, Christchurch, New Zealand

A R T I C L E   I N F O

Keywords:
River-bed
Substrate sizes
National predictions
Downstream fining

A B S T R A C T

Predictions of river bed substrate cover are required for various purposes including delineating management zones, linking with ecological status and assessing river rehabilitation options. Three contrasting methods were tested for predicting the proportion of river bed covered by seven different substrate categories: generalised linear models (GLMs), machine learning regression models (random forest), and a summed normal distribution model (SND) which incorporates distribution of predictors and substrate covers throughout the modelling framework. Various predictors representing climate, geomorphology, land cover and geology were derived from existing environmental databases to generate predictive models. Model performance was assessed through a cross-validated comparison with substrate samples collected from 229 river sites distributed across New Zealand. Model performance for 10-fold cross-validated predictions showed that the SND model performed best in predicting the proportions of riverbed covered by bedrock, boulder, cobble and fine gravel categories. Random forest models performed best in predicting coarse gravel, sand and mud plus vegetation proportions. Therefore, combined random forest and SND methods were used for estimating substrate cover proportions at unsampled sites across New Zealand. Texture analysis of predicted substrate cover consistently showed downstream fining of sediment size. The national predictions of substrate cover proportions are key descriptors that can be linked with a wide range of national scale applications for ecological assessment of New Zealand Rivers. The techniques developed and tested are applicable to other locations but it is notable that relatively poor performance in regional cross-validation tests shows that transferability of substrate models to locations with no calibration data is challenging.

1. Introduction

There is a growing requirement for exploring the controls on, and prediction of, substrate cover in rivers. Aquatic biota show strong responses to substrate movement as a direct mechanistically linked indicator of bed disturbance (Jellyman et al., 2013). Sedimentary conditions influence macroinvertebrate community structure (Rempel et al., 2000). River bed grain size also influences suitability of spawning, rearing and feeding habitats for many fish species, particularly salmonids (Kondolf and Wolman, 1993; Armstrong et al., 2003; Hedger et al., 2006). Without suitable stream habitat a given species is unlikely to exist at that particular location (Reiser, 1998; Maddock, 1999). Obtaining a detailed knowledge about the characteristics and spatial distribution of river bed substrate cover over a variety of spatial scales is therefore essential for ecological assessment of rivers.

Understanding longitudinal variations in river bed grain size is important as it has a dominant control on geomorphological and sedimentological regimes. Rivers generally show a downstream fining of sediments (Church and Kellerhals, 1978; Rice, 1998; Morris and Williams, 1999; Ferguson, 2003; Costigan et al., 2014). River bed grain size affects abrasion rates (Frings, 2008), rate and mode of sediment transport (Wilcock and Crowe, 2003; Haddadchi et al., 2013), type and dimension of river bed forms (Buffington and Montgomery, 1997; de Almeida and Rodríguez, 2011), and the size of channel bank deposits (Ten Brinke et al., 2004). Downstream fining of bed material occurs in both gravel-bed and sand-bed rivers (Frings, 2008). However, this general trend can be interrupted by: (i) sedimentation processes in lakes, reservoirs and water conveyance structures; (ii) tributaries which introduce large sedimentary inputs to significantly punctuate this fining trend (Rice, 1998; Benda et al., 2004); (iii) dominated proximal sediment sources from surface soils with dissimilar characteristics established independently from upstream catchment surface soil sources (Haddadchi et al., 2015).

Fining of river bed sediments over the longitudinal profile is commonly modelled using a downstream exponential decrease in grain size:

$$D = D_0 e^{\alpha L} \tag{1}$$

where D in Eq. (1) is particle size characteristics (i.e., median

* Corresponding author at: 10 Kyle Street, Riccarton, Christchurch 8011, PO Box 8602, Christchurch, New Zealand.
E-mail addresses: arman.haddadchi@niwa.co.nz (A. Haddadchi), doug.booker@niwa.co.nz (D.J. Booker), richard.measures@niwa.co.nz (R.J. Measures).

https://doi.org/10.1016/j.catena.2017.12.014
Received 6 January 2017; Received in revised form 27 November 2017; Accepted 13 December 2017

diameter), D_0 is initial particle size diameter (i.e., particle size of most upstream sediment), L is distance downstream (in km) and α is an empirical diminution coefficient (in km^{-1}).

In addition to the effect of chemical weathering and abrasion in situ (Miller et al., 2014; Menting et al., 2015), differential mobility of coarse and fine grains within the bed sediment mixtures leads to downstream fining (Parker and Toro-Escobar, 2002). Therefore, the diminution coefficient reflects cumulative effects of both abrasion and sediment sorting and, thus, it depends on lithology, channel morphology and flow and sediment transport conditions (Powell, 1998).

Measurements of river bed substrate proportions have been carried out for assessing stream habitat in the USA (Herbst and Suk, 2005), New Zealand (Harding et al., 2009) and elsewhere. However, because of the temporal and financial limitations of monitoring river bed grain size via direct observation (Wright et al., 1998), the application of indirect methods based on topographic mapping analyses and remote sensing are growing fast. Channel morphologic measurements derived from traditional digital elevation models together with empirical hydrologic methods have been used to predict bed grain size (Buffington et al., 2004; Gorman et al., 2011). Airborne LiDAR data has been used to identify potential habitat in catchments by estimating river bed grain size (Wilkins and Snyder, 2011; Carbonneau et al., 2012; Rinaldi et al., 2013; Snyder et al., 2013). The main limitation of these approaches is their low accuracy when applied to wetted areas of rivers, especially in rivers with high turbidity and large water depth (Groll et al., 2016).

New Zealand has strong gradients in climate, geology, topography and hydrological regime at the national scale. Various river and catchment characteristics have been mapped onto a national river network describing the spatial configuration of New Zealand's rivers (Snelder and Biggs, 2002). Each segment of the river network has characteristics assigned to it including: catchment area, stream length, elevation and slope derived from digital elevations models; catchment geology derived from geological maps; land cover from remote sensing data; and runoff, rainfall and potential evapotranspiration from climate station data (Leathwick et al., 2011). These characteristics have previously been used to predict the spatial distribution of invertebrate communities (Booker et al., 2015), various fish species (Crow et al., 2013), availability of physical habitat (Snelder et al., 2011a, 2011b; Booker, 2016), hydrological indices (Booker and Woods, 2014), and hydraulic geometry (Booker, 2010) in rivers across New Zealand.

The aim of this study was to predict spatial patterns in substrate characteristics of alluvial river channels across New Zealand from nationally available site and catchment characteristics. To do this, three models with different levels of complexity, data needs and user inputs were used to predict substrate proportions; a generalised linear model (GLM) using an ordinary linear regression, random forest (RF) using machine learning to fit a flexible regression, and summed normal distribution (SND) representing a complex model using distribution of predictors for selection procedure together with genetic algorithm procedure to optimise the results.

The study objectives were: (1) to apply various statistical techniques to elucidate the distribution of river bed substrate covers as a function of upstream catchment characteristics incorporating climate, geomorphology, land cover and hydrological factors; (2) to compare the predictive performance of these techniques when used to make predictions at unvisited sites; (3) to predict river bed substrate proportions at unsampled rivers across New Zealand based on the best performing models; and (4) to increase understanding of controls on sediment characteristics at the national scale.

Fig. 1 outlines the strategy used to predict the substrate cover proportions for river reaches across New Zealand. It involved calculating the areal proportions for each substrate category for each site, extracting predictors and selecting independent variables using combined expert opinion and chi-square tests, independency tests or automated procedures (depending on the type of model being fitted), fitting various types of model to predict each substrate category, and calculating predicted values across the entire river network. Substrate cover proportions calculated from each model type were compared. Several performance metrics were then used to quantify predictive performance.

2. Materials and methods

2.1. Site substrate observations

Field data were assembled from physical habitat studies applied by NIWA (National Institute of Water and Atmospheric Research, New Zealand) and various regional councils at 284 sites across New Zealand. At each site, areal proportions of bedrock (> 512 mm), boulder (256–512 mm), cobble (64–256 mm), gravel (8–64 mm), fine gravel (2–8 mm), sand (0.06–2 mm), mud (< 0.06 mm) and vegetation were observed visually at discrete observation locations across multiple cross-sections. Observation locations were centred at regular intervals across each cross-section except on sections with abrupt changes in bed height, where extra observations were added. Cross-sections were positioned to represent all meso-habitat types (e.g., pool, riffle, run) present within each site. Cross-section average sediment cover by each substrate category was calculated using a weighted mean, with weightings based on the separation of observation points. Reach averaged sediment cover was calculated as a weighted mean of cross-section cover with weightings based on the number of sampling cross-sections, and the proportion of the entire reach area, covered by each meso-habitat type. See Jowett et al. (2008) for further details of field procedures. In total, 73,550 observations were included in the data set (an average of 259 per site). The reach length surveyed at each site ranged from 30 to 3000 m, averaging 330 m per site. The average number of cross-sections at each site was 14, and the average spacing between observation points was 0.84 m. Sampling sites were located throughout the New Zealand river network (Fig. 2) and represented a wide range of river sizes, climatic, topographic and hydrological conditions. See Booker (2016) for further details. Particle size distribution of observed substrates varied between sites, with median diameter (D_{50}) of substrate materials ranging from < 0.06 mm to larger than 100 mm.

Grid co-ordinates and site descriptions from various data providers were used to identify which of the 570,000 reaches that comprise the New Zealand river network best represented the position of each site. There were 37 reaches that had more than one sampled site (two to five sites) assigned to them. The proportion for each substrate category averaged over all sites assigned to the same reach was used to represent substrate proportions at these reaches. This reduced the number of sampled sites from 284 to 229.

2.2. Nationally available predictors

Many environmental variables have previously been mapped onto the New Zealand river network (Leathwick et al., 2008; Booker et al., 2015) and were therefore available as potential explanatory variables for predictive models (Table 1). Climate is represented by various parameters representing different characteristics of precipitation (i.e., usRainDays10, usRainDays25, usRainDays50, usRainDays100, usRainDays200, usAnRainVar), hydrology (usFlow, SpecificMeanFlow, SpecificMALF, FRE3, SpecificAnnualFlood) and temperature (segEquiTSum, segEquiTwin, usPET). Geomorphology is represented by eight parameters such as upstream catchment area (usArea) which is strongly related to wetted width of the river segment (Booker, 2010), average slope of catchment (usAveSlope) calculated from 30-m digital elevation model (DEM), and distance from the coast (dsDistToSea) indicating the location of the site in the river network. Land cover is represented by the proportion of surface area occupied by five categories of land cover (usPastoral, usIndigForest, usExoticForest, usUrban, usScrub; see Table 1 for details). Geology of the upstream catchment, which has a strong influence on the bed material cover of downstream river reaches,

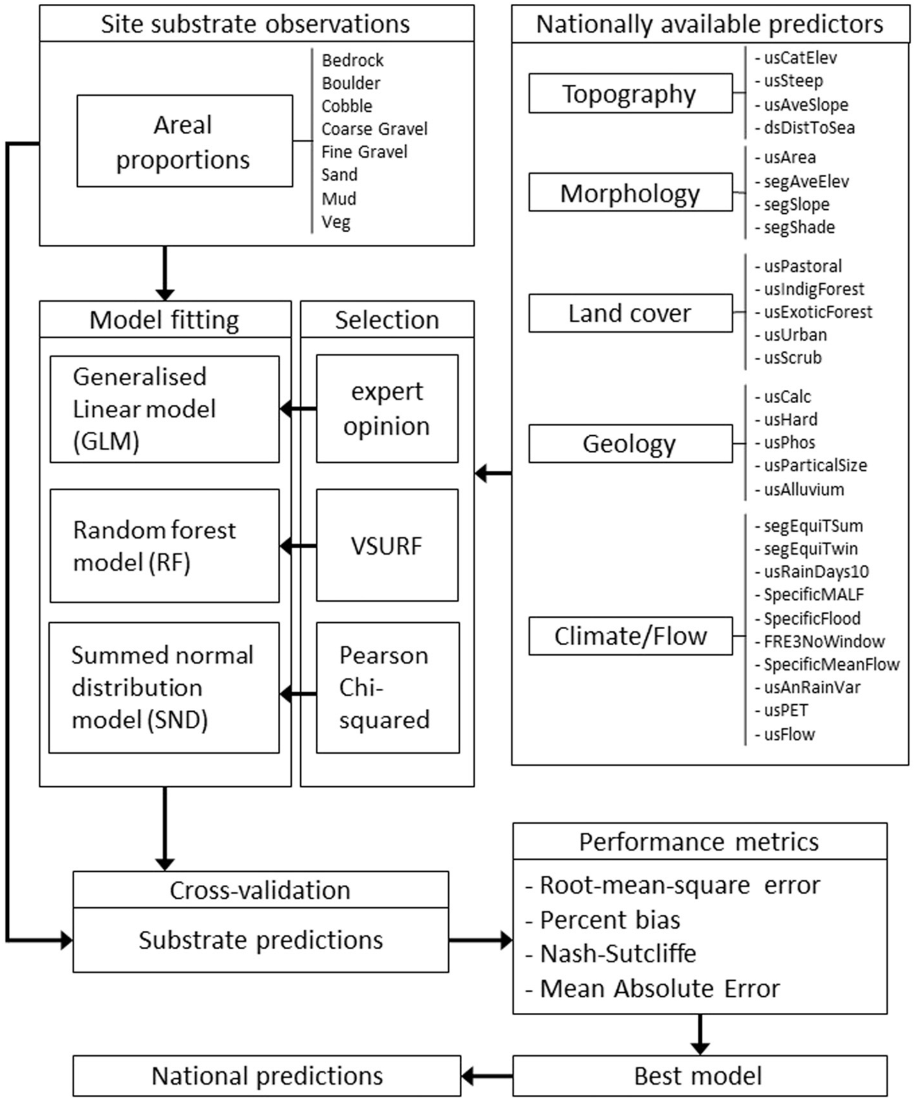

Fig. 1. Schematic showing strategy for estimating river bed substrate areal proportions including statistical tests, modelling and validation.

