Regional Flood Estimation Tool for New Zealand Part 2

Final Report

Prepared for Envirolink

June 2018

Prepared by:
Roddy Henderson
Daniel Collins
Martin Doyle, Tasman District Council
Jeff Watson, Horizons Regional Council

For any information regarding this report please contact:

Roddy Henderson
Hydrological Processes
+64-3-343 7873
roddy.henderson@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
PO Box 8602
Riccarton
Christchurch 8011

Phone +64 3 348 8987

NIWA CLIENT REPORT No:	2018177CH
Report date:	June 2018
NIWA Project:	CARW1805

Quality Assurance Statement		
	Reviewed by:	Charles Pearson
	Formatting checked by:	Fenella Falconer
	Approved for release by:	Sam Dean

Contents

Tables

Figures

Executive summary

We have derived a new model of flood magnitude for New Zealand catchments and a re-assessment of the uncertainty inherent in the existing method that this work is intended to replace. This work was prepared for the regional sector and forms an extension to an original report (finalised in August 2016) funded by a MBIE Envirolink Tools grant.

Flood estimation and its companion discipline, extreme rainfall intensity estimation, are critical aspects of the design of a large amount of the built infrastructure of New Zealand. The previous method for flood estimation, dating from 1989, needed updating because more extreme events have been observed in the interim, and because of the probable effects of climate change, which will increase into the future. The previous method was derived using subjective expert opinion to build the empirical model, and in this work, we specified a more objective procedure, to allow more frequent and convenient updating in the future.

The new dataset has twice as many sites and three times the annual maxima than the previous study. Nearly 58% of sites are operated by regional councils, 38% by NIWA, and the remaining 4% by other organisations. Preliminary analysis suggested no spatially coherent temporal trends in the annual series of flood maxima. The new dataset is systematically organised with inclusion of both monthly and annual maxima for each series, annotation of the years potentially affected by gaps and expert assessment of the true impact of gaps, and inclusion of early historic annual maxima.

Workshops held in late 2015 for regional council stakeholders and for a wider audience provided useful feedback about aspects of the new model. Changes made following these are incorporated in the current model, including the division of the dataset by island.

Over the past two years we have explored ways to better estimate mean annual flood (MAF). This has involved co-learning approaches between researchers and regional council practitioners. As a result, we have chosen regression models that seek to optimise information gain without incorporating too many variables, thus remained conceptually tractable and transparent. We have investigated alternative regional approaches, and developed objective contouring methods to account for any spatial organisation of residuals. We have also adopted an unbiased error estimator, being the ratio of logs of data-based estimates of MAF and modelled MAF.

Ordinary least squares (OLS) regressions in log space have been performed on each Island individually, using a three-variable equation (area, annual precipitation and hydrogeology) in the North Island, and a two-variable equation (area and annual precipitation) in the South Island. The residuals of these equations have been contoured and a leave-one-out cross validation performed. The result of this process is an all-New Zealand record-length-weighted factorial error of 1.82 for MAF, or a relative error of ±61%. To compare with the previous method, the worst 5% of sites are removed and the factorial error reduces to 1.62. As a relative error this is ±49%, which is as good as or better than the assessed error of the previous method for 95% of all New Zealand, at ±49% to ±70%. We propose ±50% as the standard error of estimate for mean annual flood.

The regional growth curve model of the previous study was found to be still applicable, and we propose ±20% as the standard error of estimate for q_{100} (Q_{100}/MAF).

The MAF model is combined with the regional growth curve model to provide return period estimates from 5 to 1000 years, with standard error estimates, across all stream reaches. The results are displayed on a web-based map application, with the option to download flood statistics for

selected rivers and streams. At-site annual flood series and calculated flood statistics are also displayed and may be downloaded.

The method provides a good balance of technical depth, repeatability, physical realism, and transparency, all aided by the joint application of statistical modelling and co-learning among researchers and stakeholders.

1 Introduction

In August 2016 a report by Henderson and Collins (2016) describing the work completed on the regional flood frequency study funded by a MBIE Envirolink Tools grant (C01X1308) was submitted to regional council representatives. Subsequent reviews and comments from regional council staff informed further work over 2016/17 and 2017/18. This report describes that further work, and includes notes of correspondence between NIWA and regional council staff exchanged over that time that guided these developments and led to the revised current form of the flood estimation tool.

The report is mainly chronological and describes the developments arising from a number of meetings both face to face and virtual, between the NIWA authors of the first report and a number of regional council staff, including Martin Doyle (Tasman District Council and primary contact for the work), Jeff Watson (Horizons Regional Council, LAEMG[1]), Gary Clode (Hawke's Bay Regional Council, River Managers Special Interest Group), and Craig Goodier and David Carruth (Hawke's Bay Regional Council engineers).

[1] Local Authority Environmental Monitoring Group.

2 Meetings and correspondence

2.1 September 2016 meeting and subsequent developments

The outcome of this face to face meeting was a series of emails concerning further work in late 2016, a draft MBIE EnviroLink completion document at April 2017, and an unsuccessful attempt to devise an automated regionalisation procedure, reported in the NIWA MBIE SSIF completion report of June 2017.

Various steps of the above took a number of months and at the end of June the councils were not happy to sign off the EnviroLink Tools project.

2.2 Virtual meeting of 15 September 2017

The outcome of this meeting was a plan of work for the 2017/18 year:

- Produce a number of different models ready for contouring. (Contouring the residuals from a model based on area of catchment only would be the equivalent of the previous McKerchar and Pearson approach).
- Joint decision made with council staff on the best model to which to apply the contouring algorithm.
- Apply algorithm to the model and generate uncertainty estimates.
- Analysis of uncertainty estimates and reporting.
- Meet with council representatives to present the contouring and demonstrate the web portal tool.

2.3 Virtual meeting of 9 November 2017

At this meeting we discussed new data that had become available, including a revised rain map, and the hydrologically relevant variables in the recently released QMAP from GNS. We also agreed on work to assess alternatives to regionalisation.

2.4 Workshop of 14 December 2017

At this workshop we presented work to date, including that mentioned above, and exchanged ideas about the model choice, and removal of apparent bias from the error analysis.

3 Detailed reanalysis

Following discussions between NIWA and council staff several revised approaches were advanced. These included the use of an enhanced national rainfall map, the inclusion of hydro-geological parameters from the recently released QMAP by GNS, the step-wise consideration of regression fits to select the optimal combination, and contouring of the residuals to objectively deal with the unexplained portion of each model. Optimal is defined as maximum explanation with minimal complexity.

3.1 Rain surface

The new rain surface adopted is an enhanced VCSN[2], by the addition of average rain estimates (mm/year) for raingauges supplied by Tasman District Council. These fill gaps in coverage especially in the mountainous area of northwest Nelson (Aorere headwaters etc.).

There is clearly scope for this approach to be repeated in a later version when other regional council rain networks have been incorporated into the VCSN.

3.2 QMAP parameters

The recently released QMAP (quarter million geology map) by GNS has three parameters that relate to hydrogeology, which warranted testing as potential regression variables for prediction of mean annual flood. These are:

- Depth to basement; processed as an area weighted mean value on the River Environment Classification version 1 (RECv1) catchments.
- Porosity; processed as an area weighted mean value on RECv1 catchments.
- Hydraulic Conductivity; processed as an area weighted harmonic mean on RECv1 catchments.

In the event, none of these were selected by the regression procedure as better than any of the previous set.

3.3 Step-wise regression error approach

An *a priori* order of selection of variables for regression was arranges as follows:

1. Just area alone.
2. Area and one rainfall parameter.
3. Area, annual precipitation, and one storm parameter.
4. Area, one rainfall parameter, and one catchment parameter.
5. Area, annual precipitation, one storm parameter, and one catchment parameter.
6. Area, one rainfall parameter, and two catchment parameters.
7. Area, annual precipitation, one storm parameter, and two catchment parameters.

[2] Virtual Climate Station Network, a 5x5 km coverage of New Zealand interpolated using ANUSplin at daily time steps.

8. Area, one rainfall parameter, and three catchment parameters.

9. Area, annual precipitation, one storm parameter, and three catchment parameters.

Area is the major variable accounting for flood magnitude. Rainfall is a consistent secondary explanatory variable, and after that come a variety of landscape variables that relate to hydrological response. The available variables from which we selected precipitation, storm and catchment variables are listed in Appendix A.

This variable list selected as above in nine different combinations resulted in 484,268 different models for each island. Within each grouping listed above, the best model was selected based on its Root Mean Squared Relative Weighted Error (RMSRWE) as defined in Henderson and Collins (2016):

$$RMSWRE = SQRT[SUM_i\ \{RecLen_i\ .\ [(Q_{mod}-Q_{obs})/Q_{obs}]^2\}/SUM_j(RecLen_j)] \quad (1)$$

where Q_{mod} is the modelled mean annual flood (m^3/s), Q_{obs} is the mean annual flood calculated from the observational record, and *RecLen* is the record length of each site used.

The national models, as presented, combine the same parameter search groups for the two islands, although this does not necessarily mean the same parameters for each island.

The errors for the best of each of the nine models are plotted in Figure 3-1. For the South Island, a significant improvement in model performance is achieved with the addition of a rain parameter (mean annual rain). Further significant improvement is observed by model 6 (addition of two catchment parameters) and there is little improvement after that. For the North Island, significant improvement is achieved by model 4 (rain and one catchment parameter), further improvement by model 6, and little after that.

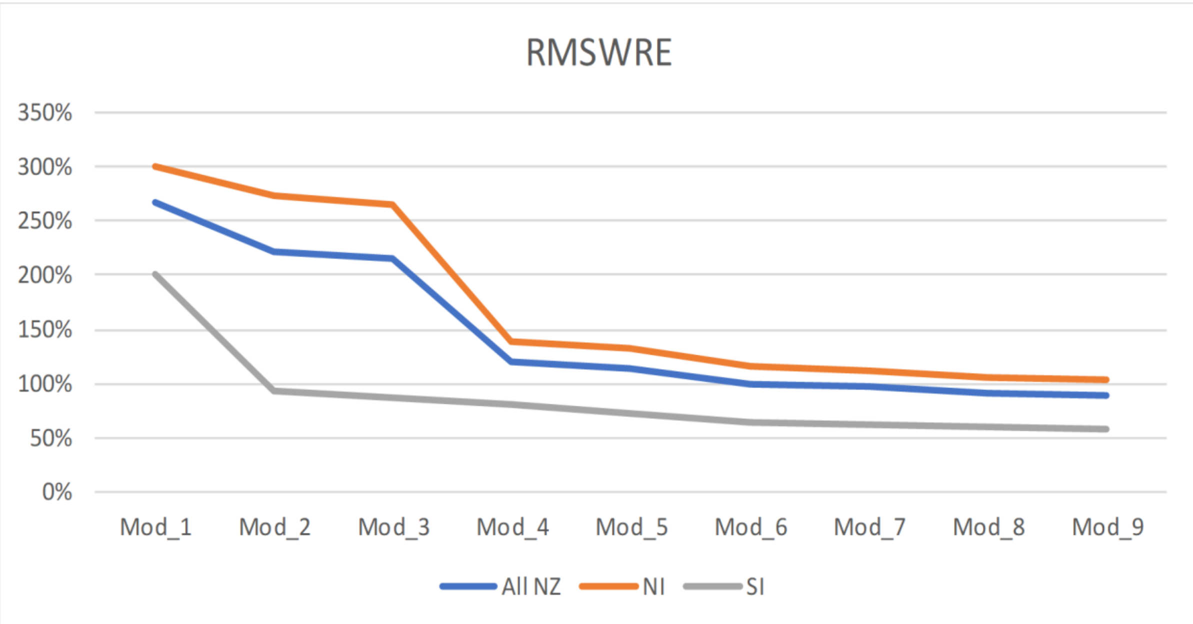

Figure 3-1: Error of best model in each of nine model styles. Error is calculated according to the RMSRWE formula in Henderson and Collins (2016).

The following figures show the spatial error distribution of models 1, 2, 4 and 6. Catchments are coloured over their boundaries, and small catchments are over-plotted on larger for clarity.

Discussion of these maps between NIWA and council staff at the virtual meeting of 9 November 2017 led to several changes to the approach, and these were reported at the face-to-face meeting of 14 December 2017.

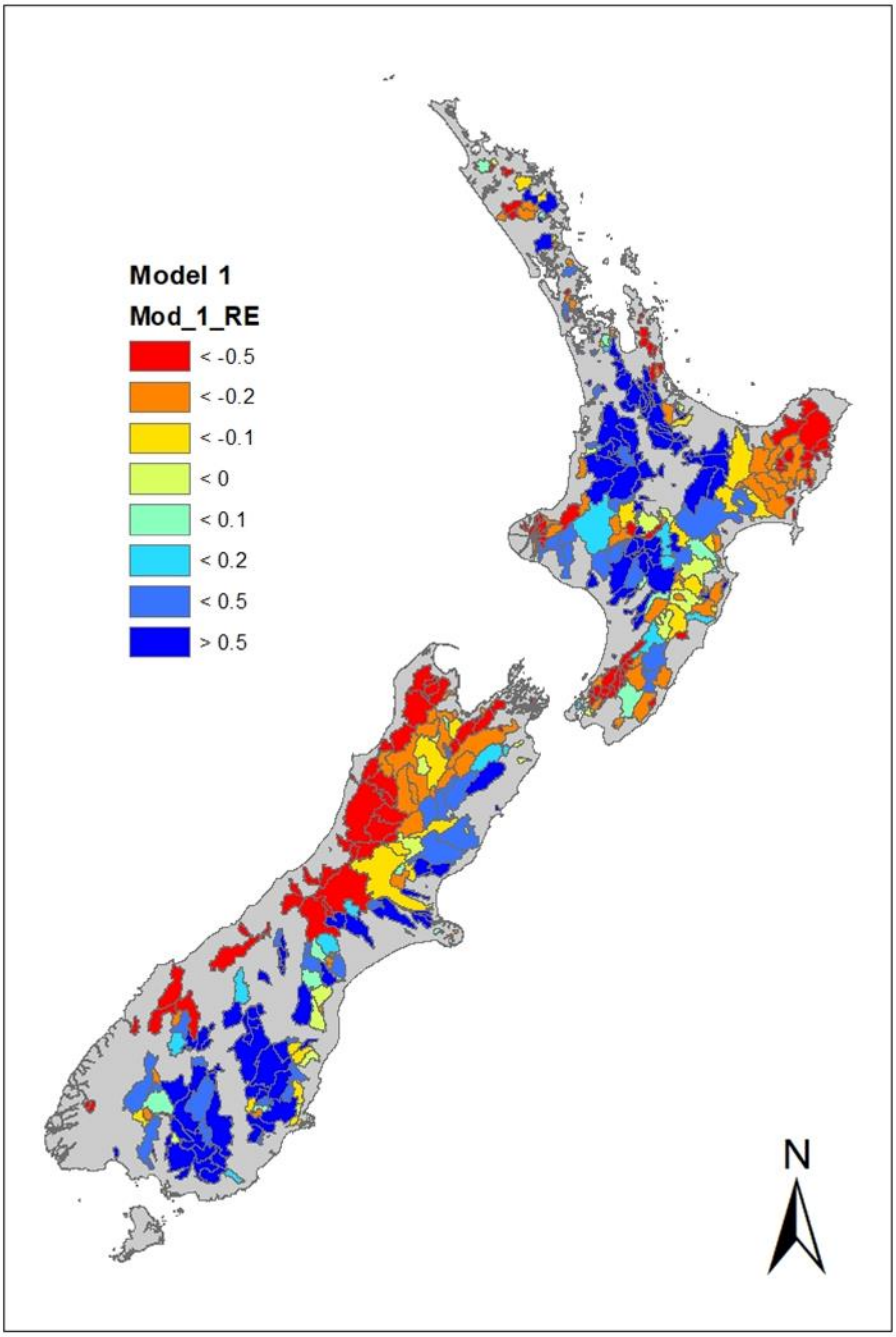

Figure 3-2: RMSWRE for model 1 (area only). Note large contiguous areas of consistent bias in both islands.

