Regional Flood Estimation Tool for New Zealand

Final Report

Prepared for EnviroLink Tools (MBIE)

August 2016

Prepared by:
Roddy Henderson
Dr Daniel Collins

For any information regarding this report please contact:

Roddy Henderson
Hydrologist
Applied Hydrology
+64-3-348 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: 2016049CH
Report date: August 2016
NIWA Project: ELF14248 and HARWXXXXX

Quality Assurance Statement		
	Reviewed by:	Charles Pearson
	Formatting checked by:	Tracy Webster
	Approved for release by:	Helen Rouse

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.

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, is in need of updating because more extreme events have been observed in the interim, and because of the probable effect of climate change, which will increase into the future. This method was derived using subjective expert opinion to build the empirical model, and a more objective procedure was another project aim, 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. No spatially coherent temporal trends were detected 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 country by island.

The uncertainty of the previous model for estimating mean annual flood is larger than originally stated, but the uncertainty of the growth curve parameter (q_{100}) is very similar. Derivation of a new model for mean annual flood proved more difficult than expected, at least to the extent that the eventual uncertainty is more than twice the previously published estimate of ~±22%.

The current dataset tested against the previous method confirms this greater uncertainty. For records used in the previous method, the root mean squared relative error (RMSRE) is ±32% for mean annual flood (Q_{mean}) and ±14% for q_{100} (Q_{100}/Q_{mean}) for 95% of the data. For flow records not used in the previous method, these errors are ±70% and ±21% respectively.

Ordinary least squares (OLS) regressions in log-log space have been performed on the entire dataset and each Island individually. These result in an all-New Zealand record-length-weighted error (RMSWRE) of ±55% for Q_{mean} with bias of 3% for 95% of the data. This is similar to the assessed error of the previous method for all New Zealand, of ±49%. The question of whether this scale of uncertainty is acceptable, and how practitioners will respond to it, remains to be resolved.

A satisfactory model to explain the flood growth behaviour (estimation of Q_T for different return periods T) was not found. The q_{100} ratio model of the previous study was found to be still applicable with small bias and standard errors of 21% or less.

Given the current performance of the new model, a web-tool originally proposed to provide access to the new estimation method has not been finalised. The next step for the project is consultation with regional council stakeholders to assist in determining future actions.

1 Introduction

This report describes progress on a co-funded project to update the national flood frequency estimation method. Funding was provided by MBIE through EnviroLink Tools (C01X1308 – Regional Flood Estimation Tool for New Zealand) and a sequence of NIWA Core Fund projects over five successive years.

Objectives

From the EnviroLink application the following describes the overall plan and demarcation of effort between funding streams:

"A high level outline of work to be done for this tool development is as follows:

Tasks

1. Project management, including meetings with council and other stakeholders.
2. Literature review of at-site frequency analysis methods.
3. At-site analysis: the assembly of annual floods series and historic values; data quality checking and dealing with gaps; generation of at-site flood statistics.
4. Literature review of methods for regional food frequency analysis.
5. Regional analysis: test various methods to predict an index flood (such as median annual or mean annual); consider choice of frequency distribution; test various methods to predict distribution parameters, including uncertainty.
6. Literature review of interpolation techniques.
7. Apply interpolation techniques to estimate flood statistics everywhere on the national digital river network.
8. Estimate a first order climate change impact on flood statistics by use of the national TopNet model (adapted from the MBIE WaterScape programme) with inputs conditioned by agreed climate change scenarios. This rainfall-runoff modelling approach is necessary because the only guidance currently available on climate change impacts is that on rainfall and temperature.
9. Provide guidance on flood volumes for use with simple design hydrographs in inundation modelling.
10. Delivery of results:
    a. Free, open-access web-based delivery of flood frequency estimates and their uncertainty on all New Zealand rivers, as well as the base-line "at-site" data (including a help file to explain use of the system).
    b. GIS data layers available for free download, giving flood frequency estimates and their uncertainty on New Zealand rivers within the scope of the tool, as well as the base-line "at-site" data.

c. A technical report which summarises the data, methods and results, provides worked examples, makes clear the intended purpose and validity of the results, and identifies limitations. This report will be available to download or on CD/DVD.

d. Instigate a series of workshops / symposia to launch the tool to councils and engineering professionals.