is represented by five parameters such as proportion of catchment area occupied by alluvium land resource inventory (usAlluvium), and average particle size of underlying rock from hillslopes distributed throughout the upstream catchment area (usParticleSize). The usParticleSize parameter is derived from the Land Environments of New Zealand (LENZ), and it represents the particle characteristics of the geological regolith on an ordinal scale (Leathwick et al., 2002). Previous studies have directly compared the size distribution of sediments in channels and adjacent hillslope sources and they suggest that there may be general relationship between the size of sediments originated from hillslopes and supplied to channel banks (Marshall and Sklar, 2012; Attal et al., 2015; Riebe et al., 2015; Sklar et al., 2017). These associated databases have previously been used to define a hierarchical classification of New Zealand's rivers called the River Environment Classification (REC). For example, a priori defined topographic groupings have been applied to classify each river segment as being either Glacial Mountain (permanent ice > 1.5%), Mountain (> 50% annual rainfall volume above 1000 m ASL), Hill (50% rainfall volume between 400 and 1000 m ASL), Lowland (50% rainfall below 400 m ASL), or Lakefed (Lake influence index > 0.033). See Snelder and Biggs (2002) and Snelder et al. (2005) for full details.

Data transformations were used to make highly skewed distributions less skewed and restore symmetry to the data. These transformations are valuable both for making patterns in the data more interpretable and for helping to meet the assumptions of inferential statistics, for instance, when applying regression methods, normal distributions are desirable. Therefore, predictor variables with skewed distributions were square-root or log10 transformed to aid interpretation of the fitted relationships by GLM and RF models. For RF models, transformation of predictors should not influence fitting because they are invariant to monotonic transformations of the predictors, but may influence calculated importance of predictors.

2.3. Estimation methods

For this study three methods for predicting river bed substrate proportions were compared. The three methods were selected and applied deliberately in particular ways such that they represented contrasting abilities to incorporate reduction of potential predictors, nonlinear responses, interactions between predictors, and interactions between responses. A brief description of the first and second methods is provided as these methods are well established. The third method is less established, therefore a full description is provided.

The first method was to fit a generalised linear model (GLM; Nelder and Wedderburn, 1972) to each substrate category separately. GLMs were deliberately designed to represent relatively simple regression models. GLMs apply a flexible generalisation of ordinary linear regression to accommodate both non-normal response distributions and

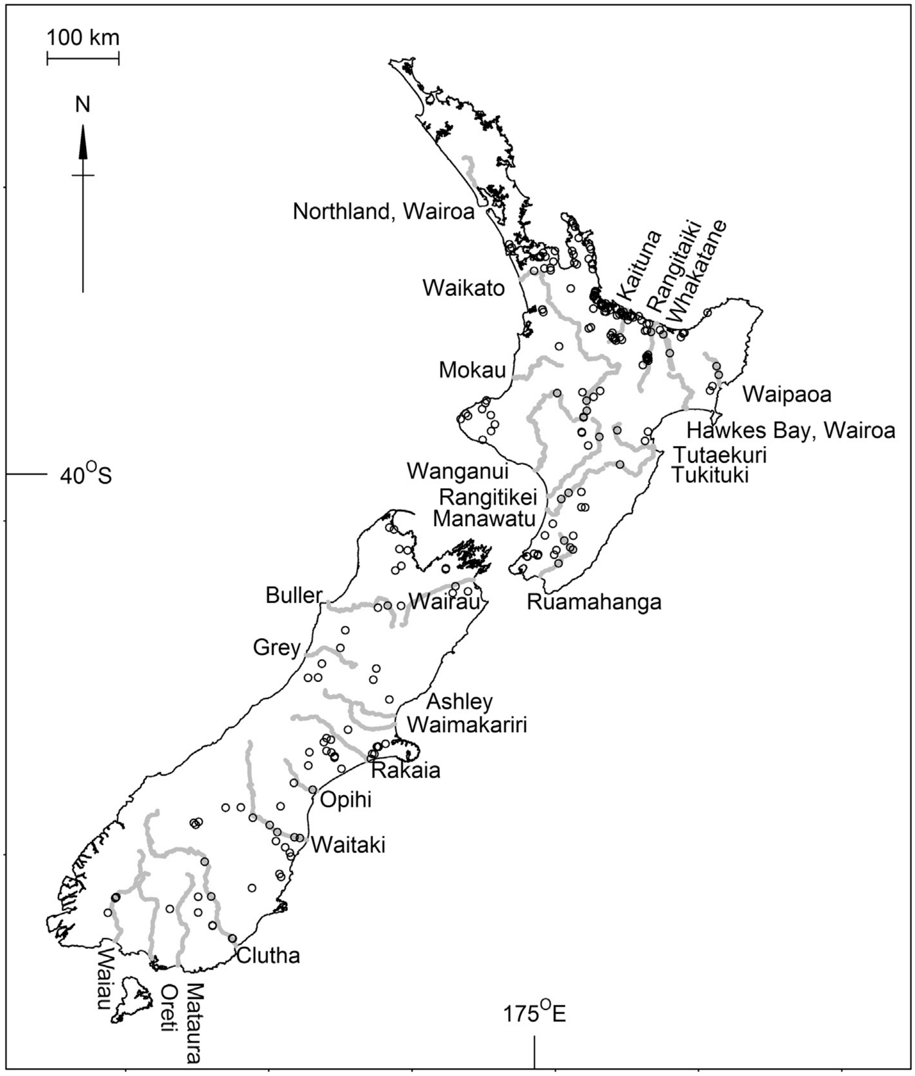

Fig. 2. Map showing sampled reaches (n = 229) and courses from highest source to the sea for some main rivers.

transformations to linearity (Venables and Ripley, 2002). They allow the linear model to be related to the response via various link functions by allowing the magnitude of the variance of each observation to be a function of its predicted value. A logit link function and binomial error distribution were applied in each GLM as is appropriate when the response is proportion data. Predicted responses were therefore always monotonous. The same limited set of predictors was applied to each substrate category with no consideration of interactions between predictors.

The second method was to fit random forest (RF) models (Breiman, 2001), similar to that applied by Snelder et al. (2011b) to predict mean grain size across the French river network. This method was deliberately selected and applied such that it represented a set of relatively complex relationships which are able to consider interactions between predictors and complex forms of responses, but not interactions between substrate categories. This method uses machine learning to fit a flexible regression representing the relationship between combinations of predictors and the areal cover proportion of each substrate category separately.

The third method was to fit summed normal distribution (SND) models. This method was deliberately selected and applied such that it represented a complex regression method that incorporates distributions of predictors and substrates rather than using central tendency parameters throughout the entire modelling framework. The technique of assigning distributions is similar to the approach described by Olley et al. (2013) and Haddadchi et al. (2014) for tracking sediment sources using distribution mixing models.

The first step in applying the SND model is to standardize the outputs (substrate cover proportions in this case) and predictor variables, and to derive distributions for individual outputs and variables by summing normal distributions centred around each sampled variable/output value using kernel density estimation. Fig. 3 shows the distribution of substrate cover proportions estimated using summed normal distributions, as well as the individual sample cover proportions and their assigned uncertainties (whisker plots below the main plot). The kernel density distribution for each predictor variable, P, and type of substrate cover, S, are given by:

$$P_i = \sum_{j=1}^{m} \int_{a}^{b} \frac{1}{\sqrt{2\pi\sigma p_{i,j}^2}} e^{-\frac{(xp_{i,j}-\mu p_{i,j})^2}{2\sigma p_{i,j}^2}} \tag{2}$$

$$S_k = \sum_{j=1}^{m} \int_{a}^{b} \frac{1}{\sqrt{2\pi\sigma s_{k,j}}} e^{-\frac{(xs_{k,j}-\mu s_{k,j})^2}{2\sigma s_{k,j}^2}} \tag{3}$$

where: i is the predictor variable index (i = 1, …, n), j is the sample number (j = 1, …, m) and k is the substrate category index (k = 1, …, z); $\mu p_{i,\ j}$ and $\mu s_{k,\ j}$ are the predictor and substrate cover values respectively (for a specific substrate/output and sample); and $\sigma p_{i,\ j}$ and $\sigma s_{k,\ j}$ are the uncertainty of each sample's predictor/substrate cover value. The uncertainty values are set equivalent to the standard error of the

Table 1
Summary of the climate/flow, geomorphology, land cover and geology predictors together with their mean and standard deviation (SD) for all 229 datasets. [L] = log transformation applied for RF and GLM models. [R] = square root transformation applied for RF models.

	Parameter	Unit	Mean	SD	Description
Climate/flow	segEquiTSum	°C	20.88	1.9	Summer (January) equilibrium air temperature at location of river segment
	segEquiTwin	°C	6.92	2.36	Winter (June) equilibrium air temperature at location of river segment
	usRainDays10	Number of days/yr	3.58	1.12	Average number of days per year within catchment with rainfall intensities > 10 mm/month
	usRainDays25	Number of days/yr	1.14	0.56	Average number of days per year within catchment with rainfall intensities > 25 mm/month
	usRainDays50	Number of days/yr	0.27	0.21	Average number of days per year within catchment with rainfall intensities > 50 mm/month
	usRainDays100	Number of days/yr	0.04	0.05	Average number of days per year within catchment with rainfall intensities > 100 mm/month
	usRainDays200[L]	Number of days/yr	0.004	0.01	Average number of days per year within catchment with rainfall intensities > 200 mm/month
	usAnRainVar	mm	175.5	21.3	Coefficient of variation of annual catchment rainfall
	usPET	mm	954.6	152.2	Annual potential evapotranspiration of catchment
	usFlow	m^3/s	26.2	82.4	Total annual runoff volume
	FRE3	events/yr	12.9	5.02	Number of events per year exceeding three times the median flow based on mean daily flows, with no windows to account for successive events applied.
	SpecificMALF[L]	$m^3/s/km^2$	0.009	0.007	Mean annual low flow per unit catchment area
	SpecificAnnualFlood[L]	$m^3/s/km^2$	0.435	0.274	Mean annual flood (derived from mean daily flows) per unit catchment area
	SpecificMeanFlow[L]	$m^3/s/km^2$	0.0364	0.0256	Mean flow per unit catchment area
Geomorphology	segSlope[R]	m/m	0.006	0.01	Average segment slope
	segShade	%	0.17	0.19	Estimated riparian shade
	usArea[L]	km^2	810.5	2523	Upstream catchment area
	segAveElev[L]	m	146.3	165.4	Average elevation of the segment
	usCatElev	m	501.3	307.9	Average elevation of catchment
	usSteep	%/100	0.14	0.17	Proportion of catchment with slope < 30°
	usAveSlope	Degree	14.2	6.5	Average slope of catchment calculated from 30 m DEM grid
	dsDistToSea	km	70.9	91	Distance of the segment to the coast
Land cover	usPastoral	%/100	0.31	0.29	Proportion of catchment covered by pastoral land cover
	usIndigForest	%/100	0.33	0.29	Proportion of catchment covered by indigenous forest land cover
	usExoticForest[R]	%/100	0.07	0.14	Proportion of catchment covered by exotic forest land cover
	usUrban[R]	%/100	0.003	0.02	Proportion of catchment covered by urban land cover
	usScrub[R]	%/100	0.08	0.11	Proportion of catchment covered by scrub land cover
Geology	usCalc	Ordinal scale	1.51	0.44	Catchment average of calcium of underlying rocks, 1 = low to 4 = high
	usHard	Ordinal scale	3.1	0.72	Catchment average of hardness of underlying rocks, 1 = low to 5 = high
	usPhos	Ordinal scale	2.33	0.94	Catchment average of phosphorous concentration of underlying rocks, 1 = low to 5 = high
	usParticleSize	Ordinal scale	2.78	1.2	Catchment average of particle size of underlying rocks, 1 = fine to 5 = coarse
	usAlluvium[R]	%/100	0.084	0.17	Area of catchment covered by alluvium land resource inventory

mean for the parameter being considered, derived from standard deviation of datasets divided by the square root of sample size.