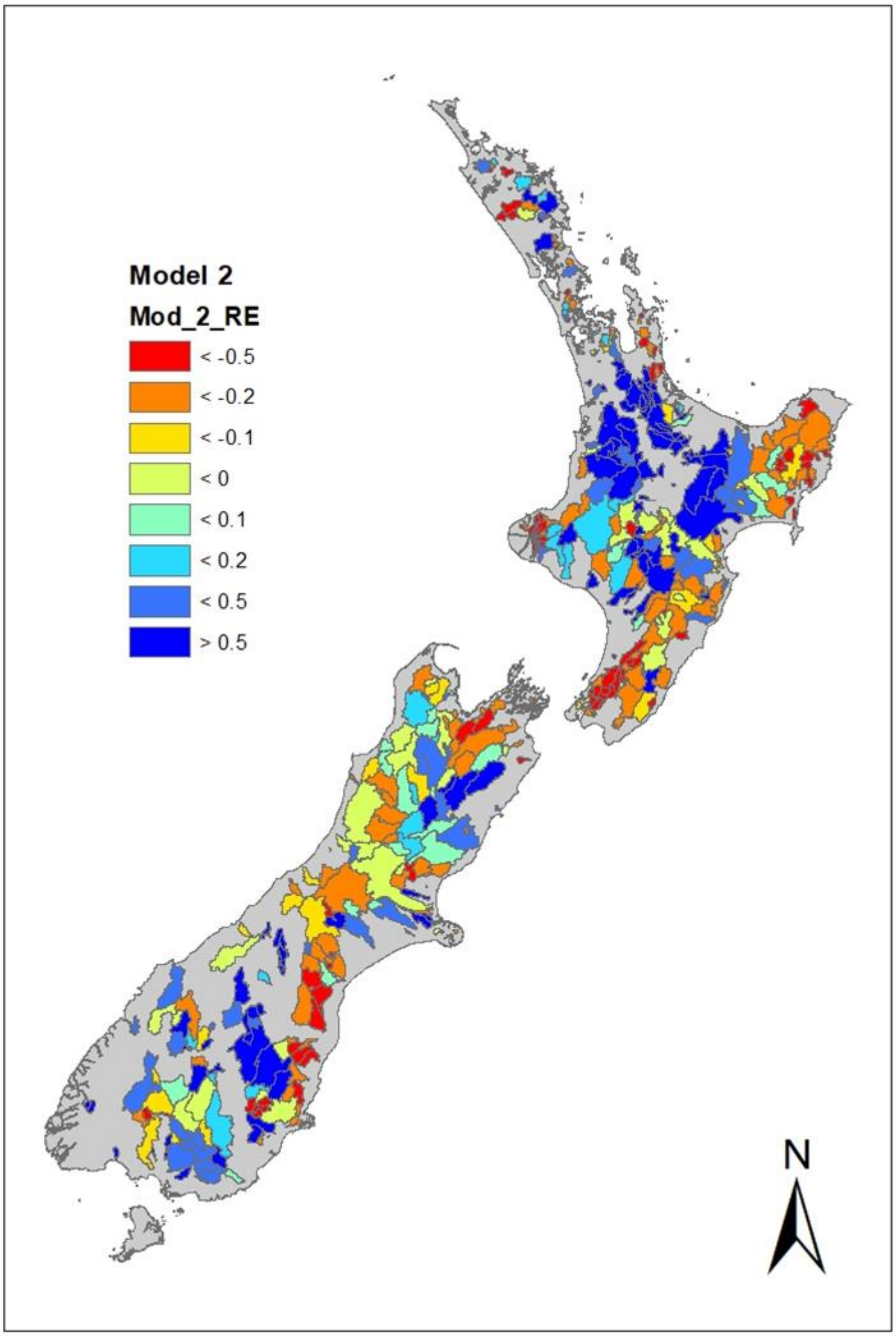

Figure 3-3: RMSWRE for model 2 (area + rain). Note reduced bias areas in South Island but still large contiguous areas of consistent bias in the North Island.

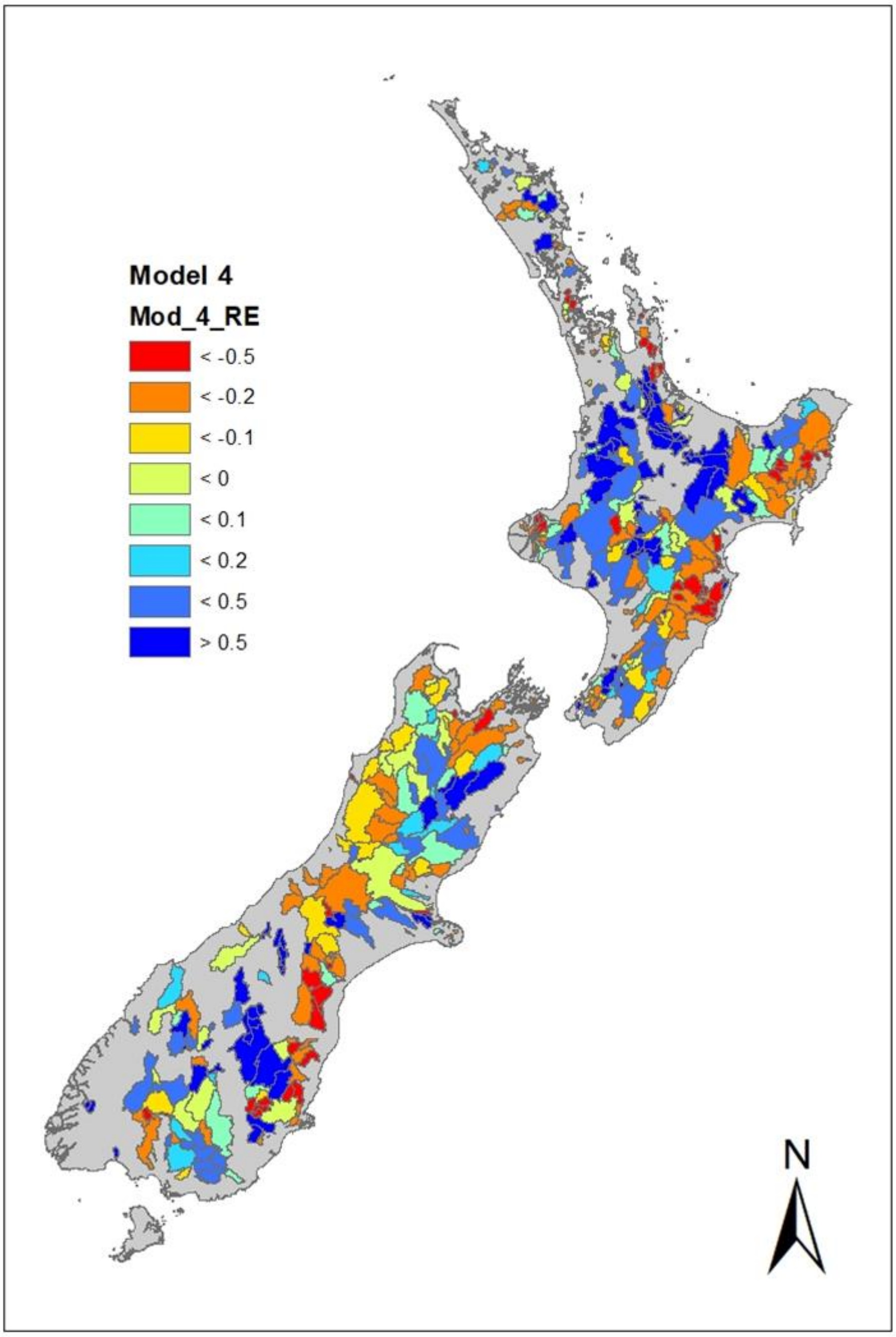

Figure 3-4: RMSWRE for model 4 (area + rain + catchment). Note still contiguous areas of consistent bias in the North Island.

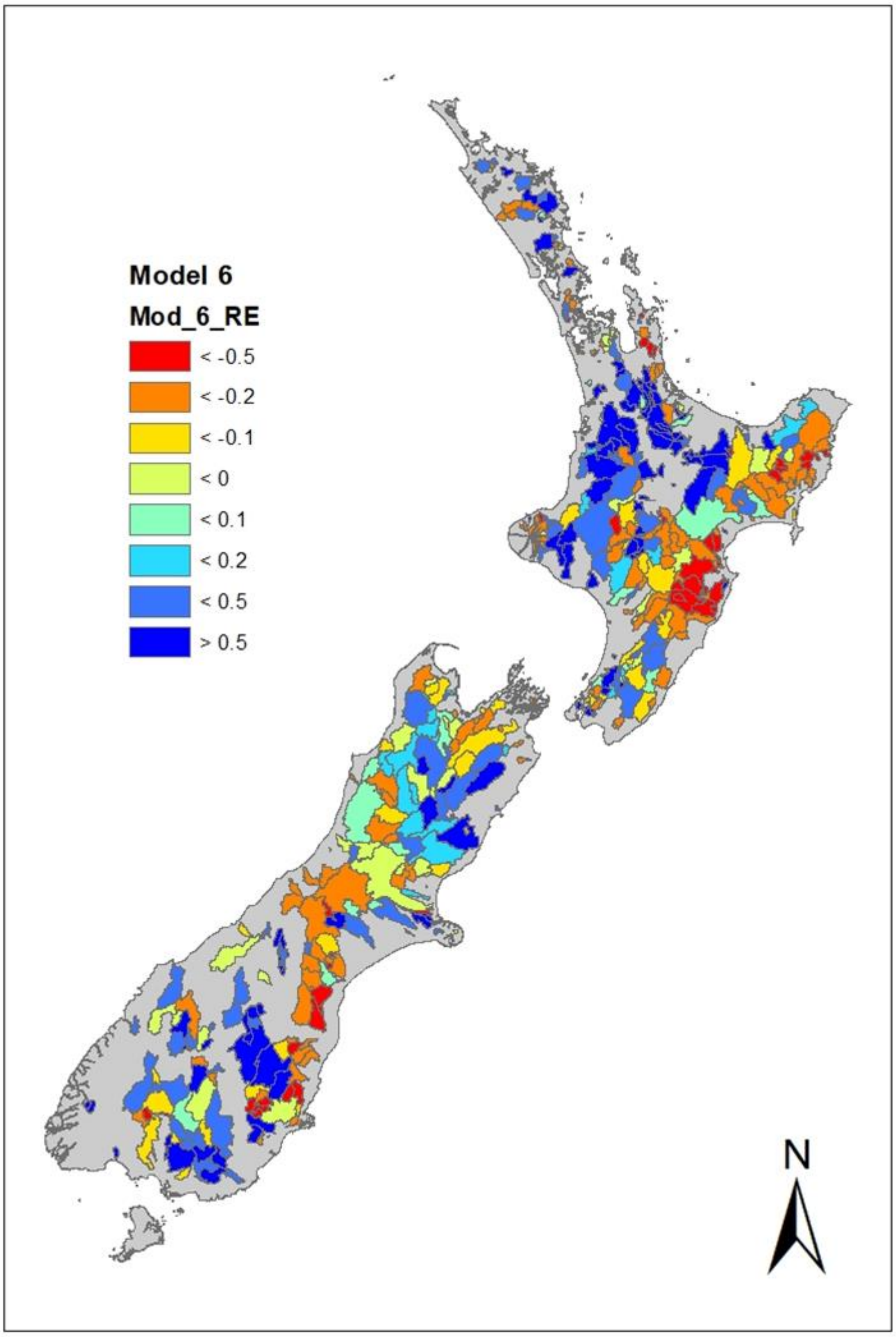

Figure 3-5: RMSWRE for model 6 (area + rain + two catchment). Note there are still contiguous areas of consistent bias in the North Island, and little change in the South Island wrt model 4 or model 2.

3.4 Contouring and error reduction

The errors calculated at each catchment (located at the catchment centroid) were contoured using ANUSplin. The contoured error surfaces were then applied to each catchment centroid to provide a corrected model result. This is akin to the approach of McKerchar and Pearson (1989) whose model is a manual version of model 2. After discussion at the face to face meeting of 14 December 2017 it was agreed that the area coloured maps do not represent the contouring approach so well, since the contouring process only sees each catchment as a point. Hence the maps in the figures below show each catchment as a point at the catchment centroid. Only models 1, 4 (NI) & 2 (SI), and model 6 are shown, as these seemed the most promising break points for model development.