It is generally proposed that the EnviroLink Tools Grant funding be used for items 3, 7, 8, 10 and part of item 1, and that the supplementary funding provided by NIWA is predominantly for items 2, 4, 5, 6 and 9. This achieves a distinction between development and application of methods.”

2 Data and pre-processing

Data for the study are derived from time-series archives held by NIWA and regional councils and others, and from various geographic data sources for catchment characteristics related to climate and land resource variability.

2.1 Time-series data

The method of flood frequency analysis described here is built on observational records of river flow collected by a number of organisations.

2.1.1 Data selection criteria and quality

The NIWA Site Information Management System (SIMS) contains meta-data for 2349 sites that were identified as having the potential to have generated river flow data. These were examined to eliminate unlikely candidates in order that data providers would only need to provide further assessment of the minimum number of sites. Criteria applied fell into two distinct categories:

- Objective criteria, such as length of record and the presence in the catchment of significant lakes and other storage;
- Expert criteria, such as the degree of artificial manipulation of flows that might affect flood size, the goodness of the top end of the stage to flow rating curves at the site, and any other site considerations that were understood by the data collectors.

Application of the objective criteria and some of the subjective where the necessary information was held by NIWA or known to NIWA staff resulted in a list of 1023 sites that was sent to recording authorities for their inspection and comments regarding data quality. Information on these 1023 sites were recorded in an Excel spreadsheet master list.

In the previous work, and especially the 1989 study, the rating curves and gauging records were assessed by the project team. In the current study questions of data suitability were asked of the data providers, being for the most part experienced regional council hydrologists or NIWA hydrology field team leaders. Their comments are preserved in the master list and will be useful in future for enhancing meta-data about flow records and their suitability for use in a variety of studies.

2.1.2 Data sources

Many of the time-series used are either currently shared funding across a number of organisations or else have changed funding between organisations over time. This means it is difficult to be categorical about the funding composition of the data used. However, we present below an estimate of the numbers of sites provided by each of the major categories once expert rejection criteria had been applied:

- NIWA – sites funded by NIWA or by the various incarnations of the Public Good Science Fund and its successors - 246 sites
- Regional councils – 369 sites
- Client data – NIWA clients' exclusive data - 19 sites; other organisations' data – 6 sites.

This gives a total of 640 sites initially selected from meta-data lists held in NIWA SIMS for the 1023 sites that were initially expected to have suitable flood flow data. This is nearly twice as many as were available for the 1989 study.

Figure 2-1 shows the distribution of catchments across New Zealand between the two studies. The figure also shows record lengths. Both the spatial coverage and record lengths are enhanced in the current work so that whereas the previous study used slightly more than 6,000 years of record across the gauging locations, the current dataset has more than 18,000 - a threefold increase.