To solve indefinite Gaussian (normal) integrals in Eqs. (2) and (3), 1000 increments with bin width of 1.2×10^{-3} were used. The terms a and b in these equations were the lower and upper limits of integration. Due to pre-normalizing of predictor values and substrate proportions, integral limits of -0.1 to 1.1 were applied.

The second step is, for models of each substrate class, to determine the relative coefficient of each predictor. In SND, instead of calculating a single value as the coefficient of each environmental factor on the model, their normal distributions were estimated. Model fitting is carried out by simultaneously minimising the sum of squared relative errors:

$$E = \sum_{k=1}^{z} \left(\frac{\sum_{i=1}^{n} P_i X_{i,k} - S_k}{S_k} \right)^2 \tag{4}$$

where X is the normal distribution of coefficients with mean value ($\mu x_{i,k}$) as coefficient value for each environmental predictor and standard deviation ($\sigma x_{i,\,k}$) as uncertainty of their calculations:

$$\mathrm{X_{i,k}} = \frac{1}{\sqrt{2\pi\sigma x_{i,k}^2}} e^{-\frac{(x_{i,k}-\mu x_{i,k})^2}{2\sigma x_{i,k}^2}} \tag{5}$$

Mean values in coefficients are modelled as truncated normal distributions:

$$-1 < \mathrm{X_{i,k}} < 1 \tag{6}$$

As local optimisation methods can fail to identify globally representative solutions of the model, genetic algorithm optimisation with 1000 iterations was deployed to find optimum coefficients. Minimising the sum of squared relative errors (E) was used as the objective function for this optimisation (Eq. (4)). The Optquest algorithm in Oracle's Crystal Ball software (Oracle, 2015) was used for this task. In addition, to select the values within the distributions, the Latin Hypercube Sampling (LHS) method with 500 iterations was applied. Although these make the optimisation procedure more time consuming, the precision of results significantly increases compared with local optimisation methods. Initial values for all predictions were selected using genetic algorithm optimisation estimates of linear regressions.

Once optimum distributions of coefficients for $\mathrm{X_{i,k}}$ have been found, the SND models can be used to derive predictions of substrate cover across new Zealand. The major advantage of SND over the RF and GLM is that for each new prediction of each substrate category, uncertainty is calculated explicitly within the model fitting procedure. Therefore, to make predictions, 10 random values (R = 10) of predictors' contributions within the range of their normal distributions ($\mathrm{X_{i,r}}$), were generated to predict substrate proportions and their related uncertainty across New Zealand rivers:

$$\sum_{\mathrm{r=1}}^{\mathrm{R}} \sum_{\mathrm{i=1}}^{\mathrm{n}} (\mathrm{P_{i\text{-}unsampled} X_{i,r}})/\mathrm{R} \tag{7}$$

where $\mathrm{P_{i\text{-}unsampled}}$ were predictors for river reaches across New Zealand.

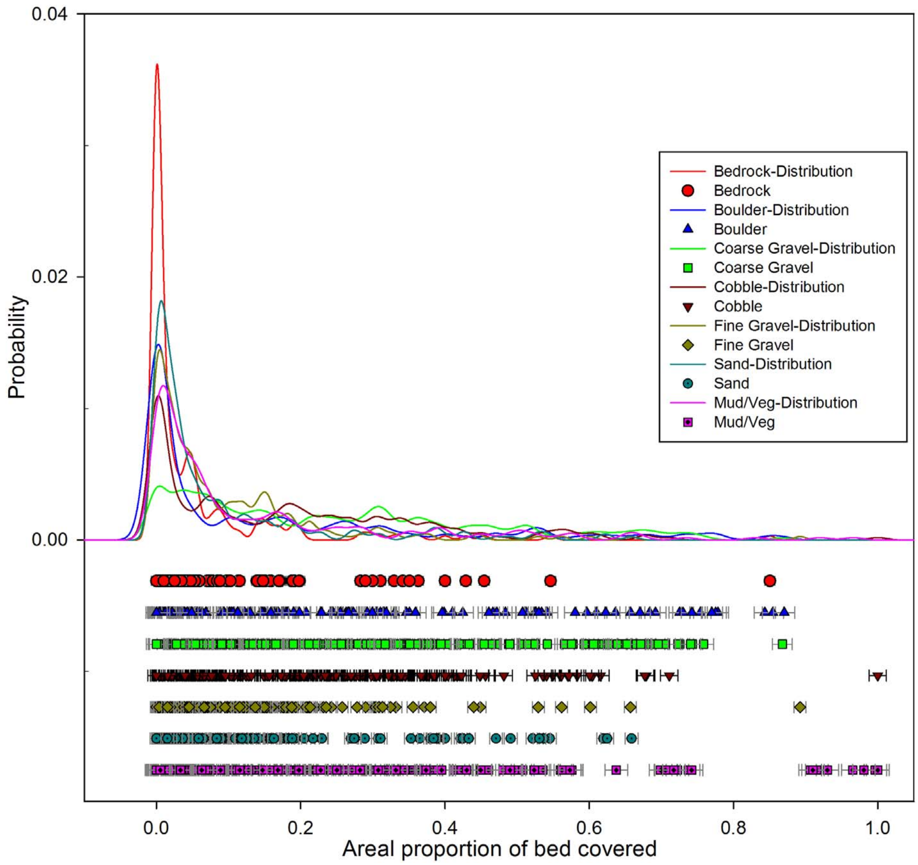

Fig. 3. Probability distribution of substrates used in the summed normal distribution model (see Eq. (4)). Points under the distributions are the 229 sampled sites colour-coded to seven river bed substrate categories used to derive theses distributions using summed normal density function as described in Eq. (4). The error bars are equivalent to one standard error on the mean and are derived from all datasets on each category.

2.4. Selection of predictor variables

Different approaches for selecting predictor variables were applied for the three methods: GLM, RF and SND. The complexity of the approaches for selecting variables was chosen in-line with complexity of the method. Key aims for variable selection were to identify the parameters which were most useful for predicting substrate, reduce the number of variables to prevent over-parameterisation of the models, and to avoid parameters which were strongly dependent.

For the relatively simple GLM method, four predictors were selected using expert opinion. The predictors were selected as being the most appropriate to characterise patterns in the dominant processes controlling sediment deposition. The availability and size-distribution of sediment supply was represented by catchment average of particle size of underlying rocks (usParticleSize). Flow magnitude was represented by catchment runoff (Log10 of SpecificMeanFlow). Flow variability was represented by the number of events exceeding three times the median flow (FRE3). Geomorphological setting including slope and valley confinement was represented by the REC Topographic class (e.g., Mountain, Hill, Lowland) as described above. No interactions were included in any GLMs and no methods for variable reduction were applied. Chi-square tests were applied to assess the significance of predictors.

For the second method of analysis, RF model, an initial long-list of 22 predictors was selected from those available (Table 1) based on expert knowledge. This reduced list was selected to eliminate some parameters which were known to be strongly dependent. An automated variable selection procedure known as VSURF was then applied in which predictors are eliminated through a step-wise ascendant strategy designed to decrease the error rate (Genuer et al., 2010; Genuer et al., 2015). VSURF was applied separately for the different substrate classes, allowing automated identification of a reduced number (between four and eight) of the most important predictors for each class, which were then used to produce the models.

Since co-varying predictors can lead to biased results in the third method of analysis, SND models, correlation between predictor variables was examined first, prior to developing the models. To assess whether predictors were independent of each other, a test of independence using two tailed Pearson's chi-squared was applied to investigate relationships between each predictor and every other predictor. Significant relationships between predictors (P-value < 0.01%) were identified, and predictors were eliminated manually to create a reduced list of 13 independent predictors (i.e. no significant relationships between them). The same predictor variables were selected across substrate classes for the SND model.

2.5. Performance metrics

Two cross-validation techniques, k-fold and regional, were applied to quantify predictive performance independently from the fitted data set. The same cross-validation techniques were applied across models to ensure consistency when evaluating performance.

A k-fold cross validation was carried out by dividing the observed sites into 10 subsets. For each model, for each subset, predicted values were obtained after having fitting to all data from the remaining 9 subsets. As SND models require a minimum number of datasets to derive distributions using kernel density function, leave-one-out (LOO) cross validation or k-fold cross-validation with smaller groups of subsamples (i.e. higher number of folds) were not appropriate for model evaluation. It should be noted that the cross-validation technique applied can be used to indicate model performance at unvisited sites within the environmental range of the observed sites, but does not elucidate on performance outside of that environmental range.

Regional cross validation was undertaken by splitting the dataset into two exclusive subsets containing sites from the North Island and South Island respectively. Predicted substrate cover proportions for

North Island sites were obtained after fitting the models to 73 datasets from the South Island, and predictions for the South Island sites, were obtained after fitting to 156 North Island sites.

Following the recommendations of Moriasi et al. (2007) and Muleta (2012), predictive performance was assessed using four performance metrics regardless of model type or substrate category. Mean absolute deviation (MAD) is an indicator of the distance between the model predictions and eventual outcomes (Hyndman and Koehler, 2006). Nash-Sutcliffe efficiency (NSE) determines the relative magnitude of the residual variance in the estimated river bed size fractions compared to their measured variance (Nash and Sutcliffe, 1970). Values of one indicate a perfect match between observed and predicted values. Values of zero indicate the same performance as would be the case if all predictions were equal to the mean observed value. Percent bias (pbias), is the ratio of the sum of the river bed size fraction residuals to the sum of the observed size fractions. It indicates the average tendency of the estimates to be larger (negative pbias) or smaller (positive pbias) than their observed counterparts (Gupta et al., 2009). Root-mean-square deviation (RMSD) determines the square root of the variance of the observed minus estimated river bed size fractions. RMSD is a commonly used performance criteria. It quantifies the absolute precision of the predicted values (Gupta et al., 2009). See Moriasi et al. (2007) for further details of how these performance metrics are calculated and how various performance metrics may complement each other when comparing observed and predicted values. The method of Piñeiro et al. (2008) was applied when plotting observed against predicted values.

3. Results and discussion

3.1. Input variables and fitted models

For the GLM models, results from chi-square tests showed considerable variation between the significance of predictors between substrate categories (Table 2). GLMs showed no evidence of overdispersion for any substrate category; residual deviance was always far less than residual degrees of freedom. For example, LogSpecificMeanFlow was significant for all categories except fine gravel and sand. The GLMs also showed strong responses in proportion areal cover to each predictor. For example, the proportion covered by the combined mud and vegetation category (mudveg) decreased with increasing FRE3, usParticleSize and SpecificMeanFlow. The proportion covered by mudveg was greatest in Lakefed and Lowland classes, followed by Hill, Mountain and Glacial Mountain classes (Fig. 4).

Predictors representing slope, elevation, rainfall variability and flow per unit area were often selected by the VSURF procedure to be retained in the RF models. Predictors relating to land cover were not strongly selected (Fig. 5). Selection of variables representing slope and rainfall was consistent with the strongest contributors to estimates of mean grain size using RF models in Snelder et al. (2011b) study.