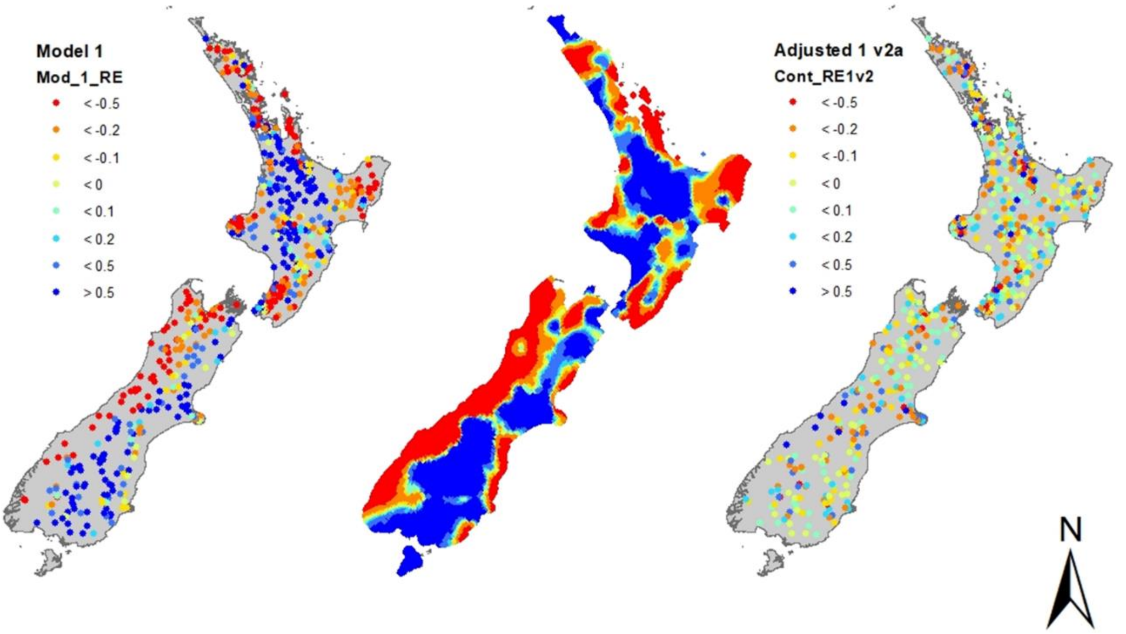

Figure 3-6: **Errors of model 1, contoured error, and the result of applying the contoured error to the modelled estimates.**

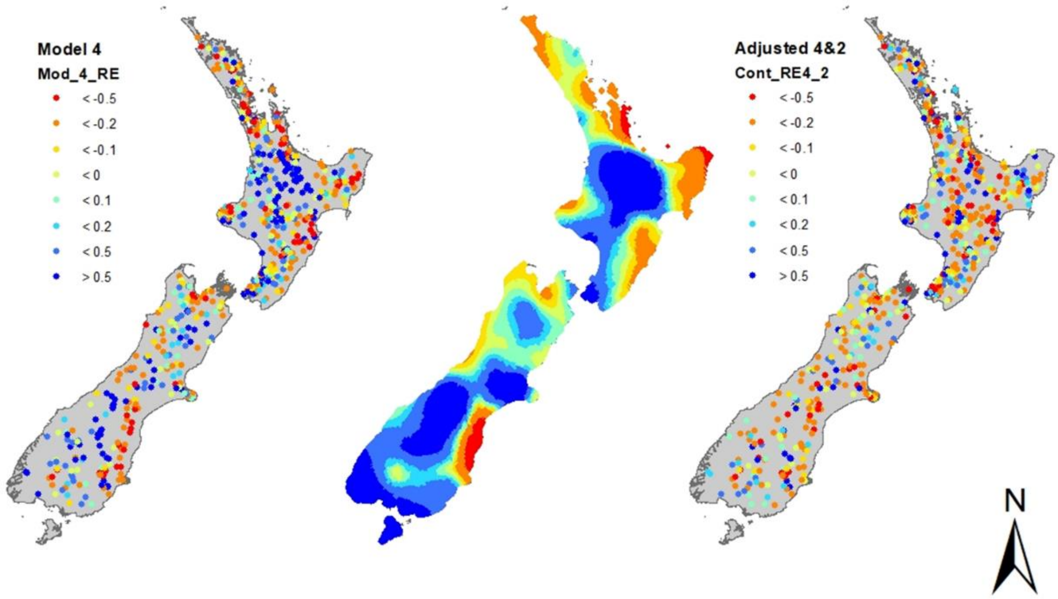

Figure 3-7: Errors of model 4(NI)&2(SI), contoured error, and the result of applying the contoured error to the modelled estimates.

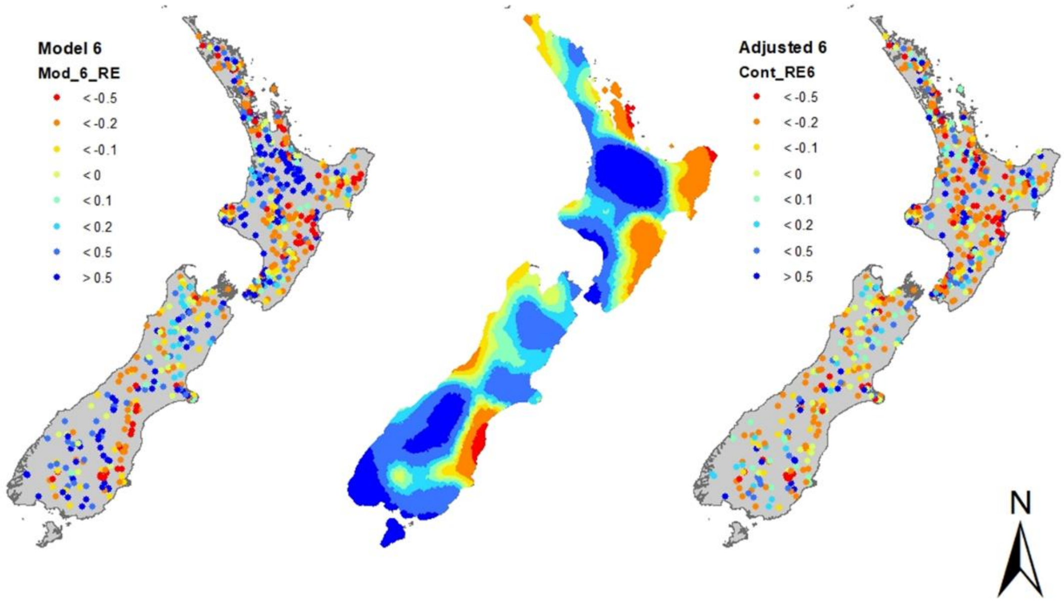

Figure 3-8: Errors of model 6, contoured error, and the result of applying the contoured error to the modelled estimates.

3.5 Two alternative regionalisation procedures

Suggestions were made by Martin Doyle and Jeff Watson at the 9 November 2017 virtual meeting to investigate dividing the North Island into two portions based on the relative influences of rain and geology. This would be tried in two ways: firstly, by applying the South Island rain equation (model 2) to the North Island, and examining the residuals of this in the North island; secondly by using expert knowledge to divide the island somewhere to the west of the main dividing range, but east of the influence of the volcanic soils of the central plateau and Rotorua areas.

3.5.1 North Island rain anomalies

Figure 3-9 shows the application of the South Island rain model, with residuals, to the North Island. A clear zone is evident in the east and south on the left-hand map, where the rain model under-estimates flood magnitude. The right-hand map shows the attribution of rain or geology dominance to catchments. Some of the northern Hawkes Bay rivers seem influenced by their headwater behaviour, but this is a limitation of the catchments available for the study in that area.

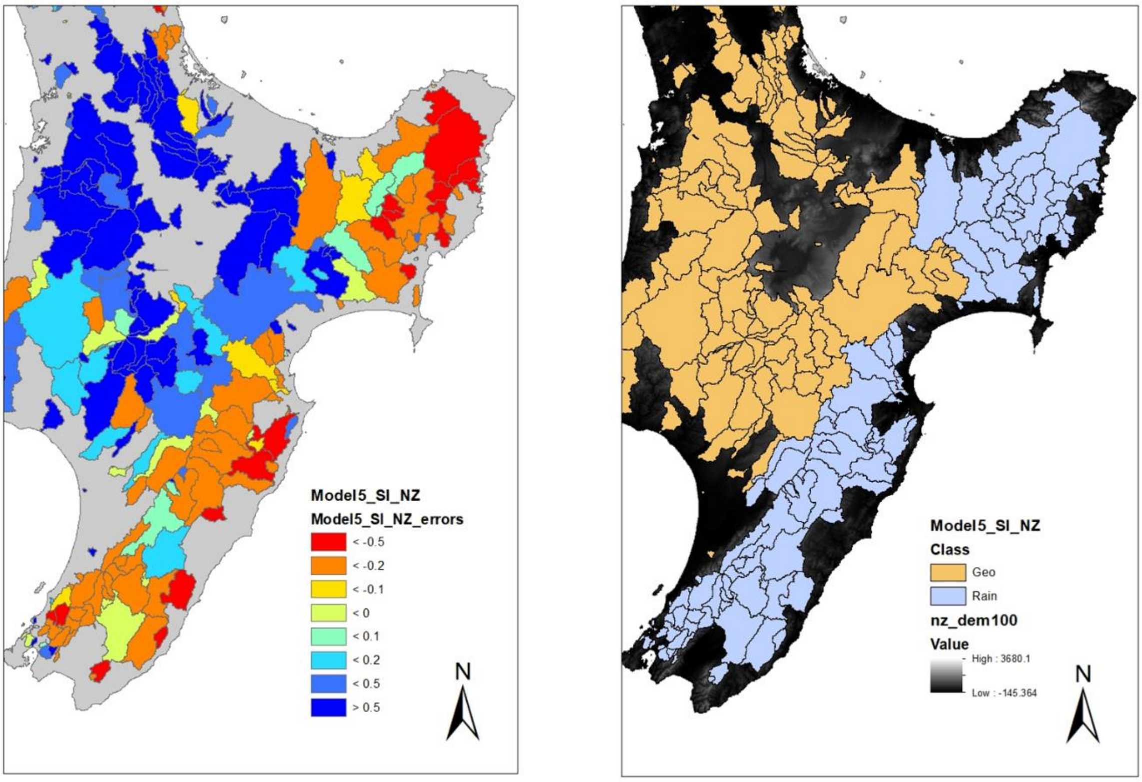

Figure 3-9: South Island rain model applied to the North Island. Left map is the residuals of the model; right map is the attribution of catchments to rain or geology influence.

The model progression described in section 3.3 was applied to these redefined 'islands' but only out to the equivalent of model 7. These produced overall and split error results very similar to those derived from the simple island division (see Figure 3-10). The enhanced South Island result was slightly better than the original, and the simplified North Island slightly worse.

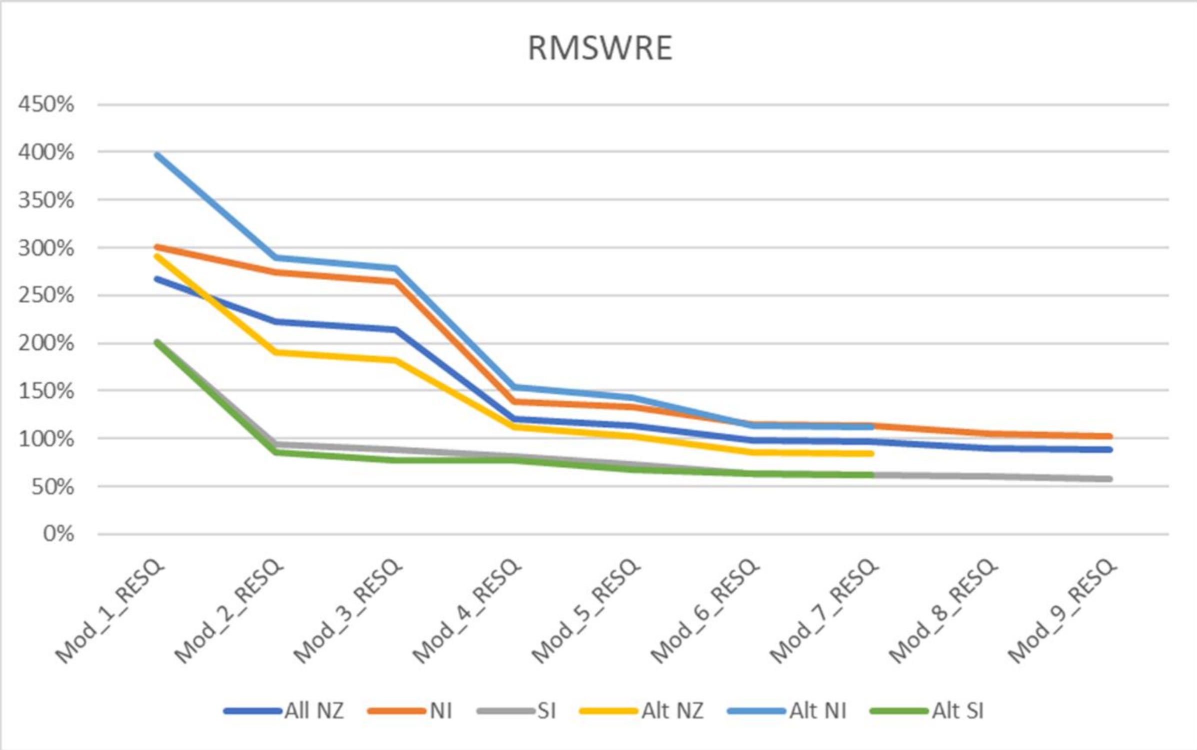

Figure 3-10: Average errors when the North Island is divided by geology and rain. Simplified North Island results are slightly worse than before, South Island slightly better.

3.5.2 Expert separation

An attempt was made to divide the North Island based on a number of previous classification systems or analyses: Toebes and Palmer (1969); Hutchinson (1980) hydrogeology classes and 5-year 7-day low flow regions; the flood regions and contours of Beable and McKerchar (1982) and McKerchar and Pearson (1989); and the hydrogeology variables from QMAP. Maps of these methods are illustrated in the following figures (Figure 3-11 to Figure 3-14).

The dividing line is based primarily on Toebes and Palmer (1969) and then located by eye on all the following maps.

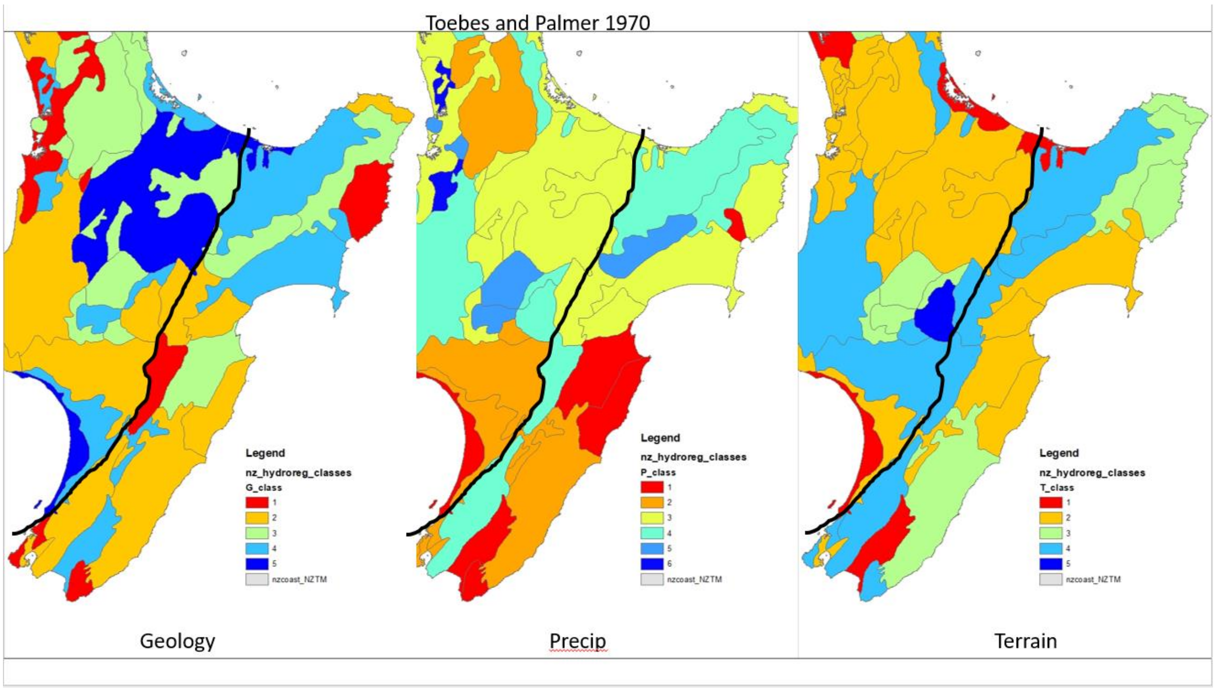

Figure 3-11: Toebes and Palmer (1969) regions. Three variables are divided into several classes for the purpose of delineating supposedly homegeneous hydrological regions.