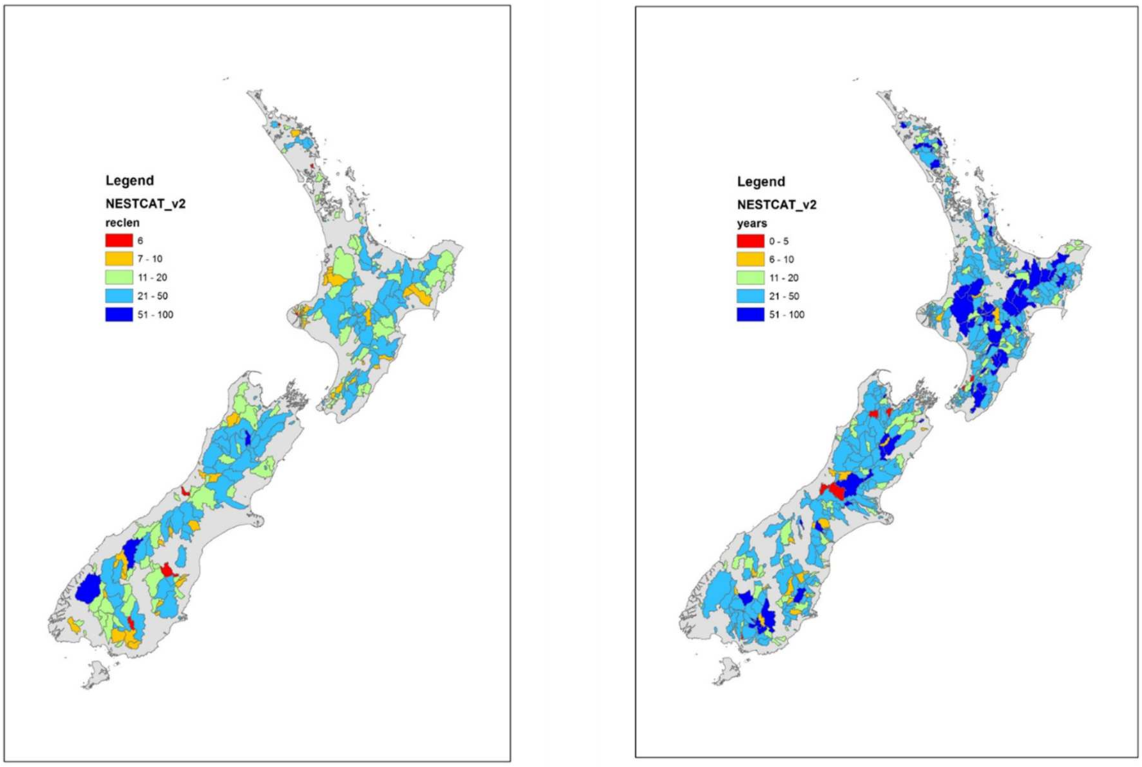

Figure 2-1: Map of catchments included in 1989 study (left) and current study (right). Legends indicate the length of record in years.

2.1.3 Combination sites

Some flow records in the assembled dataset are in fact flow data collected from nearby points on the same river over different periods of time. These 'combination' sites have been identified from meta-data held at NIWA (originally from comments in Walter (2000) and previous versions). Generally a new site was marked as "Replaces …" and the older site was marked as "Replaced by …". The result of this 'combination' process was 32 longer series derived from 68 shorter records. Most combinations are of only two time-series, but four are of three successive time-series. The resulting flow series are named and numbered for the most recent flow recording site in each case.

When catchment areas between one site and another within the same sequence were different, as a result of shifts in recorder location, the older flow records were scaled by the ratio of catchment areas raised to the power of 0.8. This factor was chosen as an approximate scaling for flood magnitude from catchment area based on the factor used in McKerchar and Pearson (1989). Sites for which the change in area was larger than 25% were not incorporated into a combination series but treated as separate time-series.

2.1.4 Early manual annual series

In previous studies (Beable and McKerchar 1982; McKerchar and Pearson 1989) there were a number of time-series where the annual maxima derived from computer records were enhanced by the addition of earlier annual maxima. These were generally derived from manual processing of chart records that had not been captured on computer. Fortunately, these annual maxima have been documented in the appendices of the two previous reports, and thus have been available to add into the dataset assembled for this study. In the two studies mentioned above, 550 and 590 station years respectively were added from this source. Instrumental data available to those studies were 3212 and 6358 station years respectively, and thus the addition of early data contributed 17% and 9% respectively.

Most of the data supplied for the current study did not include reference to these historical manual annual records. Where this is the case we have added them using the data from the earlier work. If there was evidence of some change to ratings between studies, from consideration of ratios of flood peaks, we have applied these to the historic data. Annual values were added for 637 station years, to an instrumental record total of slightly over 18,000 station years, an increase of 4%. The data and its source are inserted into the spreadsheet database (see 2.2 below) in the correct time sequence. In most cases the only timing information available is the year of the historic annual maximum.