There was some variability in predictors selected between substrate categories in the RF models (see Fig. 5). Variable importance, as defined by mean decrease in node impurity (Grömping, 2009), also varied between substrate classes, with usCatElev being an important predictor for sand and fine gravel, but usPastoral being the most importance predictor for mudveg and boulder. This variability is likely due to a combination of factors including non-independence of the predictors, complex interactions between predictors and differences in the physical processes that govern the condition of each substrate category causing differences in predictors that were selected for each sediment category. Sediment generation, transport and deposition are governed by a complex trade-off between sediment availability and power to erode or transport. Inconsistency in selected predictors between random forest models for each substrate category indicates that this trade-off may be best represented by a combination of the available predictors, including their interactions.

For the SND model, 13 independent predictor variables were retained after eliminating dependant variables (Fig. 6). From 10 parameters representing climate and hydrology of the catchments, coefficient of variation of catchment rainfall, number of days with rainfall > 200 mm month^{-1} and annual runoff were retained for the modelling analysis. Two geomorphological parameters including slope and average elevation of the segments and four land use catchment proportions (i.e. exotic forest, indigenous forest, scrub and urban) were also used as the input of the distribution model. Four out of five variables which represent geology of the upstream catchments including average phosphorus and calcium concentration, particle size of underlying rocks and percentage of catchment covered by alluvium successfully passed the independency test.

The estimated normal distribution of the predictors' contribution derived from the SND models are shown in Fig. 7 for all eight river bed substrate categories. The SND model allowed variations in each of the coefficients of Eq. (5) to vary within each river bed substrate category. Therefore, the mean and standard deviation of normal distributions were determined separately for each substrate and predictor. The absolute magnitude of different predictor coefficients for each substrate indicates the strength of a predictors influence on the modelled substrate cover. The sign of the coefficient indicates the direction of correlation.

In general, the slope of the reach on which sites were located (segSlope) had highest contributions to the prediction of river bed substrate covers. The relative contribution of slope to coverage of different substrate classes is interesting in that roughly scales with grain size. The mean coefficient for segSlope has a value of 0.40 for boulder, 0.25 for cobble, 0.15 for coarse gravel, − 0.05 for fine gravel, and − 0.10 for sand.

Other predictors tend to differentiate one substrate class from others. For example, usRainDays200 (average number of days per year within catchment with rainfall intensities > 200 mm/month) has a very strong positive coefficient for coverage of boulder substrate but small negative coefficients for all other substrates. This shows that this parameter is important to identify locations with high coverage of boulders. The physical basis for this may to be that the threshold of 200 mm/day is a very high rainfall intensity, which is likely to be associated with river conditions capable of transporting boulders (given sufficient slope, also a very strong predictor of boulders as previously mentioned).

The coverage of mud and vegetation (mud/veg), and bedrock both have very different predictor coefficients compared with other substrate types. This is unsurprising given that they are both influenced by somewhat different processes compared to the other substrate types. Mud/veg substrates have cohesive sediment unlike other substrate classes, and bedrock represents an absence of sediment rather than presence of a particular size. Predictors which have a strong influence on bedrock are slope and usAlluvium. There is a logical physical basis for both of these parameters having a strong influence: high slope causes hydraulic conditions where even the largest sediment can be scoured out of a reach, and presence of alluvial deposits means that there is less likely to be any bedrock for scour to expose.

The predictors with the strongest contribution to mud/veg coverage include the land cover parameters usUrban and usExoticForest (% land cover by urban areas and exotic forest respectively). This is interesting as no other substrate category has any land cover parameters identified as one of its important predictors (except sand which has usScrub as an important predictor). This suggests that the proportion of mud and vegetation is more strongly influenced by land cover than other substrate types. The physical basis for this may be that exotic forest and urban areas are significant contributors of fine sediment (the physical basis for usScrub as a predictor for sand coverage is not obvious).

Geology predictors including upstream catchment area covered by alluvium and average concentration of calcium within the catchments has little influence on most substrate classes.

Table 2
Results for chi-square test for GLM models of each substrate categories. Df = Degrees of freedom.

Substrate Category	Predictor	Df	Deviance	Residual Df	Residual deviance	P-value
Bedrock				228	32.0	
	usParticleSize	1	0.035	227	31.9	0.642
	LogSpecificMeanFlow	1	1.851	226	30.1	0.001
	FRE3	1	2.137	225	28.0	0.000
	Topography class	4	1.099	221	26.9	0.151
Boulder				228	80.9	
	usParticleSize	1	0.147	227	80.7	0.496
	LogSpecificMeanFlow	1	12.459	226	68.3	0.000
	FRE3	1	0.770	225	67.5	0.119
	Topography class	4	1.121	221	66.4	0.471
Cobble				228	55.6	
	usParticleSize	1	5.770	227	49.8	0.000
	LogSpecificMeanFlow	1	4.404	226	45.4	0.000
	FRE3	1	1.669	225	43.7	0.002
	Topography class	4	1.827	221	41.9	0.040
Coarse gravel				228	58.2	
	usParticleSize	1	3.644	227	54.6	0.000
	LogSpecificMeanFlow	1	2.270	226	52.3	0.001
	FRE3	1	0.443	225	51.9	0.153
	Topography class	4	1.808	221	50.1	0.080
Fine gravel				228	33.6	
	usParticleSize	1	0.004	227	33.6	0.872
	LogSpecificMeanFlow	1	0.087	226	33.5	0.477
	FRE3	1	0.139	225	33.4	0.368
	Topography class	4	1.025	221	32.4	0.201
Sand				228	39.6	
	usParticleSize	1	4.515	227	35.1	0.000
	LogSpecificMeanFlow	1	0.000	226	35.1	0.967
	FRE3	1	0.129	225	34.9	0.402
	Topography class	4	3.181	221	31.8	0.002
Mudveg				228	79.3	
	usParticleSize	1	11.270	227	68.0	0.000
	LogSpecificMeanFlow	1	8.796	226	59.2	0.000
	FRE3	1	0.015	225	59.2	0.811
	Topography class	4	7.754	221	51.4	0.000

3.2. Model validation

Model performance for 10-fold and regional cross-validated predictions of each river bed substrate category derived from each model is given in Fig. 8.

Using the 10-fold cross validation technique, RMSD values for the SND model ranged between 9.6% to predict bedrock proportions and 19.4% to predict mudveg with average RMSD of 14.2% and for GLM model from 10.4% (for bedrock) to 20.8% (for boulder) with average of 16.1%. For RF models RMSD varied from 9.7% to 19% to predict bedrock and mudveg proportions, respectively, with an average of 14.3%.

Biases were relatively small (pbias ranging from − 1.1% to 0.53%) in GLM models in comparison with the RF models (− 4.5% < pbias < 0.47%) and SND models (− 0.02% < pbias < 2.4%). Low pbias values indicate that the average tendency of simulated variables estimated by GLM models were close to the respective observed substrate classes, however high RMSD values demonstrate large variance in predicted variables. Except for fine gravel predictions using GLM, NSE was positive for all substrate categories and all models. Negative NSE values for fine gravel proportions estimated by the GLM model (i.e. − 0.01) indicate that the mean observed fine gravel proportion is a better predictor than the simulated GLM values.

In General, GLM models were unbiased, but performed less well than RF and SND models. The SND model gave better precision (lower RMSD and higher NSE values) and less biased results (pbias closer to zero) when compared to RF models in predicting bedrock, boulder, cobble and fine gravel areal cover proportions. RF models performed better for coarse gravel, sand and mud proportions.

The regional cross validations showed RMSD values for the SND model ranged from 11% (for Bedrock) to 22% (for mudveg), only slightly worse than for the 10-fold validation. For the RF model, RMSD were very similar to the SND model except for boulder and sand where the RF model performed slightly worse. For all three models, biases were much worse in the regional than 10-fold validation, ranging from − 38 to + 50. Unlike for the 10-fold validation the SND model was the least biased (− 11 to + 17) and showed a different pattern of bias compared to both GLM and RF which were similar.

NSE was negative for all GLM predictions in the regional validation except for mudveg. This indicates large uncertainties of the GLM model on predicting substrate cover proportions outside the region the model were calibrated. RF and SND models performed better but still poorly compared to the 10-fold validation. All models had a negative NSE for fine gravel in the regional validation.

Overall the performance of all three models was worse in the regional than the 10-fold cross validation. Part of the reason for the reduced model performance in the regional cross-validation is the lower number of sites used to fit the models, particularly when predicting substrate at the 156 North Island sites after fitting the model to only 73 South Island sites. However, the poor performance also suggests that the models require training using local data to perform well, despite the comprehensive range of predictor variables including climate/hydrology, geomorphology, land cover and geology. This has implications for development of predictive models of substrate cover outside New Zealand, suggesting that fitting to local data will be required.

It is notable that the SND model performed better than the RF model in the regional validation, despite having similar performance in the 10-fold validation. This indicates that the SND model may give better predictions in locations were no training data are available.

Visual inspection of observed data against 10-fold cross validation predicted values shows that all models generally struggle to predict high coverage by any single substrate type (Fig. 9). This is likely due to

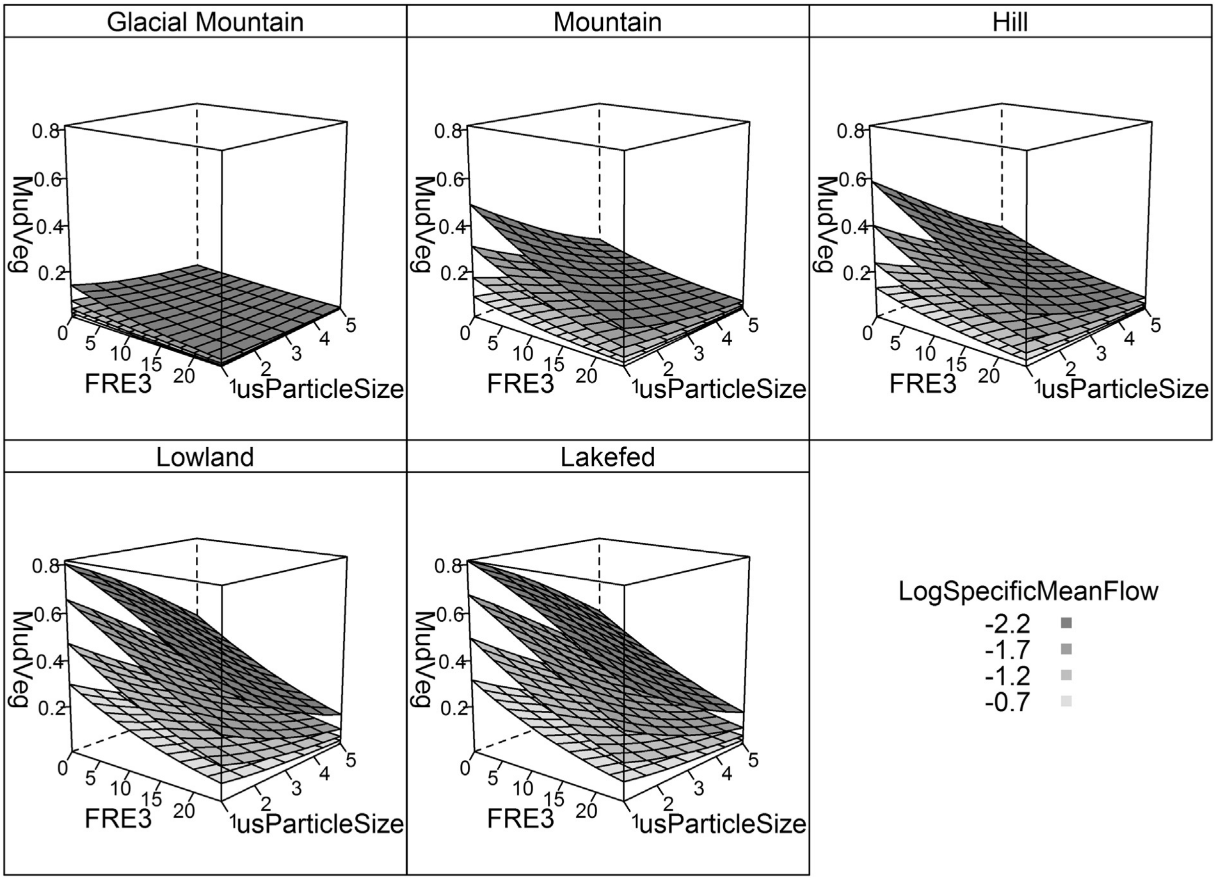

Fig. 4. GLM predicted areal cover by mudveg as a function of runoff (LogSpecificMeanFlow), flow variability (FRE3), average particle size of underlying rock (usParticleSize) for each REC topography class. See Table 1 for variable definitions and units.

difficulty of modelling the skewed distribution of the observed data with many sites having coverage between 0% and 20% of any given substrate, but very few having > 80%.