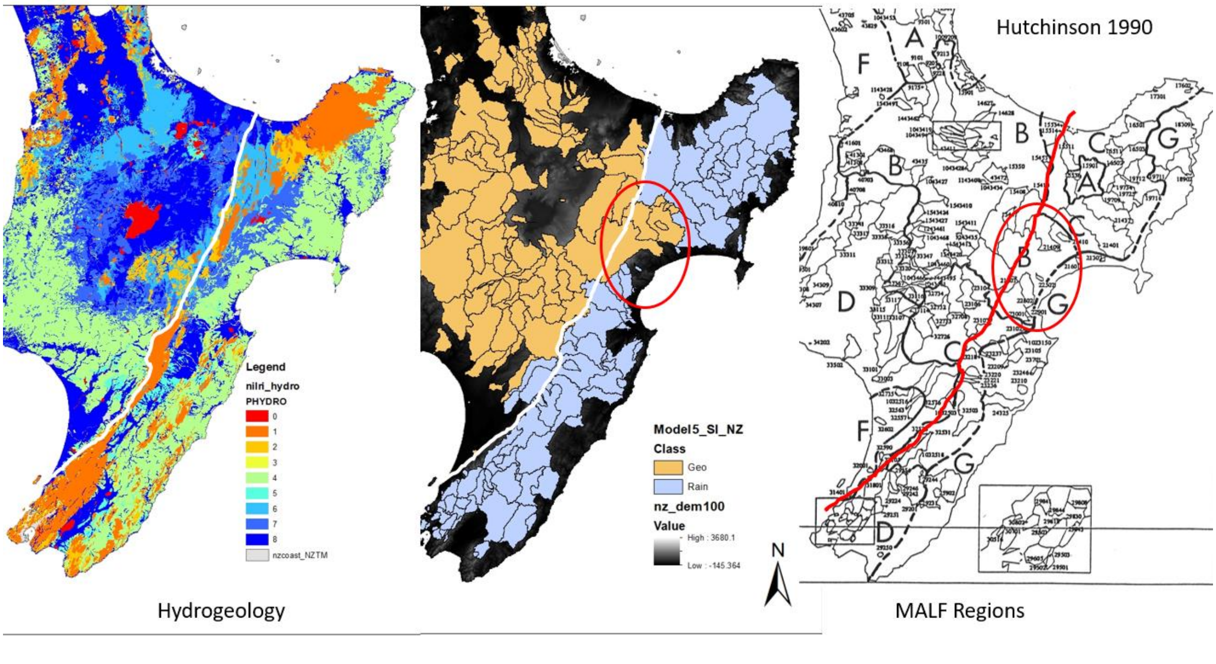

Figure 3-12: Hutchinson's (1980) hydrogeology (left) and low flow regions (right). The centre map shows the division based on rain model residuals from the previous section.

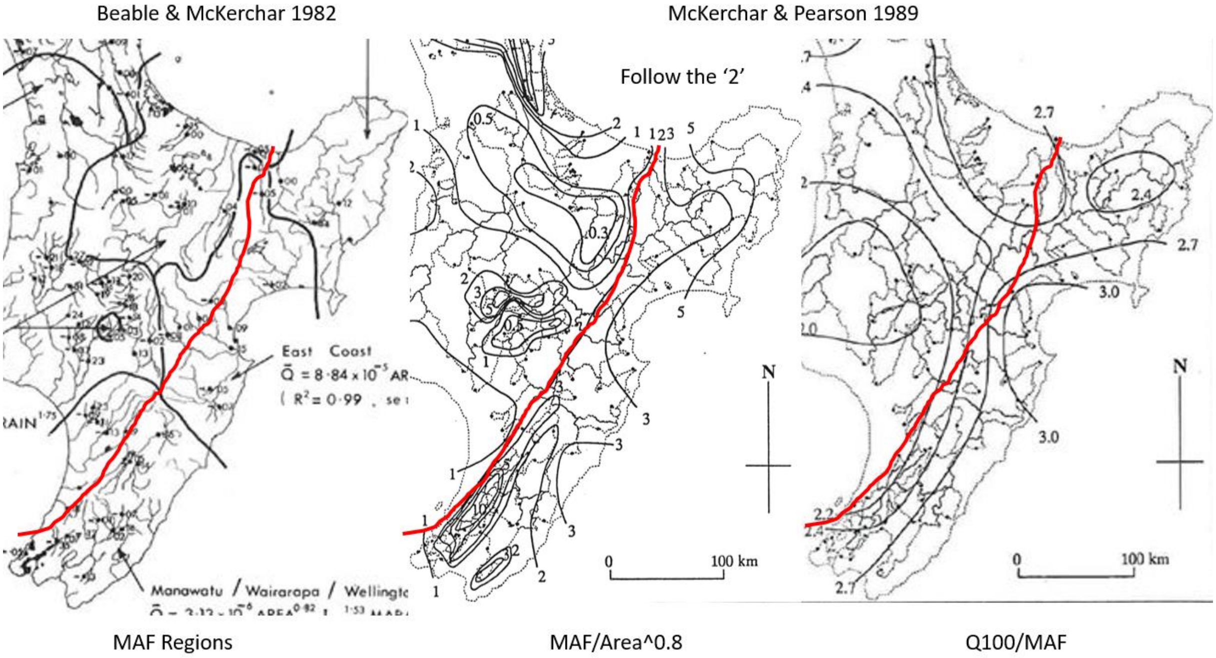

Figure 3-13: Beable and McKerchar (1982) flood regions, and McKerchar and Pearson (1989) contours. There is little correspondence with the other regions in these maps.

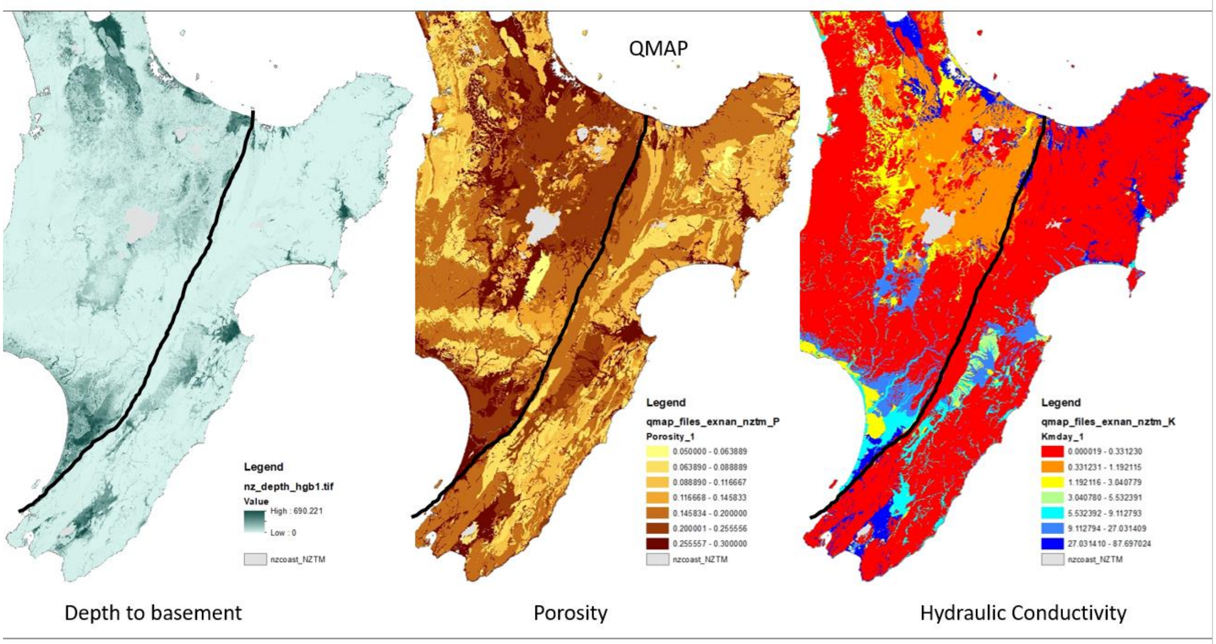

Figure 3-14: QMAP variables potentially related to hydrogeology. Some parts of the boundary are related but not consistently.

Applying the same model development process to the enhanced South Island and simplified North Island based on the boundary illustrated above, produced average errors very similar to the ones in the previous section.

3.6 Signal to noise ratio

Analysis of signal to noise ratios for the contouring showed that this was approximately 1:7, far larger than generally found when using ANUSplin for contouring climate variables. Two alternatives were trialled (1:4 and 1:1) to assess the sensitivity of the contouring to this parameter. Maps of the 1:1 contours follow (Figure 3-15 to Figure 3-17).

The centre map in each case is the adjusted contour set, which for each model, exhibit clear signs of the contouring being stretched far more tightly across the data than in the minimised generalised cross-validation (GCV) contours of the previous section.

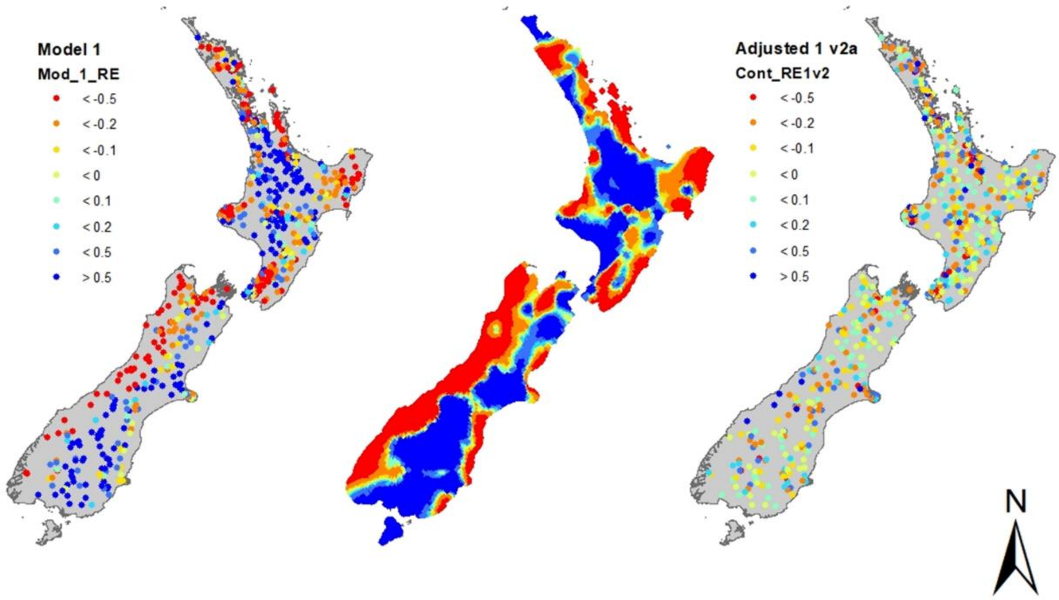

Figure 3-15: Errors of model 1, contoured error with signal to noise 1:1, and the result of applying the contoured error to the modelled estimates.

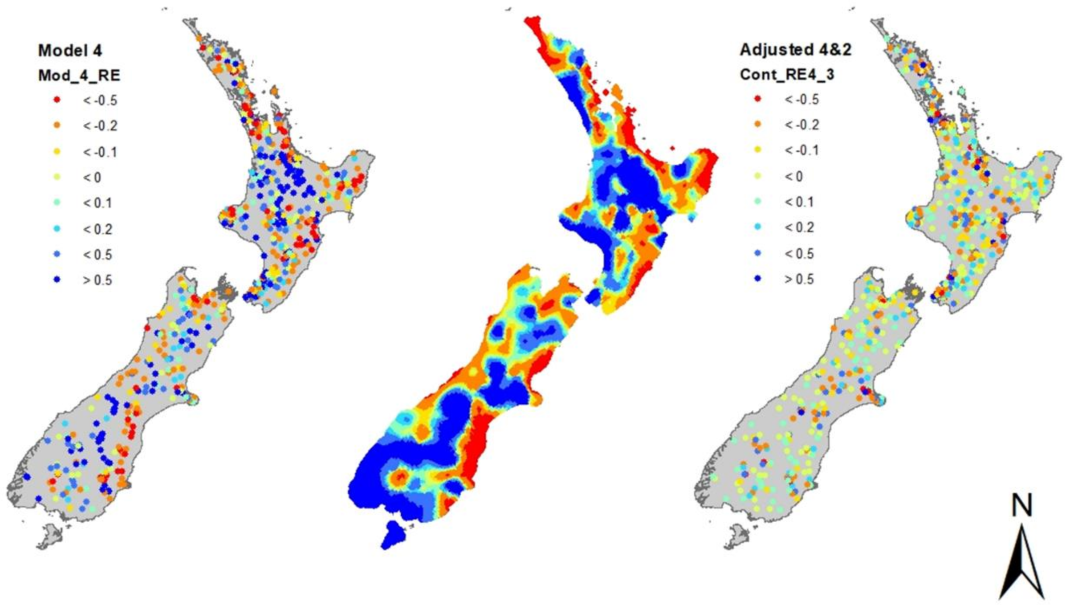

Figure 3-16: Errors of model 4(NI)&2(SI), contoured error with signal to noise 1:1, and the result of applying the contoured error to the modelled estimates.

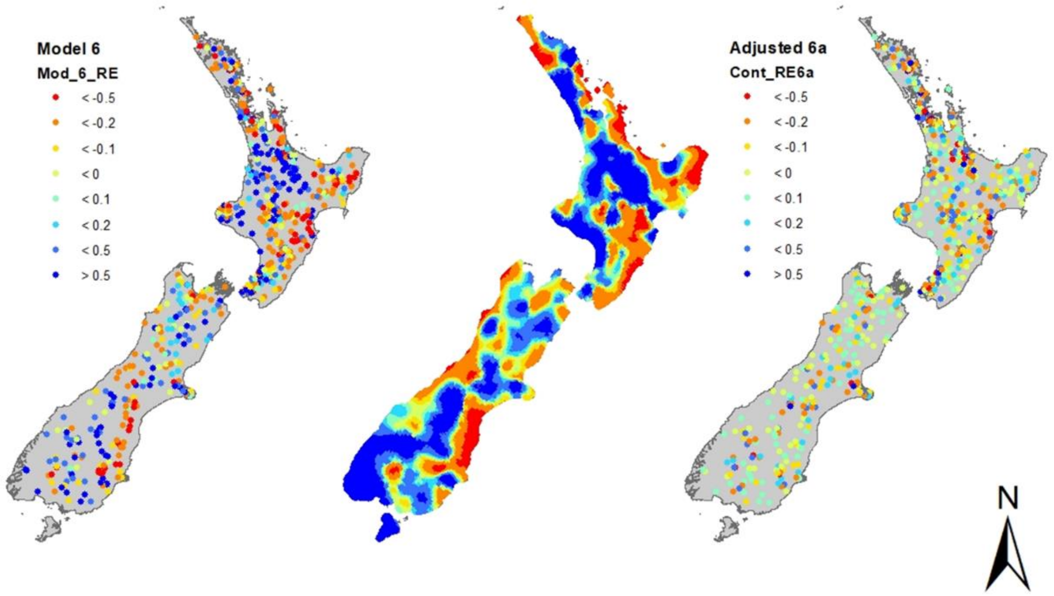

Figure 3-17: Errors of model 6, contoured error with signal to noise 1:1, and the result of applying the contoured error to the modelled estimates.

After contouring, the model estimates can be adjusted by applying the contoured error estimate, to provide an alternative model with the contours explaining some of the variability not captured in the regressions. Table 3-1 shows the regression model error for models 1 and model 4&2, the error after contouring with minimised GCV, and the error after contouring with the signal to noise ratio fixed at 1:1.