2.1.5 Temporally isolated historic flood data

In addition to the early annual maxima derived from manual inspection of charts not yet digitised, there are historic observations of large floods often dating back to the early days of settlement of an area. We call these 'temporally isolated floods' to distinguish them from those described in the previous section. Techniques to incorporate these events into a frequency analysis are available, notably the work of Stedinger in the mid-1980s (Stedinger, Vogel et al. 1993). However such techniques require provision of assumptions about the years without flood observations, and these assumptions are not susceptible to objective estimation techniques.

In the context of this work we do not recommend that these historic events be used in the regional assessment of flood magnitude or growth but recommend that such events, where known, are used to provide further detail for the at-site frequency analysis procedure that is used to blend recorded data with regional estimates.

Some of these are documented in the appendices to Beable and McKerchar (1982) and the descriptions often include assessment about the ranking of the events in the longer term context.

2.2 Extraction of flood peaks

All time-series data were stored in Tideda files (Rodgers and Thompson 1992) and the data sampling procedures of Tideda process EVENTS (translated into Matlab programs to allow greater automation) were used to extract the largest value for each month of record, with or without data gaps. Instantaneous flood peaks were extracted, and the annual period of analysis was set as 1 January to 31 December (in contrast with the typical water year for New Zealand which runs from winter to winter). Data were written to an Excel spreadsheet file, with one sheet for each site analysed. As well as the site details of site number, name, source file, and catchment area, the derived annual series was also written to the spreadsheet. The monthly data were not used further but provide a basis for possible future research into seasonal effects on floods in New Zealand.

2.3 Dealing with gaps

While flow records span many years they are not always complete. This may occur for several reasons, often because of damage or failure of the equipment during floods, particularly large floods. The presence of these data gaps means that the maximum record flood may not in actual fact have been the largest flood to occur. However, to blankly disregard all incomplete years of data would effectively sacrifice a lot of legitimate information.

In order to maximise the amount of data used in the analysis a semi-automatic procedure was developed that assessed the possibility that gaps contained the largest flood of the year. For each year of each site with missing data the data for that year were compared to those from neighbouring sites. Three neighbours were selected for the comparison: the closest catchments (centroid to centroid) that satisfied two other criteria; firstly that their catchment area was within two orders of magnitude of the target catchment's area, and secondly that they must contain some data for that year. Visual inspection allowed an operator to decide whether it was plausible that the gap in the target site's data included the largest flood of that year or not, based on the timing of the largest floods in the neighbouring records. The decision made is preserved in the annual maxima dataset, an excel spreadsheet file with a sheet for each site. Of the sites assessed, there were 19,189 potential years of annual maxima, based on start and end times of the record. If gaps were allowed to result in exclusion of a year, there would have been only 13,710 years of data. The gap assessment process described above resulted in a final total of 18,190 years of annual maxima, thus adding 4,480 years of a possible 5,479 years to the dataset.

Once these decisions had been made some re-assessment of site suitability is necessary, as some sites with a marginal number of station years have too few years when gaps are accounted for. The final total of available data is approximately 18,000 station years.

2.4 Derivation of at-site statistics

The annual series of floods from the spreadsheet of extracted monthly and annual maxima (enhanced with manual annual maxima where available) for each site was processed using a Matlab version of the Tideda process EVAN (Event Analysis) to obtain at-site flood statistics. These included:

- years: the number of years of data derived from the recorded start and finish of the time-series;
- used years: the percentage of the years that are useable after gap checking;
- statistics of the annual series, derived using the L-moments concepts of Hosking and Wallis (1993);
  - L1 (Q_{mean}); L2; L_{cv}; T3 (L_{skew}); T4 ($L_{kurtosis}$)
- statistics of the biennial series, being the largest flood value from each pair of years, based on concepts introduced in McKerchar and Pearson (1989)
  - Biennial years; L1 (Q_{mean}); L2; L_{cv}; T3 (L_{skew}); T4 ($L_{kurtosis}$);