To evaluate the performance of these predictive models, in addition to their accuracy, level of complexity, number of input parameters, and the time required for preparing and running the models should be considered in future application of these models. For instance, the number of predictors and simulation time for running the simple linear models (GLM) were significantly lower in comparison with the SND models which required distribution of predictors and used a time consuming genetic optimisation procedure.

3.3. Prediction of substrate textures across New Zealand rivers

The best performing models based on k-fold cross validation for each substrate category were used to predict areal proportion for all reaches across New Zealand. Bedrock, boulder, cobble and fine gravel proportions were predicted by the SND model and RF models were used to predict coarse gravel sand and mudveg proportions. Selected catchment and site predictors (see Figs. 5 and 6) for all reaches of the

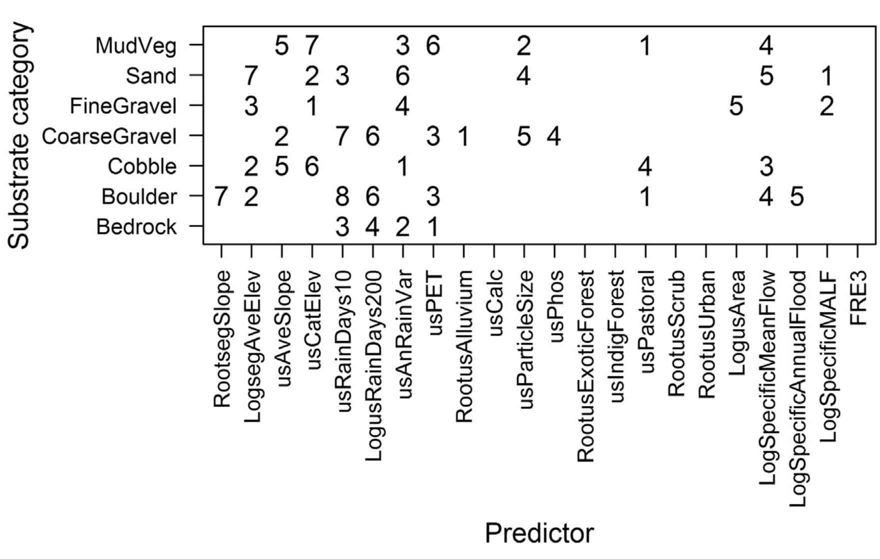

Fig. 5. For each substrate category, ranking of the importance (mean decrease in node impurity) of selected variables for random forest (RF) models after having applied the VSURF procedure.

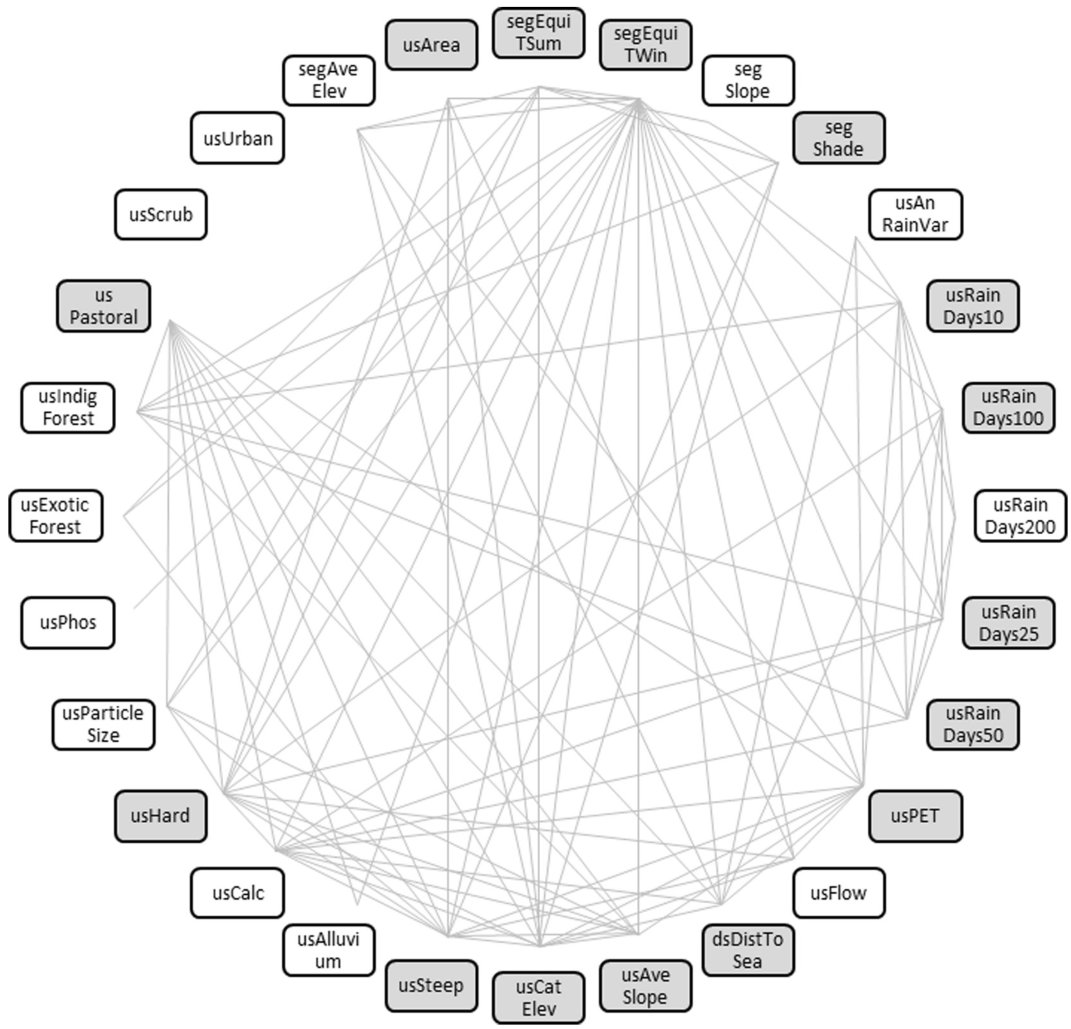

Fig. 6. Schematic output of Pearson's Chi-squared test. Lines indicate significant relationship between two variables (P-value < 0.01%). Variables in grey boxes were excluded from further analysis due to their dependency to other variables.

national river network with Strahler stream order of one to eight were used as the input parameters. The predictions for each reach were scaled such that they summed to one. The worse performance of the models when assessed using the regional cross validation suggests that caution should be applied when applying the models to locations where few calibration data were available.

Spatial patterns of the percentage of sand and finer sediments (Fig. 10a) and their median diameter (Fig. 10b) were mapped across New Zealand for rivers of Strahler order greater than four. There were coherent spatial patterns in the proportion of sediments with fine sediments. Greater proportions of finer sediments were predicted for smaller lowland rivers on the east Coast of the South Island, whereas smaller proportions of finer sediments (and therefore greater proportions of coarser sediments) were predicted for more mountainous rivers across the Southern Alps (which run along the spine of the South Island) and larger mountain-fed rivers as they cross lowland areas. In the North Island, greater proportions of finer sediments were predicted for lowland areas, particularly toward the east and north coast.

Median grain size (D_{50}) for each river reach was calculated by linear interpolation between the high and low values of the size fraction range in which the percentile was found.

Fig. 11 shows the variability of D_{50} for different river types on each stream order, separately. Based on the predicted substrate cover proportions derived from the combined RF and SND models, the median value of D_{50} was 70 mm ranging from 5 to 256 mm for all river first order reaches. The median value of D_{50} for all 4th, 5th, 6th and 7th order streams were 60, 58.5, 52 and 42 mm, respectively. Decrease in median values of D_{50} from first to seventh order streams indicate a general trend for downstream fining of surface D_{50} in rivers across New Zealand. Strong relationship between mean grain size and position of the reach in the river networks were contrary with findings by Snelder et al. (2011b) on river reaches throughout France. Note that Snelder et al. (2011b) directly estimated mean grain size from catchment variables, whereas in this study, we first modelled each size fraction and then mean grain size was calculated based on size fraction estimates. As expected this trend is not followed along river's whose sediment transport is highly affected by lakes at their upstream (i.e. Lakefed river types in Fig. 11). In addition, there are stronger relationships between mean grain size and distance from the headwaters, indicated generally by stream order, in hill type rivers (ranging from 105 mm in 1st stream order to 40 mm in 7th stream order) in comparison with Lowland classes (ranging from 49 mm in 1st stream order to 40 mm in 7th stream order). This is because the number of tributaries contributing to Lowland river reaches is higher than reaches in the hill topographic class and thus as observed by Rice (1998) and Benda et al. (2004) sedimentary inputs from tributaries interrupt the general downstream fining trend.

Therefore, texture analysis of predicted substrate covers demonstrated that rivers with lower stream order which originate from the upper parts of the catchments had the highest proportions of sediment

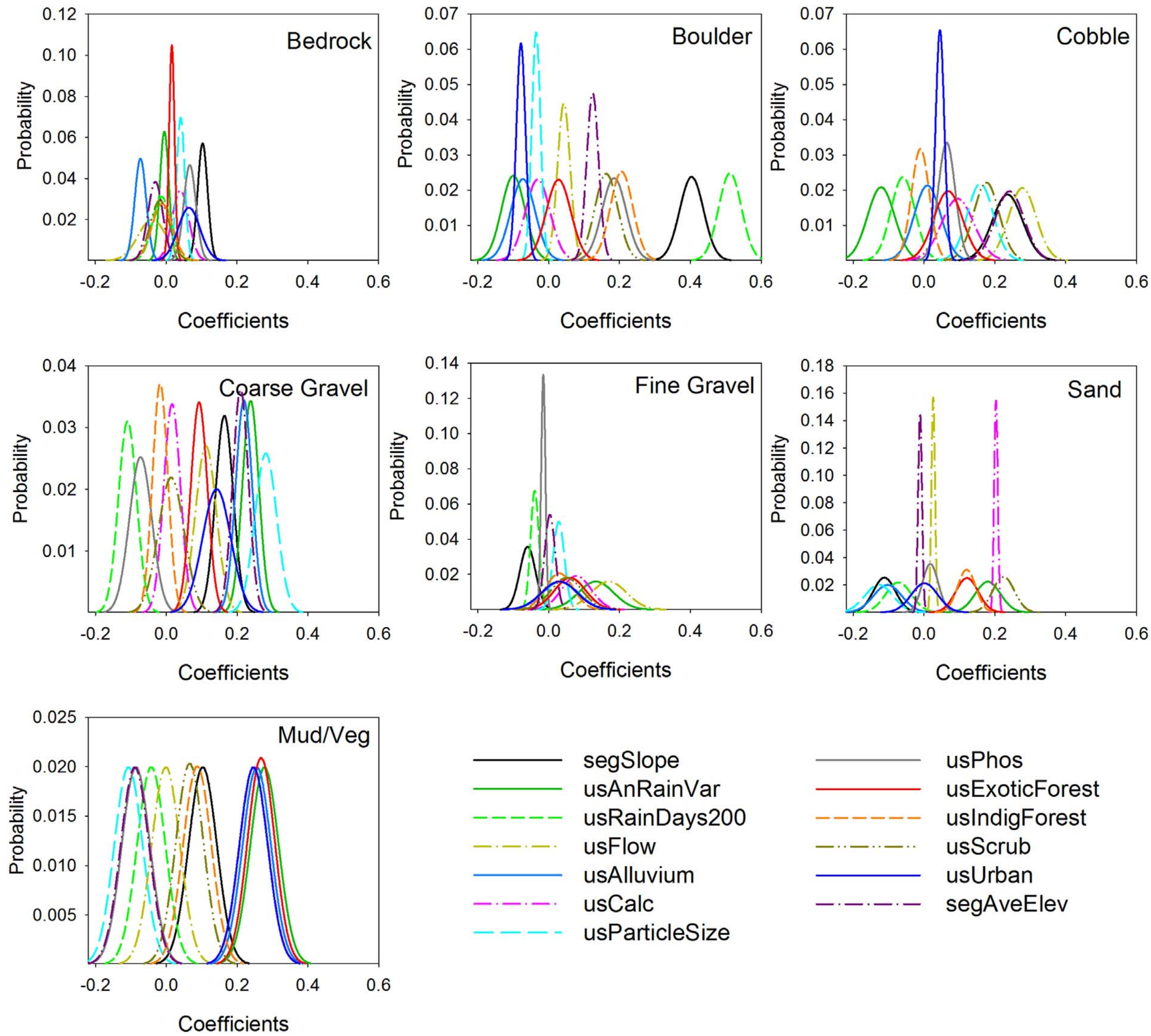

Fig. 7. Normal distribution of coefficients for all 13 predictors and all eight substrates (derived from Eq. (5)).