Table 3-1: Relative errors before and after contouring. Rows are for NZ overall, North Island and South Island respectively. Columns are: extent; Model 1 regression errors; Model 1 errors after contouring with minimum GCV; model 1 errors after contouring with 1:1 signal to noise; Model 4&2 regression errors; Model 4&2 errors after contouring with minimum GCV; model 4&2 errors after contouring with 1:1 signal to noise.

Region	RE Regression Model 1	RE Model 1 adjusted by contouring to minimise GCV	RE model 1v2a adjusted by contouring	RE Regression Model 4&2	RE Model 4&2 adjusted by contouring to minimise GCV	RE Model 4&2v3a adjusted by contouring
NZ	267%	75%	49%	124%	50%	26%
NI	301%	61%	30%	139%	52%	29%
SI	202%	94%	70%	93%	46%	20%

As expected, assessed errors are reduced after contouring, and further reduced after the contouring algorithm is tightened. However, this is not an adequate assessment of the likely error in an ungauged catchment.

3.7 Leave one out error assessment

The objective of the contouring is to, if possible, smooth out errors in the model and enhance the regional aspect by this. Excessive tightness in the contouring did not advance this objective, but simply stretched the contour fit more closely to the data from flow recorders. To test this, we examine the leave-one-out (LOO) properties of the contouring. A dataset for each flow record, with that flow record left out, is contoured to mimic the performance of the contours at points without data. The error assessment provided by this approach is close to the actual uncertainty of the overall model.

Table 3-2: Relative errors from the LOO assessment. Rows are for NZ overall, North Island and South Island respectively. Columns are: extent; Model 1 regression errors; Model 1 errors after contouring with minimum GCV; LOO result; model 1 errors after contouring with 1:1 signal to noise; LOO result; Model 4&2 regression errors; Model 4&2 errors after contouring with minimum GCV; LOO result; model 4&2 errors after contouring with 1:1 signal to noise; LOO result.

Region	RE Regression Model 1	RE Model 1 adjusted by contouring to minimise GCV	Leave one out	RE Model 1v2a adjusted by contouring with fixed signal to noise	Leave one out	RE Regression Model 4&2	RE model 4&2 adjusted by contouring to minimise GCV	Leave one out	RE Model 4&2v3a adjusted by contouring with fixed signal to noise	Leave one out
NZ	267%	75%	175%	49%	231%	124%	50%	71%	26%	100%
NI	301%	61%	200%	30%	256%	139%	52%	84%	29%	112%
SI	202%	94%	124%	70%	184%	93%	46%	42%	20%	77%

Table 3-2 shows that each LOO column error is larger than the error assessed from the contouring alone, in the column immediately to the left. The LOO errors from the tight contouring (columns 6 and 11) are nearly as large as the errors of the original regression models (columns 2 and 7). Those for the GCV minimum contouring (columns 3 and 8) are smaller, and clearly the lowest errors are from the minimum GCV contouring of model 4&2, making this the preferred model (see Figure 3-7).

To further examine the effect of the signal to noise ratio, another run was carried out with signal to noise forced to 1:4. Figure 3-18 shows the effect of this for two models; model 1 and model 4&2, and for all New Zealand and each island separately. The error effect of signal to noise varies monotonically, and since the leave one out errors shown in Table 3-2 are greater than those for MinGCV for the signal to noise ratio at 1:7, we assume that they will also be greater for signal to noise 1:4, although not by such a margin.

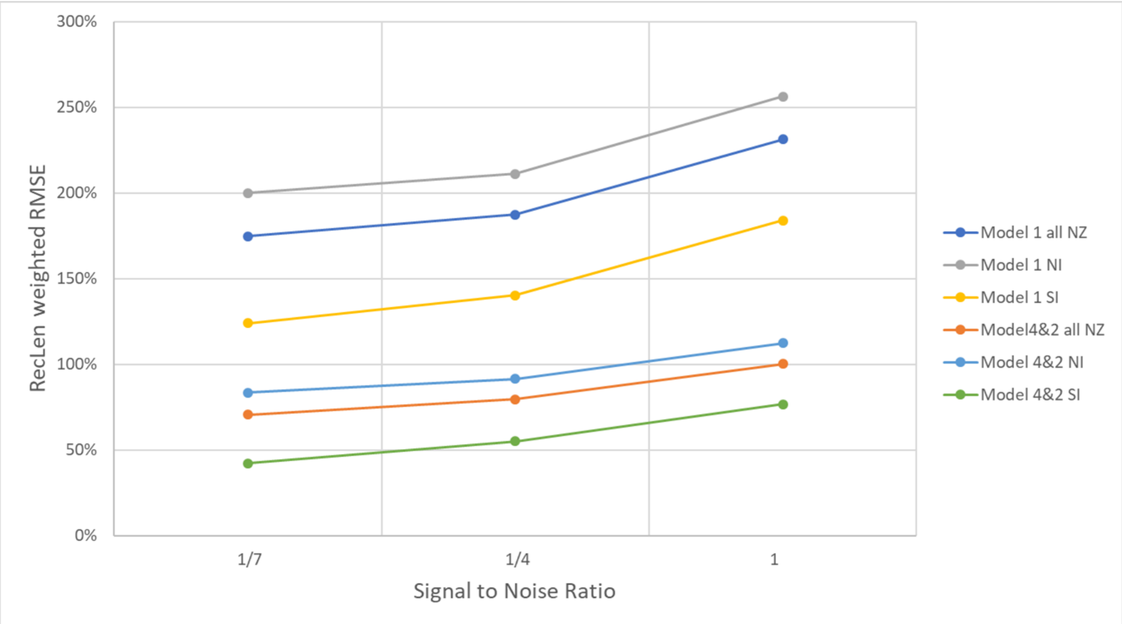

Figure 3-18: Effect of signal to noise forcing on overall model error.

3.8 Error distribution

While errors overall were improved by the application of a contoured surface, there remained a concern about error distribution. Figure 3-19 shows the error distribution resulting from three different models and three different estimation techniques.

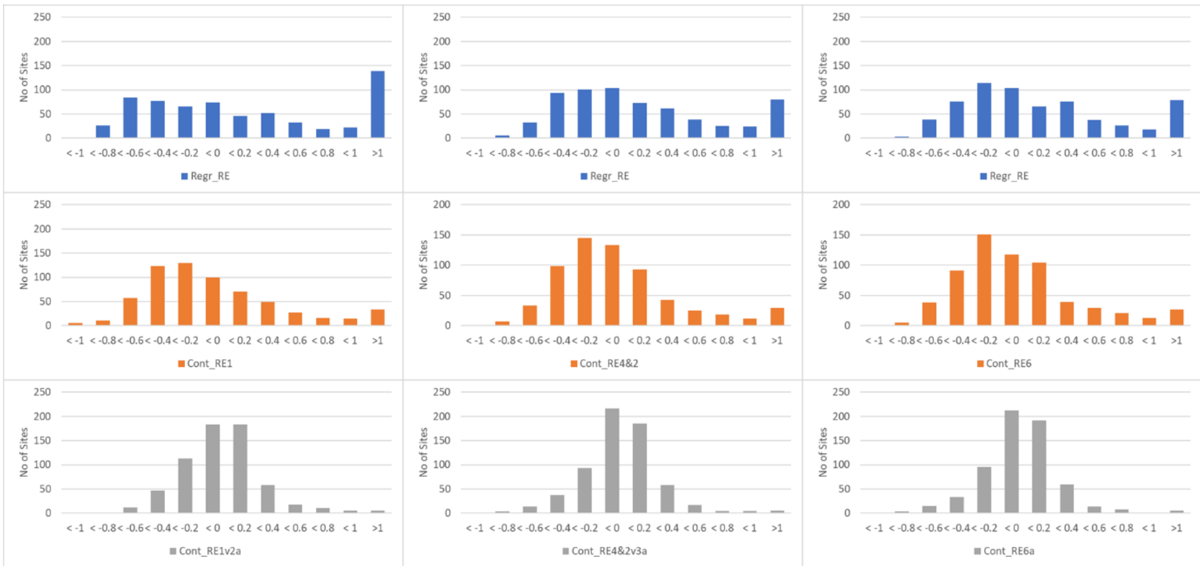

Figure 3-19: Distribution of errors from three models and three estimation techniques. Figures in columns show the error distribution from Model 1, Model 4&2 and Model 6 respectively. Figures in rows show the error distribution of the regression model, the GCV minimum contouring, and the 1:1 signal to noise contouring respectively.

Errors from the straight regression approaches are bi-modal, with an excess of values greater than one. This largely results from the errors being expressed in terms of MAF rather than log_10(MAF). Ordinary Least Squares regression, which was applied to the log-transformed variables, produces an unbiased regression fit; by calculating the error on the non-log-transformed MAF values this introduced a bias. A further complication is use of the error formula from McKerchar and Pearson (1989). This results in errors that range from -1 to +infinity, and indeed the contouring of these errors can produce negative values that are outside this range, and similarly some very large positive values. This tendency persists into the Minimised GCV contours, and is only removed with the very tight contouring at signal to noise 1:1. In all cases the error distribution is biased, positive for the regression, and up to 10% negative for the contoured results.

Discussion between NIWA and regional council staff at a workshop on 8 December 2018, where the above results were presented, led to the proposition that a more standardised error formula using the log ratio (or difference of logs) should be used, and this is the subject of the following section.

4 Final model derivation

4.1 Regression model selection

The brute force analysis for regression model selection was re-run, but using the log error formula,

$$LogErr = \log(MAF_{data}) - \log(MAF_{model}) \quad (2)$$

weighted by record length as in equation 1, to assess the best regression parameter set. The factorial errors of the nine models described in section 3.3 are presented in Figure 4-1. Factorial error is

$$Factorial\ Error = 10^{LogErr} \quad (3)$$

and the standard error ranges from Q_{mod} . Factorial Error to Q_{mod}/Factorial Error.

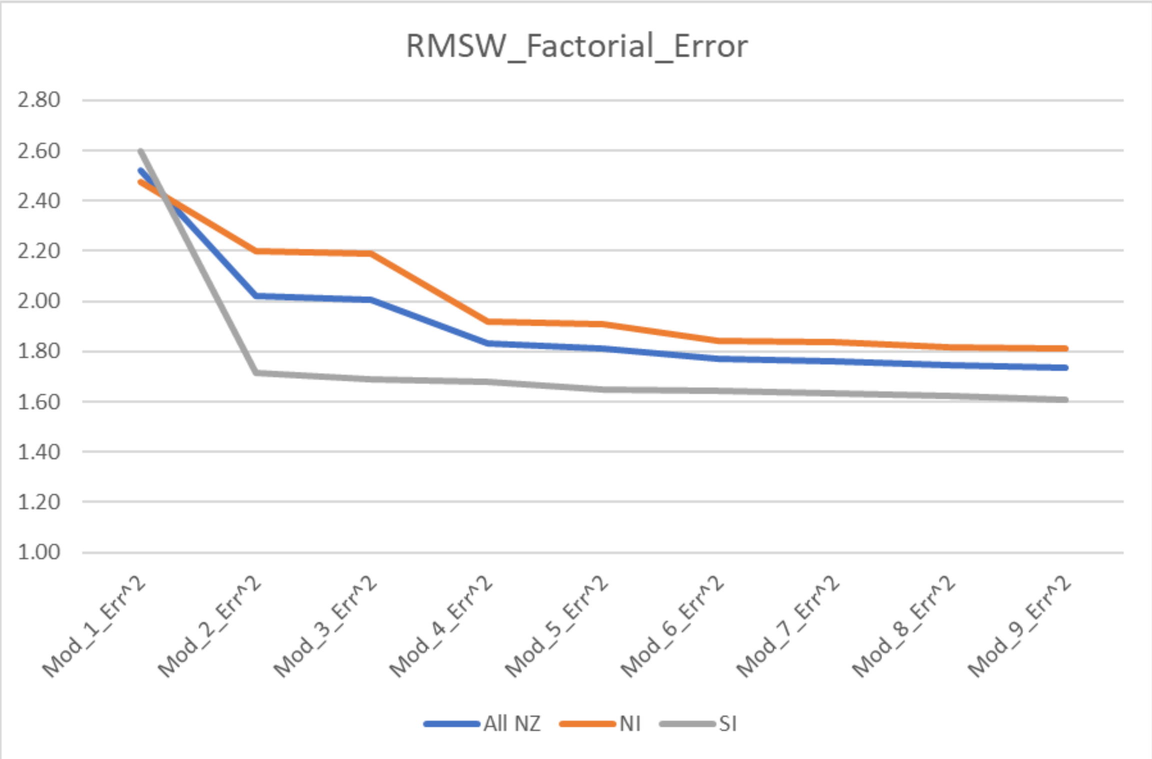

Figure 4-1: Factorial record length weighted errors of best model in each of nine categories. Vertical axis truncated at 1 as this is the minimum value of an RMS factorial error.

As for Figure 3-1, model 4 for the North Island and model 2 for the South Island, provide the largest change in error for the smallest number of parameters.

The chosen regression model is thus model 4 for the North Island, and model 2 for the South Island, with model formulae as follows:

$$MAF(NI) = k1 \,.\, A^{c1} \,.\, MeanAnnualRain^{c2} \,.\, HI68^{c3} \quad (4)$$

$$MAF(SI) = k2 \,.\, A^{c4} \,.\, NewPrecip^{c5} \quad (5)$$

Where: MAF is mean annual flood in m^3/s
A is catchment area in km^2
MeanAnnualRain is average rainfall in mm/annum
HI68 is the fraction of the catchment that is hydrogeology 6,7 or 8 (plus 0.01)
and
$k1 = 2.0099 * 10^{-6}$
$c1 = 0.9041$
$c2 = 1.7351$
$c3 = -0.3338$
$k2 = 9.7694 * 10^{-5}$
$c4 = 0.7982$
$c5 = 1.3577$

The estimates from this model combination were sent for contouring as before.

4.2 Contouring

Contouring of the log errors using ANUSplin produced two estimates; the first minimising GCV, and resulting in a signal to noise ratio of 1:4 (cf. 1:7 when contouring the relative errors); the second setting the signal to noise ratio to 1:1. The results of these two approaches are shown in Figure 4-2 and Figure 4-3.

The adjusted estimate of MAF error mapped in the right-most map of each figure is derived as follows:

Adjusted MAF = Regression model estimate * 10 ^ (contoured error), (6)

and the adjusted error value is calculated by:

Adjusted MAF error = log (MAF_{data}) – log (Adjusted MAF). (7)

Error reduction is substantial under both approaches, and the altered distribution of errors is shown in Figure 4-4.