- parameters of a Gumbel (EV1) distribution fitted to the annual series by the method of L-moments
  - Annual Gumbel u; alpha
- parameters of a Gumbel (EV1) distribution fitted to the biennial series by the method of L-moments
  - Biennial Gumbel u; alpha
- parameters of a GEV distribution fitted to the annual series by the method of L-moments
  - Annual GEV u; alpha; k; z
- parameters of a GEV distribution fitted to the biennial series by the method of L-moments
  - Biennal GEV u; alpha; k; Hosking Z
- statistics of the fitted frequency distribution
  - Gumbel Annual Q_{100} (1% aep flood); $Q_{2.33}$ (average annual flood); $Q_{2.0}$ (median annual flood)
  - Gumbel Biennial Q_{100}; $Q_{2.33}$; $Q_{2.0}$
  - GEV Annual Q_{100}; $Q_{2.33}$ (for comparison but not the average annual); $Q_{2.0}$ (median annual flood)
  - GEV Biennial Q_{100}; $Q_{2.33}$; $Q_{2.0}$
- statistics for selected distribution based on a priori limits around the Hosking Z factor
  - Selected Q_{100}; $Q_{2.33}$; $Q_{2.0}$; Q_{mean}; Q_{median}; $Q_{mean}/A^{0.8}$; Q_{100}/Q_{mean}; Q_{100}/Q_{median}.

As an additional check of data quality and suitability $Q_{mean}/A^{0.8}$ was plotted for each site, allowing preliminary potential outliers to be identified. These were then investigated as to whether there was an error in the units field held in the SIMS database, or whether there was something else wrong with the data. Errors in units were suspected if outliers were 3 or 6 orders of magnitude above or below the main cluster of points. SIMS database errors or the list of included sites were amended accordingly.

This extended data processing procedure showed up a number of problems with the general area of flow site meta-data. These problems have largely been dealt with but some remain. Notable among these are issues about accurate recording of the time extent of time-series data holdings, the agency actually holding the data, and the files in which it is stored. To an extent these problems arise in other research areas where bulk time-series data are sought by scientists (e.g., the current EnviroLink funded revision of HIRDS (High Intensity Rainfall Design System). Considerable time was spent in reconciling mismatches of data holdings or details of site meta-data.

2.5 Trend detection

Prior to the full frequency analysis it is prudent to ascertain whether the individual data sets contain any trends. The presence of a compelling trend would mean that past flood statistics are not completely indicative of future flood statistics, making their use in a flood frequency model problematic.

To test for potential trends we applied the Mann Kendall ranking test (Helsel and Hirsch, 2002) and adapted it to incomplete time-series. This is a non-parametric test (i.e., it does not require values to be normally distributed) that detects the presence of a monotonic increase or decrease in the data. This is applied to each record separately. A p-value of 5% is considered the threshold for significance, and the time-series must be long enough and dense enough (i.e., without too many gaps) for the method to be applicable.

2.6 Catchment climate and physiographic data

The spatial variables are mapped onto each gauging site catchment based on the River Environment Classification (REC), version 1. This version of the REC has a number of known errors affecting river network connectivity, particularly in gently sloping land. The resulting catchment properties will reflect these errors. One such catchment is the Avon River site (site number 66602); its REC-based area is 119 km^2 while its actual area is 63.5 km^2.

2.6.1 Rock type

Information on rock type was extracted from the “toprock” variable in the New Zealand Land Resources Inventory (NZLRI). Rock type classifications for each GIS polygon in the LRI were aggregated into one of the following groups:

- surficial
- sedimentary (weakly indurated)
- sedimentary (strongly indurated)
- igneous
- metamorphic
- other.

The fractional coverage of each of these groups was then calculated for each flow recorder catchment.

2.6.2 Soil properties

Information in soil properties was extracted from the NZLRI. Numerical variables included:

- potential rooting depth (mid);
- plant available water (mid);
- shallow macroporosity (mid); and
- deep macroporosity (mid).

An additional categorical variable is included – soil type. This is recorded as the fractional coverage within the catchment, and the categories include:

- sandy
- loamy
- silty
- clayey
- organic soil
- skeletal
- bedrock.