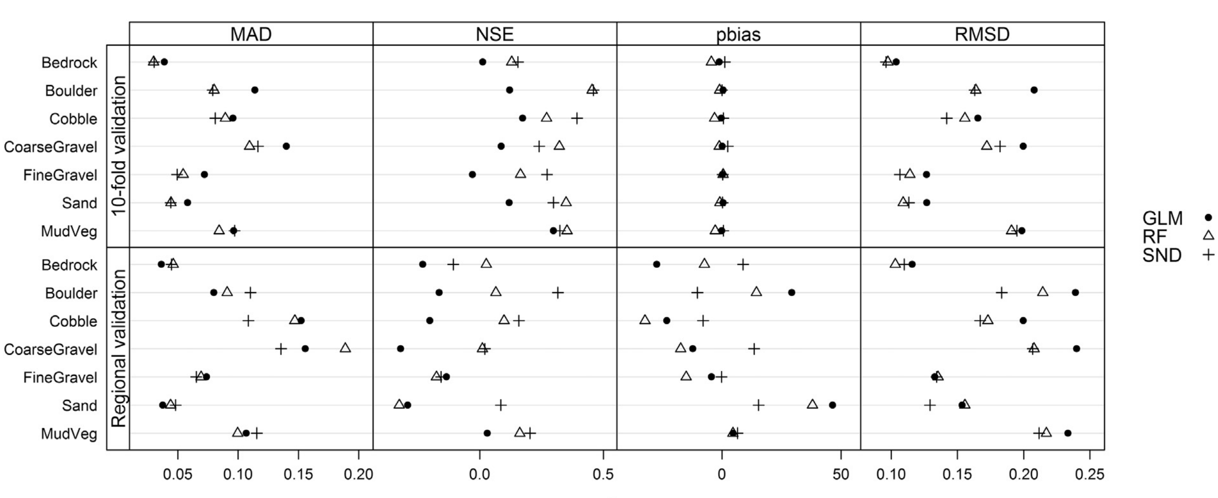

Fig. 8. Efficiency criteria values obtained for 10-fold (top plot) and regional (bottom plot) cross validations of normal distribution (SND), random forest (RF) and generalised linear (GLM) model results. MAD is mean absolute deviation, NSE is Nash-Sutcliffe model efficiency, pbias is percent bias and RMSD is Root Mean Square Deviation.

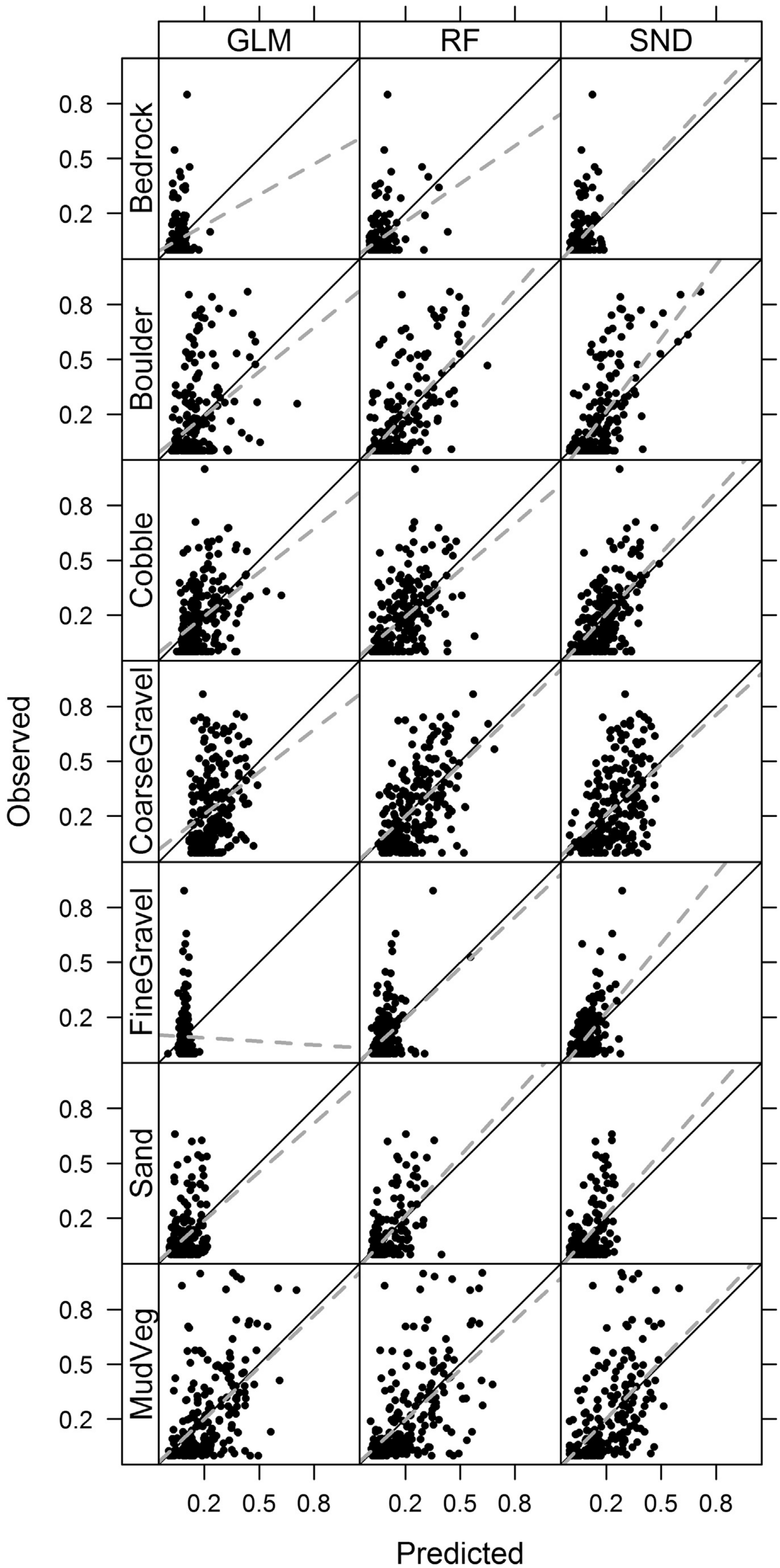

Fig. 9. Observed against K-fold cross validation (CV) predicted values of proportion areal cover for each size fraction of substrate category. Solid line represents 1:1. Grey dashed line is a linear regression.

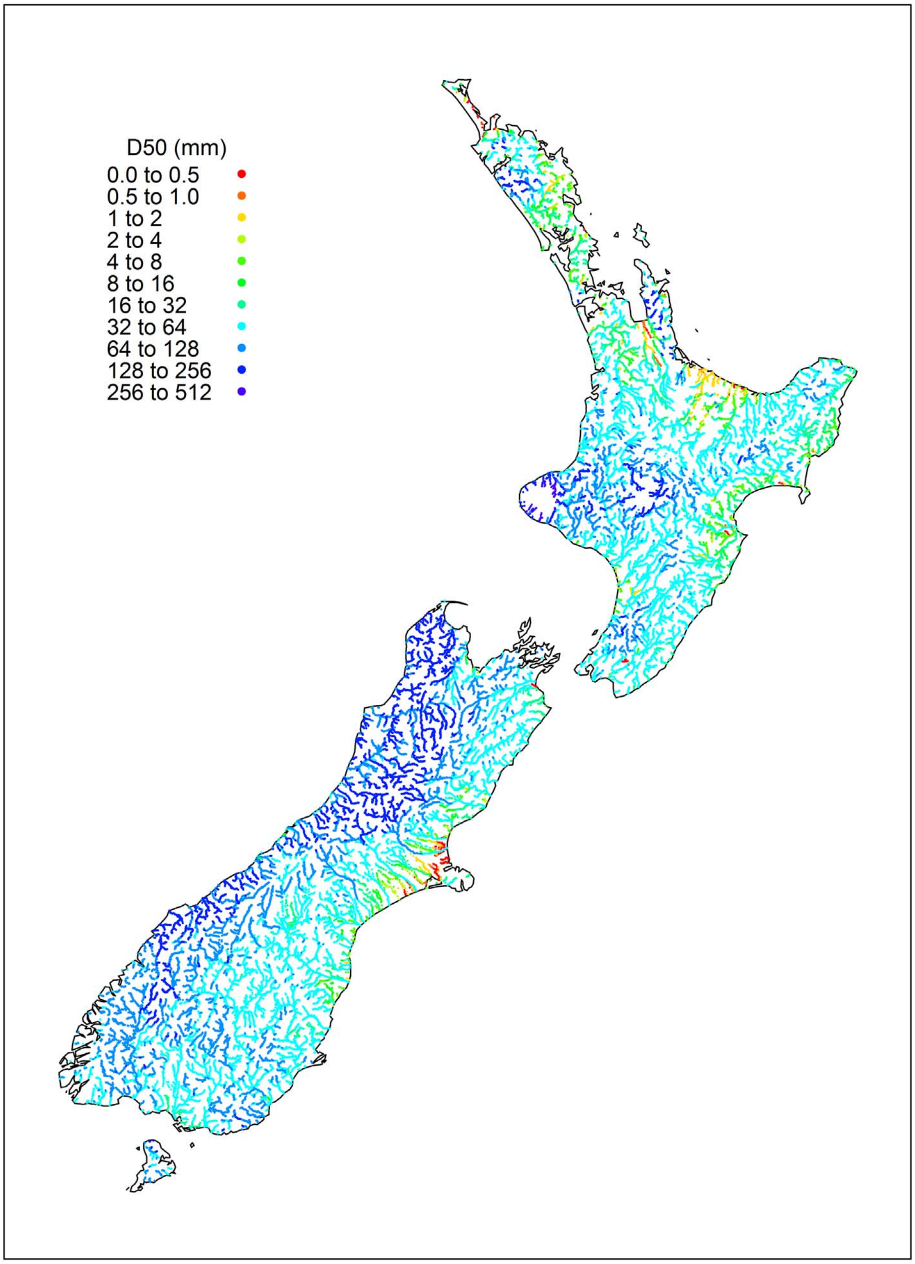

Fig. 10. Map of median diameter (D_{50}) for the entire New Zealand network with Strahler order 4 and above.

in the gravel size fractions and the further downstream (i.e. rivers with high stream orders) the more sand sized fractions were present.

To better demonstrate downstream trends of decreasing particle size on each river reach, longitudinal profiles of the predicted sediment sizes are presented for 26 major New Zealand rivers (Fig. 12). We fitted an exponential curve to the data (Eq. (1)) consistent with textbook geomorphic relationships for downstream fining. The exponential curve provided a better fit to the median diameter (D_{50}) data rather than diameter for which 84% of the sediments are finer (D_{84}) as regressions fitted to D_{50} values were significant for all rivers at $P < 0.0001$ shown in Fig. 12. This indicates that our model predicted substrate predictions show patterns of downstream fining consistent with this well-established relationship.

A very wide range of empirical diminution coefficients have previously been reported (Rice, 1999), with overall dependence of the coefficient on catchment area and thus on the scale of the river system (Hoey and Bluck, 1999). In this study, diminution coefficients for the D_{50} ranged from $1.18 \times 10^{-6}\,\mathrm{km}^{-1}$ for the Waitaki River to $5.08 \times 10^{-5}\,\mathrm{km}^{-1}$ for the Kaituna River (Bay of Plenty region), and predicted downstream patterns were not always smooth exponential curves. This indicates simple representations of downstream fining (e.g. Eq. (1)) may not be sufficient to represent national scale patterns in substrate patterns because between-catchment variations in diminution coefficients result from landscape-scales processes driven by differences in slope, sediment supply and stream power. Furthermore, changes in downstream conditions, such as changes in sediment supply or hydrology which might occur at lakes or at major tributaries join together, also interfere with downstream fining patterns. For example, many of New Zealand's larger rivers flow through lakes (e.g. the Waitaki River) and transition from mountain to lowland landscapes occur (e.g. the Rakaia and Waimakariri) as they flow to the sea.