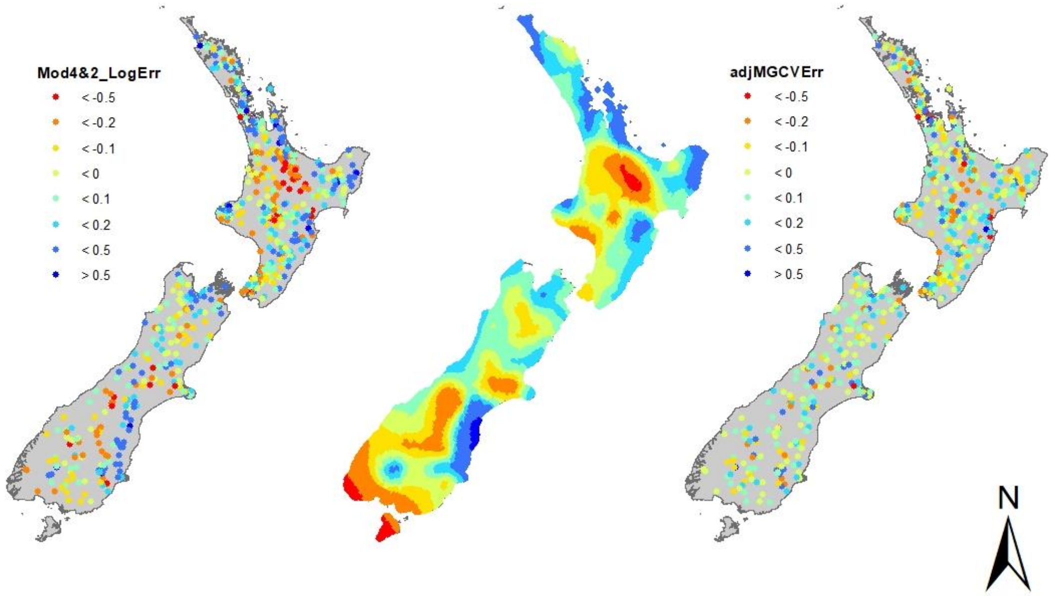

Figure 4-2: Contouring with minimised GCV. Original model errors at sites (left map), contoured error surface (centre map), and adjusted errors at sites (right map).

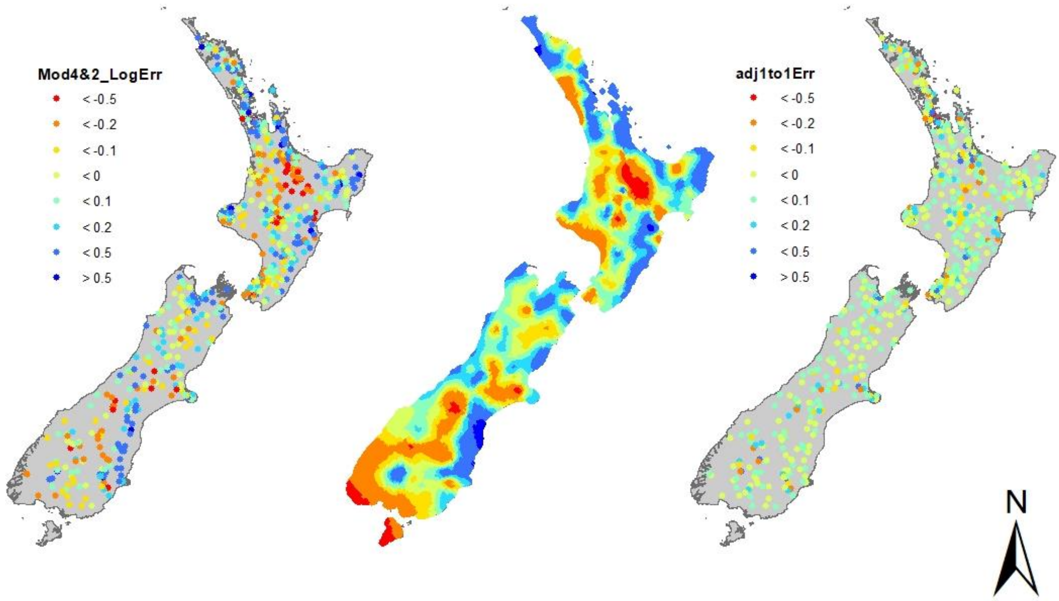

Figure 4-3: Contouring with signal to noise 1:1. Original model errors at sites (left map), contoured error surface (centre map), and adjusted errors at sites (right map).

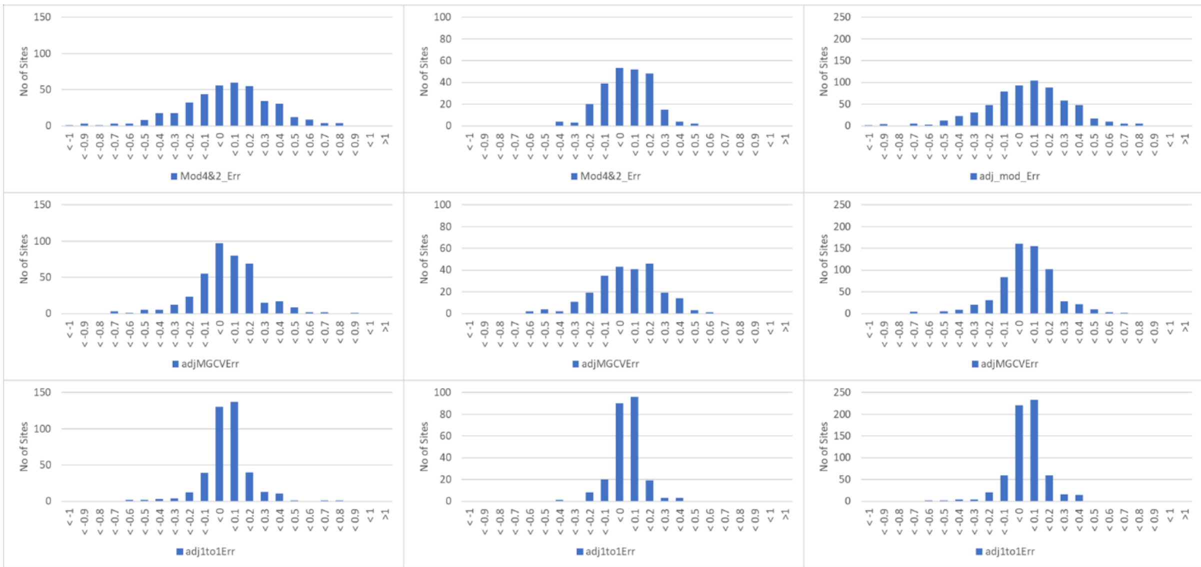

Figure 4-4: Error distributions before and after contouring. Left graphs show North Island sites; middle graphs show South Island sites; right graphs show All NZ sites. Top graphs show at site regression log errors, middle graphs show log errors with minimised GCV, bottom graphs show log errors with signal to noise ratio set at 1:1.

4.3 Leave one out error assessment

While the overall result at the flow sites seems best with the 1:1 signal-to-noise ratio, the real test of this is when a leave-one-out cross validation (LOOCV) analysis is performed. LOOCV analysis measures how well the model would fare in predicting values at sites not used in the fitting process. We have carried out two LOOCV analyses. The first is on the raw regression result, by leaving each site out of the regression analysis in turn, and deriving an individual regression equation. Each of these equations are of the same form (have the same variables) as the North and South Island equations, but have different parameters. These individual equations are then used to estimate the value of MAF at their respective left-out sites. Errors for this LOOCV analysis of the regression equation can be found in column three of Table 4-1. The second stage LOOCV analysis calculates estimates of MAF with the individual equations above, at the 635 included sites, and the errors of the 636 sets thus derived are contoured. The contour surfaces generated are then interrogated to find the contoured error at the left-out site, and an adjusted MAF estimate and error is calculated at that location. Columns five and seven of Table 4-1 contain the error analysis from this second LOOCV process.

Equivalent relative error is calculated by applying the as the multiplied and divided factorial errors, calculating the relative error of each, and taking the geometric mean of the result, according to equation 8.

Equivalent relative error = SQRT [(Factorial error – 1). –(1/Factorial error -1)]. (8)

An alternative error formulation is mean absolute percentage error (MAPE), expressed as:

MAPE = sum[ABS{(MAFdata – MAFmodel)}/MAFdata)]/n (9)

Table 4-1: Effect of contouring and leave one out on RMSLE and factorial error. Each contouring set is shown as the straight application of contours and as the leave one out result. Log errors are the root mean squared weighted error; factorial errors are 10^(log error); equivalent relative error is the geometric mean of the result of multiplying and dividing by the factorial error.

Extent	Models 4 (NI) and 2 (SI)		ANUSplin with MinGCV		ANUSplin with 1:1 signal:noise	
	Raw regression	LOOCV regression	Contour	LOOCV	Contour	LOOCV
			Log Errors			
All NZ	0.266	0.269	0.173	0.260	0.118	0.277
North Island	0.283	0.286	0.188	0.251	0.133	0.256
South Island	0.235	0.237	0.146	0.273	0.089	0.308
			Factorial Errors			
All NZ	1.84	1.86	1.49	1.82	1.31	1.89
North Island	1.92	1.93	1.54	1.78	1.36	1.80
South Island	1.72	1.73	1.40	1.87	1.23	2.03
			Equivalent Relative Errors			
All NZ	62%	63%	40%	61%	27%	65%
North Island	66%	67%	44%	59%	31%	60%
South Island	55%	55%	34%	64%	20%	72%
			Mean Absolute Percentage Errors			
All NZ	64%	65%	35%	67%	22%	72%
North Island	69%	71%	39%	66%	25%	67%
South Island	55%	55%	29%	70%	16%	79%

As noted above, the best apparent results are obtained with the 1:1 signal-to-noise ratio, but when the LOOCV analysis is done, this result worsens, so that the best result is for the leave one out Minimise GCV approach. Thus, we choose the MinGCV contour set to apply to the selected regression models to produce the result everywhere in New Zealand. Errors from column 5 of Table 4-1 provide the error estimate for each island.

4.4 Comparison with previous study

Two comparisons can be made with the previous method.

Firstly, their approach was basically a regression (of MAF vs. area) and a contour map of the residuals. In our analysis, this is equivalent to the application of the best regression equation (model 4 & 2 in NI and SI respectively), followed by contouring without further analysis. Additionally, the top 5% of outliers should be removed. The error contours of the previous method appear to be between our MinGCV and 1to1 contour methods for complexity vs. smoothness, so we average these two methods. The resulting error assessment is a standard error of ±26%, compared with their ±22%.

Secondly, we can attempt a leave one out style of analysis of the previous method, by assessing it against either all sites used in that study that are part of the present study, or against sites in the present study that were not used in the previous study. A true leave one out analysis was not possible previously as noted in their report, as it was impractical to repeat the drawing of contours many times. For all sites in common between studies, the relative error assessed against the present dataset, is ±49%, and for sites not used previously, ±70%. Removing the top 5% of errors from our final chosen full LOOCV model, the factorial error reduces to 1.62, or an equivalent relative error of ±49%, which is at the lower end of the assessed errors from the previous work. Thus following the previous study, we propose the adoption of ±50% as the standard error of estimate for the mean annual flood.

4.5 Sites with large errors

In the previous study some 5% of sites were identified as having larger errors, and these were excluded from the overall error assessment. Applying the same filter to the final model in the current study, produces the list of flow records shown in Table 4-2. There are many potential reasons why these sites do not fit the overall model as well as others. These include potential rating curve issues, length of record, limited numbers of sites with similar catchment characteristics in the vicinity, etc. Detailed consideration of such factors in a future revision is strongly recommended.

Table 4-2: Sites with the largest errors. Assessed against the final model after regression and contouring, ranked on the highest log error squared across both islands. Those marked with * appear in the previous report's list of large errors.

Site ID	Name	MAF observed	MAF modelled	Factorial error
7604	Wairau at Motorway	8.1676	40.2609	20%
7912	Opanuku at Vintage Reserve	49.0321	9.4013	522%
8304	Mangemangeroa at Breadman	6.6609	41.6515	16%
15408*	Rangitaiki at Murupara	39.3032	150.3073	26%
22715	Pakuratahi at Forest Glade	2.2586	9.4693	24%
23169	Irongate at Clarks Weir	6.7425	1.4761	457%
23302	Maraetotara at Waimarama Rd	58.6432	319.273	18%
29819	Taita at Exotic	0.0161	0.1147	14%
46641	Waipao at Draffins Rd	14.941	61.006	24%
64615	Robinson Stm at Cascades Waitakere	0.2086	1.0686	20%
1009213	Oraka at Pinedale	14.7021	73.7653	20%
1009230	Kuhatahi at Weir	1.3815	14.0401	10%
1043419	Pokaiwhenua at Puketurua	31.5374	236.9283	13%
1043476	Otutira at Otutaru	0.1443	0.0373	387%
1114629	Waipa at Retail Yard	3.21	20.5682	16%
1143408	Purukohukohu at Puruorakau	0.0136	0.2561	5%
1443433	Puruwai at Gorge	0.0321	0.3891	8%
1643460	Clarkes Rd Stm at Clarkes Rd	4.0314	1.0141	398%
2043441	Waipapa at Mulberry Rd	6.7882	25.7002	26%
60197	Grovetown Lagoon from Brin Williman	2.6497	10.5409	25%
60901	Flaxbourne at Corrie Downs	56.3515	13.0373	432%

Site ID	Name	MAF observed	MAF modelled	Factorial error
66204	Ashley at Gorge	312.6498	72.352	432%
66213	Okuku at Fox Ck	163.7503	32.7565	500%
67805	Halswell at Ryans Bge	5.9354	71.3568	8%
68529*	Dry Acheron at Water Race	2.7622	11.0106	25%
68602	Dry Ck at RDR Syphon	10.4802	78.5438	13%
68806	Sth Ashburton at Mt Somers	89.3978	345.545	26%
70902	Waihao at McCulloughs Br	272.2837	60.4817	450%
71122*	Mary Burn at Mt MacDonald	4.1859	22.5169	19%
71130	Mary Burn at SH8	4.8794	40.5357	12%
75271	Mill Ck at Fish Trap	3.3982	16.4947	21%
78607	Oreti at Lumsden	492.64	90.5068	544%

4.6 Final model

After application of the MinGCV contouring surface to the estimates from the regression models for each island, a final dataset is derived that represents the best solution for the estimation of mean annual flood. Figure 4-5 shows the model fit in log space, in keeping with the log-log regression analysis reported above.

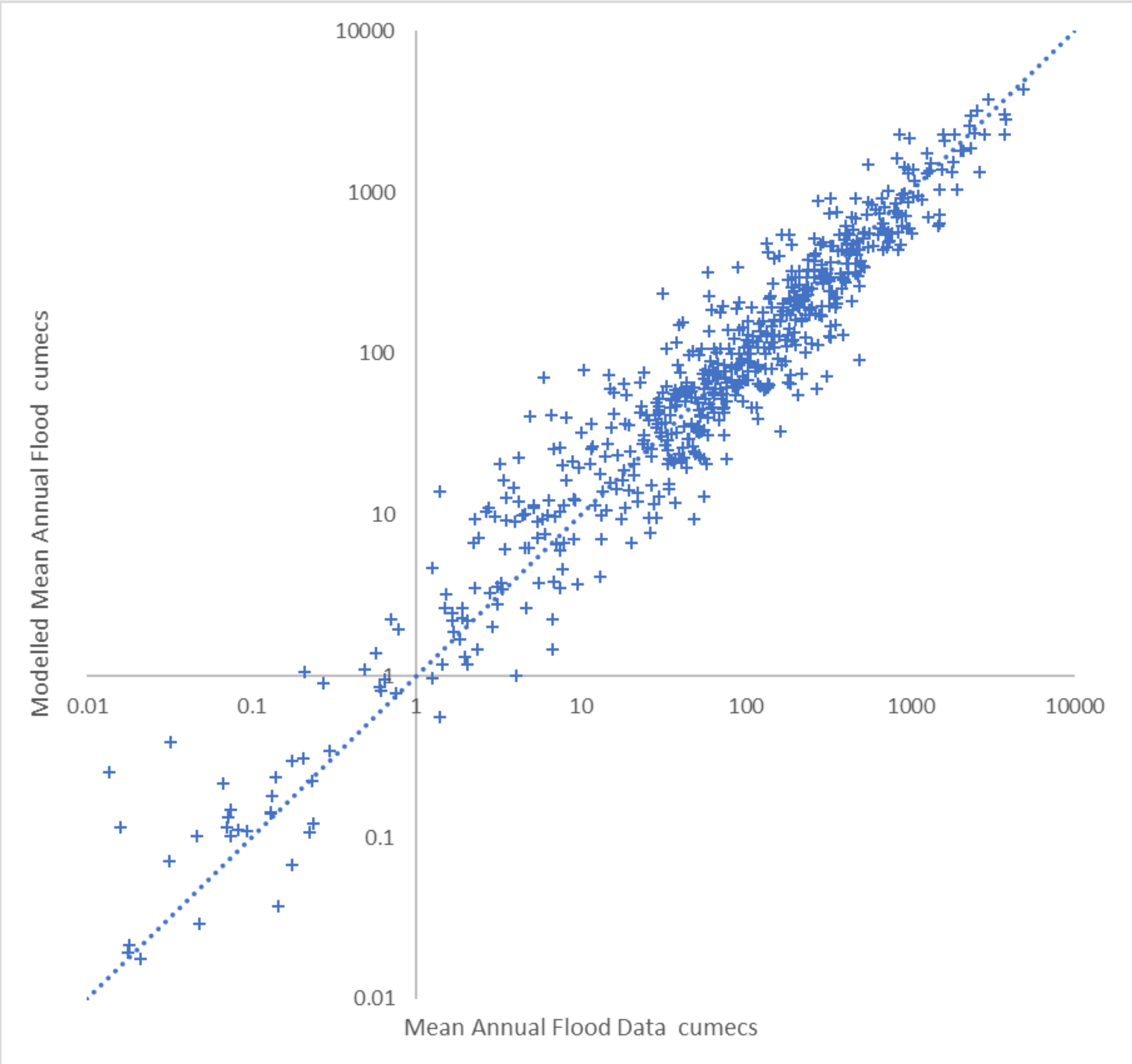

Figure 4-5: Fit of final model to mean annual flood data. Scales are logged, to reflect the log-linear regression.

The scatter about the 1:1 line and the density of data points shows some variation with flood size. The scatter may indicate greater uncertainty for floods from small catchments, as reported in the previous study; the varying density reveals a continuing sampling issue in which smaller catchments are less well represented as can be seen in Figure 4-6 below.

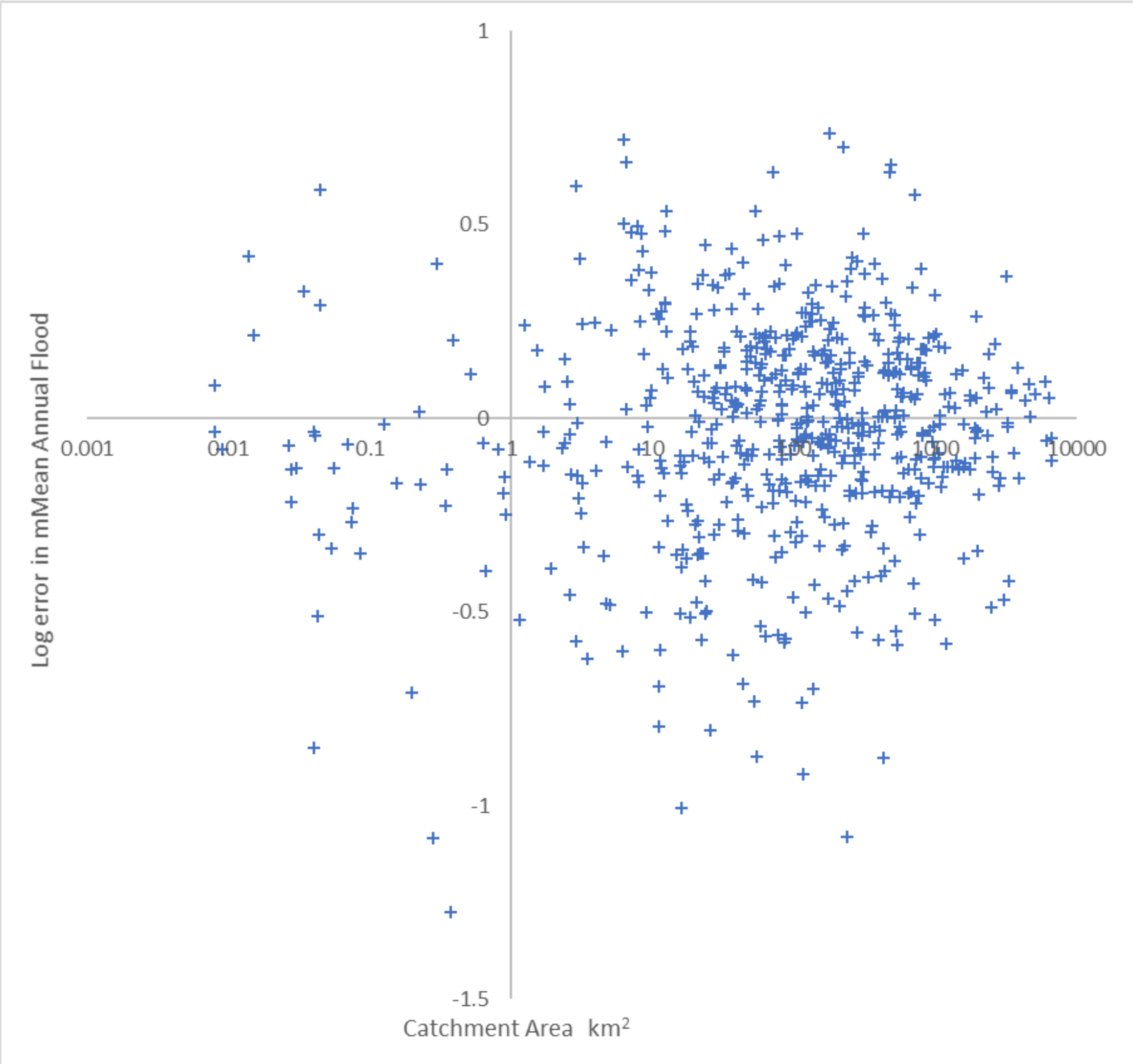

Figure 4-6: Log error vs. catchment area. A lower density of catchments smaller than 10 km^2 is evident, but bias related to area is small.

5 Preparation of a map

The chosen regression model is applied to catchment characteristics at the centroid of each digital network catchment. These values are then adjusted by the contoured surface described above, and the resulting value is the answer for mean annual flood for each reach in the network.

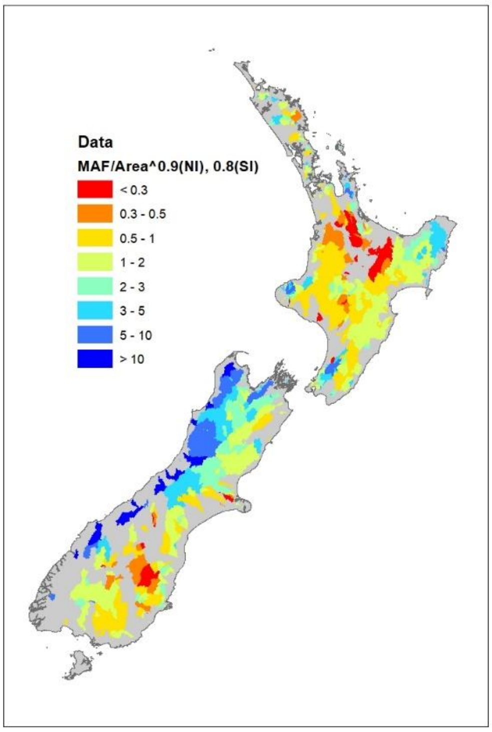

Figure 5-1: Mean annual flood data. Values are the mean annual flood divided by area raised to the power from the regression equation in each island. Both maps thus represent the remainder of the equation; for the North Island rain and geology; for the South Island just rain.

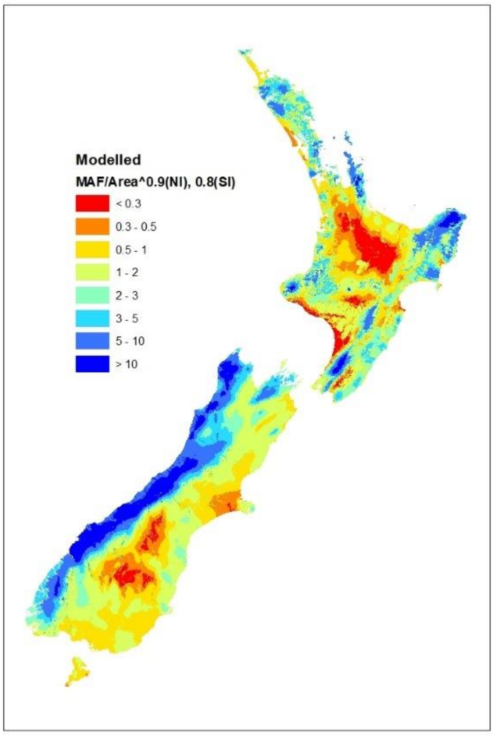

Figure 5-2: Modelled mean annual flood everywhere. Values are as per Figure 5-1. The map shows the values for each catchment at the sub-catchment polygon at the catchment outlet, so more detail is visible than in the map of data, which is plotted over the whole catchment area.

Figure 5-1 and Figure 5-2 show maps of the mean annual flood from data and model respectively. The model is applied everywhere on the national digital network version 1. The variable plotted is the mean annual flood divided by area raised to the power from the regression equation in each island (0.9 for the North Island; 0.8 for the South Island). The map thus represents the remaining terms of the regression equation (rain and geology for the North Island; rain only for the South Island), as modified by the application of contours.

Values for higher return period estimates are generated using the methods of McKerchar and Pearson (1989) for the Gumbel distribution. Calculation of the standard errors for these estimates are also covered in McKerchar and Pearson (1989). The standard error to be adopted for the q_{100} estimate (Q_{100}/MAF) is ±20%, based on analysis of q_{100} errors in Henderson and Collins (2016). Flood estimates are provided for return periods up to 1,000 years. As discussed in section 6 and illustrated in Figure 6-2 of Henderson and Collins (2016), there are regions in the east of the country where the EV2 distribution may be better used. However, a satisfactory method assigning distribution type has not been found. It is recommended that flood estimates for 50-1,000 year return periods in these regions are used with caution as they will underestimate the flood peaks.

6 Web delivery

Web delivery is by way of an Arc Online application (Figure 6-1), delivered from NIWA map servers.

Follow the link to the map application. The map can be zoomed in and when the scale gets below 1:50,000, the 1:50,000 Topomap background will appear, and the full coverage of streams down to order one will show.

A selection tool is available in the top left corner, which allows the selection of one or more reaches by dragging a rectangle on the map. All the reaches shown on the current view can be seen in the table view at the bottom of the map by clicking on the arrow. The selected reaches can be shown on their own by using the options button at the left of the table display. The selected reaches may also be exported to a csv file. Instructions are viewable by clicking the info button next to the selection tool.

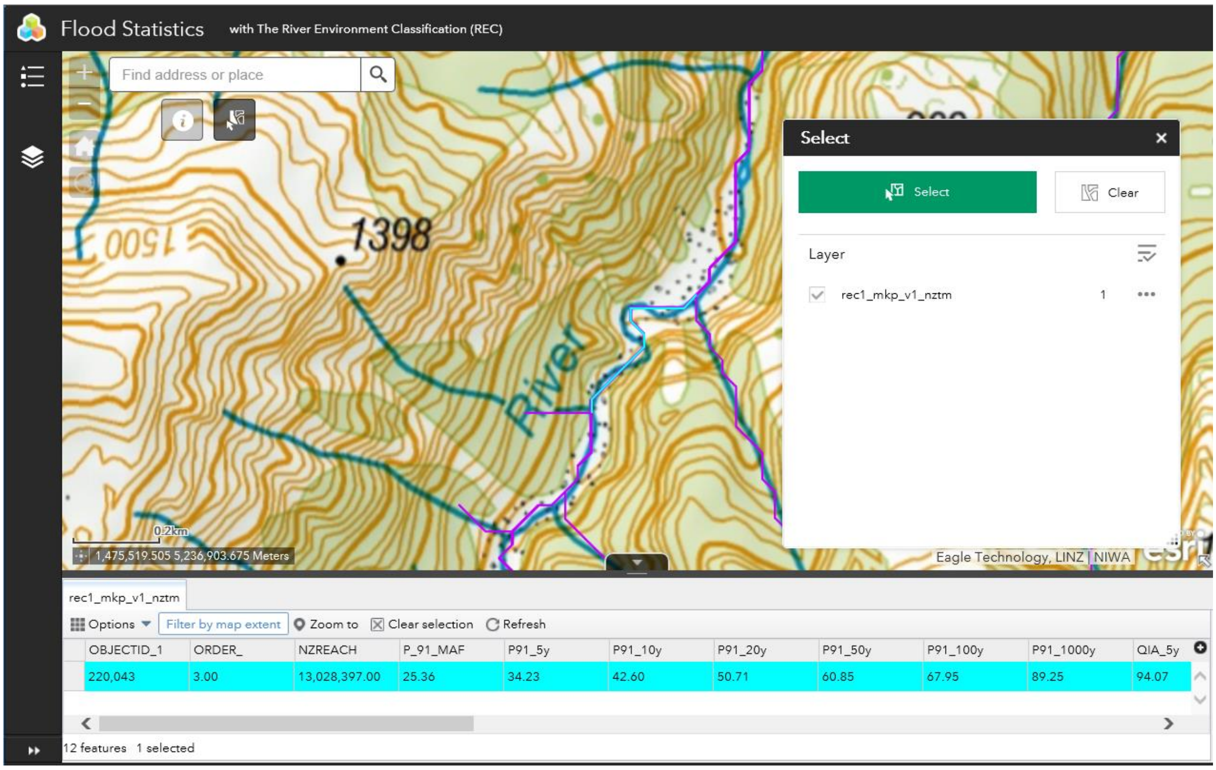

Figure 6-1: Arc Online application for stream-explorer. A reach is selected on the Anti-Crow River at Arthur's Pass, and its flood attributes are displayed at the bottom of the screen. These may be exported into a csv file. Instructions are available by clicking the info button at top left.

Attributes are listed in Table 6-1. Flood estimates and standard error estimates are given from the mean annual flood to a 1000-year return period estimate. Some caution should be exercised in using the more extreme values, especially in areas of the country such as South Canterbury, where there is some justification for use of the GEV distribution.

Table 6-1: Attributes for the stream-explorer application in Arc Online. These relate to older versions of the flood estimation, as currently displayed in stream-explorer.niwa.co.nz. All flows in cumecs.