Three important issues should be noted when applying the predicted values nationally. First, the predictive models were derived from visual inspection of proportions of river bed surface covered by different categories carried out as part of fish habitat studies. Surface layer grain-size distributions will differ to sub-surface distributions, although conversions can be applied to allow comparison (Kellerhals and Bray, 1971). Second, there is a variety of methods that can be used to

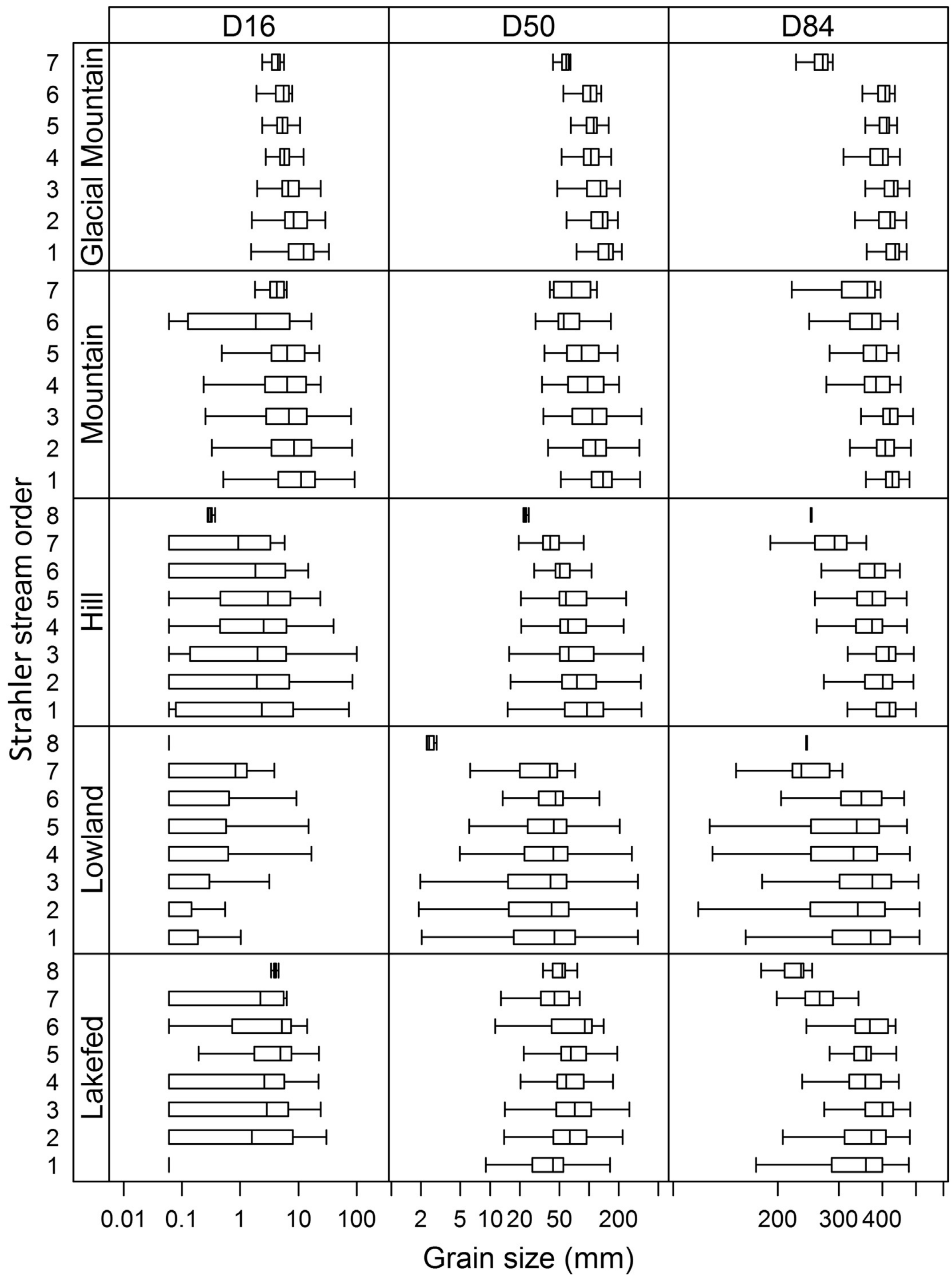

Fig. 11. Box and whisker plots of D_{50} values for different type of rivers and stream orders. Middle horizontal lines show medians. Boxes indicate quantiles. Whiskers at left and right of the box indicates the 97.5th and 2.5th percentiles. See Snelder et al. (2005) for definitions of river types.

quantify sediment size distributions (Bunte and Abt, 2001), and care is required when comparing predicted values with those collected using contrasting sampling methods. Third, whilst the models were fitted to data collected from sites across New Zealand, and from a variety of river types, these data did not include deep (non-wadable) rivers. Therefore, any applications and interpretation of the predictions should note that the predictions represent bed surface observations rather than sub-surface samples and that predictions made for large, deep rivers may over-estimate proportions of larger size-fractions.

4. Conclusion

Three models were tested for predicting substrate cover proportions across New Zealand. The models deliberately represent varying approaches to selecting predictors, interactions between predictors, and ability to include uncertainty. Final predictions of each substrate category were based on the best performing model for that category (as assessed using 10-fold cross-validation). 10-fold cross validation showed summed normal distribution models provided the best estimates for four, out of seven, categories of substrate cover proportions. Random forest models provided better performing estimations for three

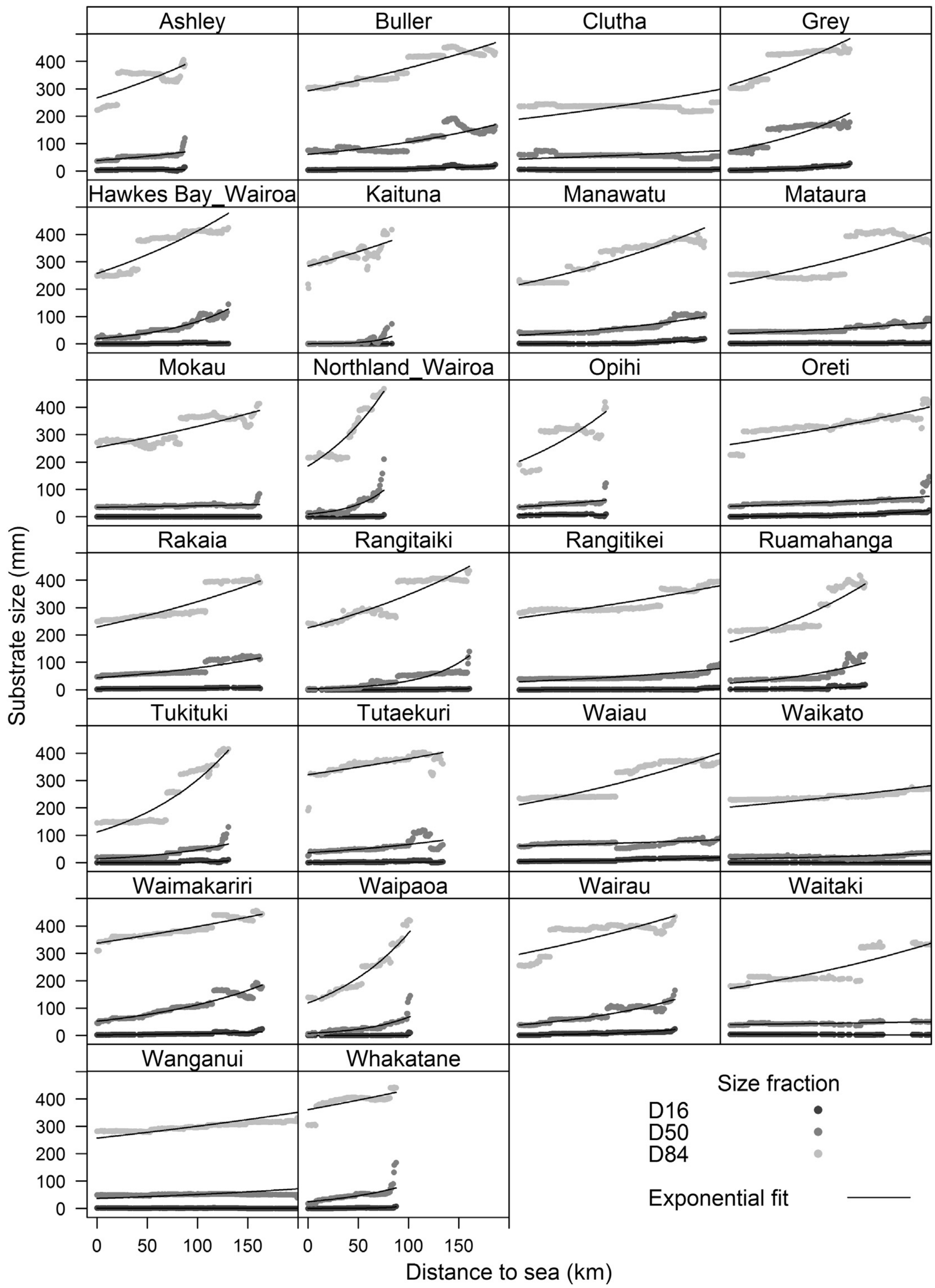

Fig. 12. Long-profiles of predicted sediment sizes for rivers mapped in Fig. 1. D_{16}, D_{50} and D_{84} are grain sizes for which 16%, 50% and 84% of the predicted sediments are finer, respectively.

of the substrate categories. Generalised linear models fitted using a deliberately simple set of predictors did not out-perform the other two methods, but were still able to provide explanatory power on patterns of substrate cover proportions. This indicates that patterns in substrate are related to simple representations of hydrological patterns and landscape setting, but that predictor interactions and flexible relationships were required to allow improvements in predictive power. Dividing the dataset geographically into North Island and South Island sites to perform regional cross validation was used to test the generality of model predictions outside the area they were calibrated in. These results were worse than the 10-fold cross-validation suggesting that model results should be treated with caution in regions where few calibration data were available. In particular, pbias results were much worse in the regional validation highlighting that significant bias is likely in areas where no training data are available. The summed normal distribution model was less biased and showed better predictive power than the other models in the regional calibration.

A combination of best performing methods (as assessed in the 10-

fold validation) was used to derive the most accurate models for all substrate categories across the New Zealand river network. Results indicate general increase of the fine fractions and decrease of coarse fractions in higher stream orders. This is consistent with textbook geomorphological relationships for downstream fining. The national predictions will be useful for a wide range of national scale applications such as: assessing the representativeness of monitoring sites, species distribution monitoring, and assessing habitat quality.

One disadvantage of selecting the best performing models is that they were not designed to explore the impact of potential future conditions resulting from: climate change (e.g. changes in rainfall), land cover management (e.g. decreasing conservation areas) or hydrological alteration (e.g. increased abstraction). Despite this methodology, some relationships between predictors and the various substrate categories aligned with a theoretical physically basis for landscape-scale controls on sediment deposition. For example, landcover influenced finer sizes only, whereas the importance of slope decreased as grain size decreased.

Acknowledgements

This work was funded by NIWA's Sustainable Water Allocation Programme [grant number FWWA1605]. Many thanks are due to the various data collectors.

References

Armstrong, J.D., Kemp, P.S., Kennedy, G.J.A., Ladle, M., Milner, N.J., 2003. Habitat requirements of Atlantic salmon and brown trout in rivers and streams. Fish. Res. 62 (2), 143–170.
Attal, M., Mudd, S.M., Hurst, M.D., Weinman, B., Yoo, K., Naylor, M., 2015. Impact of change in erosion rate and landscape steepness on hillslope and fluvial sediments grain size in the Feather River basin (Sierra Nevada, California). Earth Surf. Dyn. 3 (1), 201–222.
Benda, L., Poff, N.L., Miller, D., Dunne, T., Reeves, G., Pess, G., Pollock, M., 2004. The network dynamics hypothesis: how channel networks structure riverine habitats. Bioscience 54, 413–427.
Booker, D.J., 2010. Predicting wetted width in any river at any discharge. Earth Surf. Process. Landf. 35 (7), 828–841.
Booker, D.J., 2016. Generalized models of riverine fish hydraulic habitat. J. Ecohydraul. 1 (1-2), 31–49.
Booker, D.J., Woods, R.A., 2014. Comparing and combining physically-based and empirically-based approaches for estimating the hydrology of ungauged catchments. J. Hydrol. 508, 227–239.
Booker, D.J., Snelder, T.H., Greenwood, M.J., Crow, S.K., 2015. Relationships between invertebrate communities and both hydrological regime and other environmental factors across New Zealand's rivers. Ecohydrology 8 (1), 13–32.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Buffington, J.M., Montgomery, D.R., 1997. A systematic analysis of eight decades of incipient motion studies, with special reference to gravel-bedded rivers. Water Resour. Res. 33 (8), 1993–2029.
Buffington, J.M., Montgomery, D.R., Greenberg, H.M., 2004. Basin-scale availability of salmonid spawning gravel as influenced by channel type and hydraulic roughness in mountain catchments. Can. J. Fish. Aquat. Sci. 61 (2085–2096).
Bunte, K., Abt, S.R., 2001. Sampling Surface and Subsurface Particle-size Distributions in Wadable Gravel-and Cobble-bed Streams for Analyses in Sediment Transport, Hydraulics, and Streambed Monitoring. General Technical Report RMRS-GTR-74. USDA, Forest Service, Rocky Mountain Research Station, Fort Collins, CO, 2001.
Carbonneau, P., Fonstad, M.A., Marcus, W.A., Dugdale, S.J., 2012. Making riverscapes real. Geomorphology 137 (1), 74–86.
Church, M., Kellerhals, R., 1978. On the statistics of grain size variation along a gravel river. Can. J. Earth Sci. 15 (7), 1151–1160.
Costigan, K.H., Daniels, M.D., Perkin, J.S., Gido, K.B., 2014. Longitudinal variability in hydraulic geometry and substrate characteristics of a Great Plains sand-bed river. Geomorphology 210, 48–58.
Crow, S.K., Booker, D.J., Snelder, T.H., 2013. Contrasting influence of flow regime on freshwater fishes displaying diadromous and nondiadromous life histories. Ecol. Freshw. Fish 22 (1), 82–94.
de Almeida, G.A.M., Rodríguez, J.F., 2011. Understanding pool-riffle dynamics through continuous morphological simulations. Water Resour. Res. 47 (1), W01502.
Ferguson, R.I., 2003. Emergence of abrupt gravel-to-sand transitions along rivers through sorting processes. Geology 312, 159–162.
Frings, R.M., 2008. Downstream fining in large sand-bed rivers. Earth Sci. Rev. 87 (1–2), 39–60.
Genuer, R., Poggi, J.-M., Tuleau-Malot, C., 2010. Variable selection using random forests. Pattern Recogn. Lett. 31 (14), 2225–2236.
Genuer, R., Poggi, J.M., Tuleau-Malot, C., 2015. VSURF: an R package for variable selection using random forests. R J. 7 (2), 19–33.
Gorman, A.M., Whiting, P.J., Neeson, T.M., Koonce, J.F., 2011. Channel substrate prediction from GIS for habitat estimation in Lake Erie tributaries. J. Great Lakes Res. 37 (4), 725–731.
Groll, M., Thomas, A., Jungermann, L., Schäfer, K., 2016. Typology of riverbed structures and habitats (TRiSHa) – a new method for a high resolution characterization of the spatial distribution and temporal dynamic of riverbed substrates and microhabitats. Ecol. Indic. 61 (Part 2), 219–233.
Grömping, U., 2009. Variable importance assessment in regression: linear regression versus random forest. Am. Stat. 63 (4), 308–319.
Gupta, H.V., Kling, H., Yilmaz, K.K., Martinez, G.F., 2009. Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling. J. Hydrol. 377 (1–2), 80–91.
Haddadchi, A., Omid, M.H., Sdehghani, A.A., 2013. Total load transport in gravel bed and sand bed rivers case study: Chelichay watershed. Int. J. Sediment Res. 28 (1), 46–57.
Haddadchi, A., Olley, J., Laceby, P., 2014. Accuracy of mixing models in predicting sediment source contributions. Sci. Total Environ. 497–498, 139–152.
Haddadchi, A., Olley, J., Pietsch, T., 2015. Quantifying sources of suspended sediment in three size fractions. J. Soils Sediments 15 (10), 2086–2100.
Harding, J., Clapcott, J., Quinn, J., Hayes, J., Joy, M., Storey, R., Greig, H., Hay, J., James, T., Beech, M., Ozane, R., Meredith, A., Boothroyd, I., 2009. Stream Habitat Assessment Protocols for Wadeable Rivers and Streams of New Zealand. (Christchurch, New Zealand).
Hedger, R.D., Dodson, J.J., Bourque, J.F., Bergeron, N.E., Carbonneau, P.E., 2006. Improving models of juvenile Atlantic salmon habitat use through high resolution remote sensing. Ecol. Model. 197 (3–4), 505–511.
Herbst, D.B., Suk, T., 2005. Quality Assurance Project Plan - Aquatic Invertebrate Bioassessment Monitoring in the Eastern Sierra Nevada - Development of Biological Criteria for Stream Water Quality and Evaluations of Environmental Impacts and Restoration Projects in the Lahontan Region. Lahontan RegionalWater Quality Control Board Final Report. pp. 15.
Hoey, T.B., Bluck, B.J., 1999. Identifying the controls over downstream fining of river gravels. J. Sediment. Res. 69 (1), 40–50.
Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. Int. J. Forecast. 22 (4), 679–688.
Jellyman, P.G., Booker, D.J., McIntosh, A.R., 2013. Quantifying the direct and indirect effects of flow-related disturbance on stream fish assemblages. Freshw. Biol. 58 (12), 2614–2631.
Jowett, I.G., Hayes, J.W., Duncan, M.J., 2008. A Guide to Instream Habitat Survey Methods and Analysis. NIWA Science and Technology Series. 54. pp. 121.
Kellerhals, R., Bray, D.I., 1971. Sampling procedures for coarse fluvial sediments. J. Hydraul. Div. 97 (8), 1165–1180.
Kondolf, G.M., Wolman, M.G., 1993. The sizes of salmonid spawning gravels. Water Resour. Res. 29 (7), 2275–2285.
Leathwick, J.R., Morgan, F., Wilson, G., Rutledge, D., McLeod, M., K, J., 2002. Land Environments of New Zealand; A Technical Guide. Landcare Research Technical Report. pp. 237.
Leathwick, J.R., Elith, J., Chadderton, W.L.W., Rowe, D., Hastie, T., 2008. Dispersal, disturbance, and the contrasting biogeographies of New Zealand's diadromous and non-diadromous fish species. J. Biogeogr. 35, 1481–1497.
Leathwick, J.R., Snelder, T., Chadderton, W.L., Elith, J., Julian, K., Ferrier, S., 2011. Use of generalised dissimilarity modelling to improve the biological discrimination of river and stream classifications. Freshw. Biol. 56 (1), 21–38.
Maddock, I., 1999. The importance of physical habitat assessment for evaluating river health. Freshw. Biol. 41 (2), 373–391.
Marshall, J.A., Sklar, L.S., 2012. Mining soil databases for landscape-scale patterns in the abundance and size distribution of hillslope rock fragments. Earth Surf. Process. Landf. 37 (3), 287–300.
Menting, F., Langston, A.L., Temme, A.J.A.M., 2015. Downstream fining, selective transport, and hillslope influence on channel bed sediment in mountain streams, Colorado Front Range, USA. Geomorphology 239, 91–105.
Miller, K.L., Szabó, T., Jerolmack, D.J., Domokos, G., 2014. Quantifying the significance of abrasion and selective transport for downstream fluvial grain size evolution. J. Geophys. Res. Earth Surf. 119 (11), 2014JF003156.
Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. Am. Soc. Agric. Biol. Eng. 50 (3), 885–900.
Morris, P.H., Williams, D.J., 1999. A worldwide correlation for exponential bed particle size variation in subaerial aqueous flows. Earth Surf. Process. Landf. 24 (9), 835–847.
Muleta, M.K., 2012. Model performance sensitivity to objective function during automated calibrations. J. Hydrol. Eng. 17 (6), 756–767.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I — a discussion of principles. J. Hydrol. 10 (3), 282–290.
Nelder, J.A., Wedderburn, R.W.M., 1972. Generalized linear models. J. R. Stat. Soc. Ser. A 135 (3), 370–384.
Olley, J., Brooks, A., Spencer, J., Pietsch, T., Borombovits, D., 2013. Subsoil erosion dominates the supply of fine sediment to rivers draining into Princess Charlotte Bay, Australia. J. Environ. Radioact. 124 (0), 121–129.
Oracle, 2015. Crystal Ball 11.1.2.4. 32 Bit.
Parker, G., Toro-Escobar, C.M., 2002. Equal mobility of gravel in streams: the remains of the day. Water Resour. Res. 38 (11), 1264.
Piñeiro, G., Perelman, S., Guerschman, J.P., Paruelo, J.M., 2008. How to evaluate models: observed vs. predicted or predicted vs. observed? Ecol. Model. 216 (3–4), 316–322.
Powell, D.M., 1998. Patterns and processes of sediment sorting in gravel-bed rivers. Prog. Phys. Geogr. 22 (1), 1–32.
Reiser, D.W., 1998. Sediment in gravel bed rivers: ecological and biological

considerations. In: Klingeman, P.C., Beschta, R.L., Komar, P.D., Bradley, J.B. (Eds.), Gravel-bed Rivers in the Environment. Water Resources Publications, Colorado, pp. 199–225.
Rempel, L.L., Richardson, J.S., Healey, M.C., 2000. Macroinvertebrate community structure along gradients of hydraulic and sedimentary conditions in a large gravel-bed river. Freshw. Biol. 45 (1), 57–73.
Rice, S., 1998. Which tributaries disrupt downstream fining along gravel-bed rivers? Geomorphology 22 (1), 39–56.
Rice, S., 1999. The nature and controls on downstream fining within sedimentary links. J. Sediment. Res. 69 (1), 32–39.
Riebe, C.S., Sklar, L.S., Lukens, C.E., Shuster, D.L., 2015. Climate and topography control the size and flux of sediment produced on steep mountain slopes. Proc. Natl. Acad. Sci. 112 (51), 15574–15579.
Rinaldi, M., Surian, N., Comiti, F., Bussettini, M., 2013. A method for the assessment and analysis of the hydromorphological condition of Italian streams: the Morphological Quality Index (MQI). Geomorphology 180–181, 96–108.
Sklar, L.S., Riebe, C.S., Marshall, J.A., Genetti, J., Leclere, S., Lukens, C.L., Merces, V., 2017. The problem of predicting the size distribution of sediment supplied by hillslopes to rivers. Geomorphology 277, 31–49.
Snelder, T.H., Biggs, B.J.F., 2002. Multiscale river environment classification for water resources management. JAWRA J. Am. Water Resour. Assoc. 38 (5), 1225–1239.
Snelder, T.H., Biggs, B.J.F., Woods, R.A., 2005. Improved eco-hydrological classification of rivers. River Res. Appl. 21 (6), 609–628.
Snelder, T.H., Booker, D.J., Lamouroux, N., 2011a. A method to assess and define environmental flow rules for large jurisdictional regions. J. Am. Water Resour. Assoc. 47 (4), 828–840.
Snelder, T.H., Lamouroux, N., Pella, H., 2011b. Empirical modelling of large scale patterns in river bed surface grain size. Geomorphology 127 (3–4), 189–197.
Snyder, N.P., Nesheim, A.O., Wilkins, B.C., Edmonds, D.A., 2013. Predicting grain size in gravel-bedded rivers using digital elevation models: application to three Maine watersheds. Geol. Soc. Am. Bull. 125, 148–163.
Ten Brinke, W.B.M., Schulze, F.H., van Der Veer, P., 2004. Sand exchange between groyne-field beaches and the navigation channel of the Dutch Rhine: the impact of navigation versus river flow. River Res. Appl. 20 (8), 899–928.
Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, fourth ed. Springer, New York.
Wilcock, P.R., Crowe, J.C., 2003. Surface-based transport model for mixed-size sediment. J. Hydraul. Eng. 129 (2), 120–128.
Wilkins, B.C., Snyder, N.P., 2011. Geomorphic comparison of two Atlantic coastal rivers: toward an understanding of physical controls on Atlantic salmon habitat. River Res. Appl. 27 (2), 135–156.
Wright, J.F., Furse, M.T., Moss, D., 1998. River classification using invertebrates: RIVPACS applications. Aquat. Conserv. Mar. Freshwat. Ecosyst. 8 (4), 617–631.