Groups	Field Header	Meaning
	FID	
Arc Fields	Shape	
	OBJECTID_1	
	ORDER_	Stream order
Network	NZREACH	REC v1 reach number
	DEMTAREA	Upstream catchment area m^2
	P_91_MAF	Mean annual flood
	P91_5y	5-yr return period flood
	P91_10y	10-yr return period flood
Pearson (1991)	P91_20y	20-yr return period flood
	P91_50y	50-yr return period flood
	P91_100y	100-yr return period flood
	P91_1000y	1000-yr return period flood
	QIA_5y	5-yr return period I.A
	QIA_10y	10-yr return period I.A
Rational Method	QIA_20y	20-yr return period I.A
	QIA_50y	50-yr return period I.A
	QIA_100y	100-yr return period I.A
	QIA_1000y	1000-yr return period I.A
	MP89_MAF	Mean annual flood
	MP89_5y	5-yr return period flood
	MP89_10y	10-yr return period flood
	MP89_20y	20-yr return period flood
	MP89_50y	50-yr return period flood
	MP89_100y	100-yr return period flood
	MP89_1000y	1000-yr return period flood
McKerchar & Pearson (1989)	MPse_MAF	Standard error of mean annual flood
	MPse_5y	s.e. of 5-yr return period flood
	MPse_10y	s.e. of 10-yr return period flood
	MPse_20y	s.e. of 20-yr return period flood
	MPse_50y	s.e. of 50-yr return period flood
	MPse_100y	s.e. of 100-yr return period flood
	MPse_1000y	s.e. of 1000-yr return period flood

Rational method estimates can be made by multiplying the table estimate by a suitable C value (the runoff coefficient of the Rational Method equation Q = C.I.A).

For the new model, variables displayed will be like the McKerchar and Pearson variable set in Table 6-1 but with the 'MP' replaced by 'HC' and the 89 replaced by 18, for (Henderson & Collins, 2018). Return periods from mean annual to 1000 years, and standard errors for these will be the basis of the model output. A separate mapped layer will contain the flow recorders used in the regression, with statistics derived from the annual flood series at each point, the fitted Gumbel curve with return period estimates and standard errors (see Table 6-2).

Table 6-2: Flow recorder flood statistics viewable in a map layer.

Group	Field Header	Meaning
	Siteno	Tideda site number
	Name	Site name and river
	NZTM_E, NZTM_N	NZTM location map reference
Site details	Region	Regional council/unitary auth. area
	Operator	Site operator
	Funder	Site funder
	Area_km2	Catchment area
	Annual_series	Number of annual floods
	L1_mean	Average annual flood
	L2	L-moment second moment
	Lcv	L-moment CV
	T3_Lskew	L-moment third moment ratio
	T4_Lkurt	L-moment fourth moment ratio
Flood Statistics	Gumb_u	Gumbel u parameter
	Gumb_alpha	Gumbel alpha
	GEV_u	GEV u parameter
	GEV_alpha	GEV alpha parameter
	GEV_k	GEV k parameter
	GEV_z	Hosking Z for GEV
	Data 2.33y	Mean annual flood
	Data 5y	5-yr return period flood
	Data 10y	10-yr return period flood
Flood quantile estimates	Data 20y	20-yr return period flood
	Data 50y	50-yr return period flood
	Data 100y	100-yr return period flood
	Data 1000y	1000-yr return period flood
	se_Data 2.33y	Standard error of mean annual flood
	se_Data 5y	s.e. of 5-yr return period flood
	se_Data 10y	s.e. of 10-yr return period flood
Flood estimate standard errors	se_Data 20y	s.e. of 20-yr return period flood
	se_Data 50y	s.e. of 50-yr return period flood
	se_Data 100y	s.e. of 100-yr return period flood
	se_Data 1000y	s.e. of 1000-yr return period flood

A third data layer will contain the annual series as used in the analysis described in this report.

7 Discussion

We have explored methods for estimation of mean annual flood in a co-learning environment where researchers and practitioners exchanged ideas about the relative efficacy of different approaches. As a result, we have combined log-linear regression, error contouring, and leave-one-out cross validation techniques to provide a variety of models and to choose among them.

The diagram in Figure 7-1 describes the chain of processes that led to the final model.

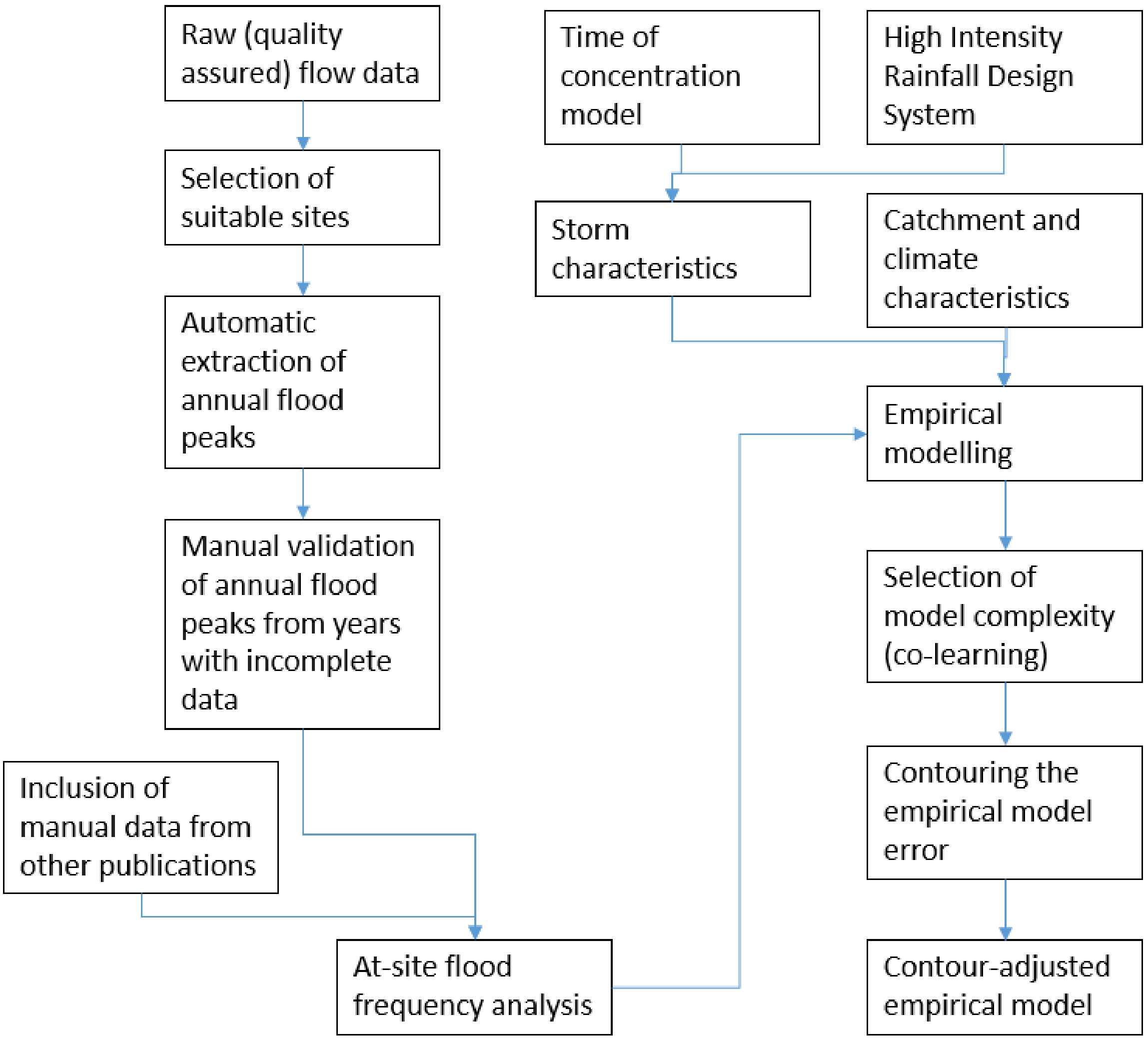

Figure 7-1: Flow chart of estimation process for mean annual flood.

In the end, improvements were made by the application of a contouring approach to the selected regression model. Model errors are still large (±61% for Al New Zealand), which suggests that more research could make improvements, however it is unlikely that substantial improvements would be made using the same method; i.e., spatial and annual average parameters regressed with a simple equation. New science is required.

That the contouring approach improved the results indicates that there were some spatial patterns that were not accounted for in the empirical model that were still important. The contours improved the fit to the dataset of 636 catchments, but the leave-one-out analyses showed that the contouring only produced a small improvement in error level over the straightforward regression approach. This demonstrates that there is some degree of regionalisation unexplained by the regressions, as shown in the contour map of Figure 4-2.

In the future, adoption of machine learning techniques, both to assess variables and to examine the possibilities of more detailed regionalisation, may yield a better result and should be the subject of research effort. However, it may be that there is little to be gained from ever more sophisticated attempts to model simple flood statistics in such an empirical fashion.

Applying Generalised Least Squares (GLS) instead of Ordinary Least Squares would also be a technical improvement upon our method. GLS accounts for serial correlation within each record as well as spatial correlation among sites. However it is unknown how much of an improvement GLS would confer.

Moving away from an empirical approach to a more process-based approach, such as using a physically-based hydrological model, may allow us to reduce errors – provided the model is good enough. Physical processes that could affect flooding, but which were not fully accounted for in the empirical approach here, may include:

- Antecedent soil moisture conditions.
- Storm type.
- Location of storms within the catchment.
- Hydrological recession characteristics.

Useful fundamental research would be to analyse historical floods on a case-by-case basis, analysing what climatic and antecedent conditions are important, and then how these factors vary with landscape characteristics. Such approaches would be enhanced by development of a long series of detailed weather model re-analysis to provide the detailed rain input necessary.

8 Summary

We have derived a new model of flood magnitude for New Zealand catchments and a re-assessment of the uncertainty inherent in the existing method that this work is intended to replace.

Flood estimation and its companion discipline, extreme rainfall intensity estimation, are critical aspects of the design of a large amount of the built infrastructure of New Zealand. The previous method for flood estimation, dating from 1989, needed updating because more extreme events have been observed in the interim, and because of the probable effects of climate change, which will increase into the future. The previous method was derived using subjective expert opinion to build the empirical model, and in this work, we specified a more objective procedure, to allow more frequent and convenient updating in the future.

The new dataset has twice as many sites and three times the annual maxima than the previous study. Nearly 58% of sites are operated by regional councils, 38% by NIWA, and the remaining 4% by other organisations. Preliminary analysis suggested no spatially coherent temporal trends in the annual series of flood maxima. The new dataset is systematically organised with inclusion of both monthly and annual maxima for each series, annotation of the years potentially affected by gaps and the expert assessment of the true impact of gaps, and inclusion of early historic annual maxima.

Workshops held in late 2015 for regional council stakeholders and for a wider audience provided useful feedback about aspects of the new model. Changes made following these are incorporated in the current model, including the division of the dataset by island.

Over the past two years we have explored ways to better estimate mean annual flood. This has involved co-learning approaches between researchers and regional council practitioners. As a result, we have chosen regression models that seek to optimise information gain without incorporating too many variables, thus remained conceptually tractable and transparent. We have investigated alternative regional approaches, and developed objective contouring methods to account for any spatial organisation of residuals. We have also adopted an unbiased error estimator, being the ratio of logs of data-based estimates of MAF and modelled MAF.

Ordinary least squares (OLS) regressions in log space have been performed on each Island individually, using a three-variable equation (area, annual precipitation and hydrogeology) in the North Island, and a two-variable equation (area and annual precipitation) in the South Island. The residuals of these equations have been contoured and a leave-one-out cross validation performed. The result of this process is an all-New Zealand record-length-weighted factorial error of 1.82 for MAF, or a relative error of ±61%. To compare with the previous method, the worst 5% of sites are removed and the factorial error reduces to 1.62. As a relative error this is ±49%, which is as good as or better than the assessed error of the previous method for 95% of all New Zealand, at ±49% to ±70%. We propose ±50% as the standard error of estimate for mean annual flood.

The regional growth curve model of the previous study was found to be still applicable as shown in the 2016 report, and we propose ±20% as the standard error of estimate for q_{100} (Q_{100}/MAF).

The MAF model is combined with the regional growth curve model to provide return period estimates from 5 to 1000 years, with standard error estimates, across all stream reaches. The results are displayed on a web-based map application, with the option to download flood statistics for selected rivers and streams. At-site annual flood series and calculated flood statistics are also displayed and may be downloaded.

The method provides a good balance of technical depth, repeatability, physical realism, and transparency, all aided by the joint application of statistical modelling and co-learning among researchers and stakeholders.

9 Acknowledgements

We would like to thank:

- Council hydrologists and NIWA field staff for original data and quality assessments.
- Dr Trevor Carey-Smith, Dr Sam Dean and Charles Pearson for report reviews.
- Ude Shankar and Julian Sykes for GIS analysis and assistance.

Funding for the work documented in this report was provided by the “Impacts and risks of weather related hazards” research programme in NIWA’s National Climate and Atmosphere Centre.

10 References

Beable, M.E., McKerchar, A.I. (1982) Regional Flood Estimation in New Zealand. *Water and Soil Technical Publication* No. 20: 132.

Henderson, R.D., Collins, D.B.G. (2016) Regional Flood Estimation Tool for New Zealand: Final Report Part 1. *NIWA Client Report*, 2016049CH for MBIE: 38.

Hutchinson, P.D. (1990) Regression Estimation of Low Flow in New Zealand. *Publication of the Hydrology Centre*, No. 22: 51.

McKerchar, A.I., Pearson, C.P. (1989) Flood Frequency in New Zealand. *Publication of the Hydrology Centre*, No. 20: 87.

Toebes, C., Palmer, B.R. (1969) Hydrological Regions of New Zealand. *Miscellaneous Hydrological Publication*, No. 4: 45.

Appendix A Variables used in regression testing

Parameter Class	Parameter	Parameter Class	Parameter
Catchment morphology	Area	Geology	Weak sedimentary
	Channel Length		Strong sedimentary
	Channel Slope		Igneous
	FWENZ slope		Metamorphic
	Catchment slope		Other (rocks)
	Catchment elevation	Land use/cover	Artificial
Climate	New precip		Unvegetated
Storm Intensity or depth from HIRDS v3	2-y 10 min Intensity		Water
	5-y 10 min Intensity		Marsh
	2-y 20 min Intensity		Grass/crop
	5-y 20 min Intensity		Shrub
	2-y 30 min Intensity		Forest
	5-y 30 min Intensity		Shrub + forest
	2-y 1 h Intensity	Soil properties	Plant rooting depth
	5-y 1 h Intensity		Plant available water
	2-y 2 h Intensity		Shallow macroporosity
	5-y 2 h Intensity		Deep macroporosity
	2-y 6 h Intensity		Bedrock
	5-y 6 h Intensity		Skeletal
	2-y 12 h Intensity		Sandy
	5-y 12 h Intensity		Loamy
	2-y 24 h Intensity		Silty
	5-y 24 h Intensity		Clayey
	2-y 48 h Intensity		Organic soil
	5-y 48 h Intensity		Not soil
	2-y 72 h Intensity	Hydro-geological Index (fraction of catchment)	Hydroindex 0
	5-y 72 h Intensity		Hydroindex 1-3
	2-y ToC Intensity		Hydroindex 4-5
	5-y ToC Intensity		Hydroindex 6-8
	2-y ToC depth	Location	Centroid Easting
	5-y ToC depth		Centroid Northing
QMAP	Hydraulic conductivity	Depth to basement	Porosity