2.6.3 Land use

Information on land use was extracted from the Land Cover Database (LCDB), specifically the 2012 version. The land use characteristics associated with each gauging site catchment is the fraction of the catchment covered by different groups of land use as follows:

- forest
- shrub
- forest and shrub combined
- grass and crops combined
- marsh
- water
- unvegetated
- artificial (paved areas etc.).

2.6.4 Climate

Climate data were obtained from several sources. Spatial variables include:

- mean annual precipitation (REC)
- mean annual precipitation from Freshwater Environments of New Zealand (FWENZ)
- mean annual temperature (FWENZ)
- mean annual evaporation (FWENZ).

2.6.5 Weather

Variables related to the storms that cause floods were extracted from the High Intensity Rainfall Database (HIRDS):

- storm intensities extracted from HIRDS for 2- and 5-year return periods and with storm durations ranging from 10 minutes to 72 hours as well as the time of concentration of the catchment;
- storm depths extracted from HIRDS for 2- and 5-year return periods and with storm durations equal to the time of concentration.

2.6.6 Topography

Topographic variables included:

- catchment area
- mean catchment elevation
- channel length (longest channel)
- mean channel slope (longest channel)
- mean catchment slope (REC)
- FWENZ slope

2.6.7 Location data

- Centroid easting
- Centroid northing

2.6.8 Hutchinson Hydrogeological Index

An aggregate measure of the hydrological responsiveness of the geology is represented by the Hutchinson Hydrogeological Index (Hutchinson 1990) representing the combined attributes of water storage capacity and transmissivity. This is represented in the data as fractional coverage within the catchment of any of four groups of indices:

- 0 - water
- 1-3 - low
- 4-5 - medium
- 6-8 – high.

3 Methods for estimation of flood statistics

3.1 Review of methods

As a first step we undertook a literature review of technical material and a survey of current international practice in flood estimation both at a site and for a region. In doing this, we bore in mind that the project had relatively limited time and funding to be completed (compared with previous estimates of the scale of work to produce a working flood estimation model). Consequently there is a need to employ proven techniques and accessible software and not to embark on developmental work of uncertain scope.

The product required is a method of estimating peak flood magnitude, Q, of a given return period, T, at a specified site for which there may or may not be annual maxima data available. The user must be able to employ the method on a PC or equivalent without having to input any information other than T at a location somewhere on the New Zealand stream network.

3.1.1 Current practice

Here we outline current practice for estimating mean annual flood, Q_{mean}, and dimensionless flood magnitude $q = Q_T/Q_{mean}$ at a site and throughout a region and recommend a general approach for use in New Zealand.

At site methods

In the USA and Australia the LP3 (Log-Pearson 3) distribution is used to model annual maxima flood peak data at a site (US Water Resources Council 1981;Institution of Engineers Australia 2012). In the UK, Austria, Italy and NZ the GEV distribution is employed (Robson and Reed 1999; Beable and McKerchar 1982;McKerchar and Pearson 1989;Bocchiola, De Michele et al. 2003).

Regional methods

Almost all practitioners use an index flood approach, or a variant thereof, which involves estimation of Q_{mean} and q for any location in a region using various techniques. For example in the UK, Q_{mean} (actually the median is used) is estimated from flood data, or from catchment descriptors or from data transfer from analogue basins and q from growth curves derived from pooling groups. In the USA and Australia, maps of regional skew are employed. In New Zealand, McKerchar and Pearson (1989) used spatial contours of $Q_{mean}/Area^{0.8}$ and q_{100}. Other variants include the employment of hierarchical regions (Gabriele and Arnell 1991), fractional membership (Wiltshire 1986), clusters (Acreman and Sinclair 1986) and regions of influence (Burn 1990).

3.1.2 General approach

At site

As the GEV distribution has some theoretical justification and as it has been successfully employed in New Zealand previously (McKerchar and Pearson 1989) it ought to be employed where fitting the data is done using L-moments which have been shown to be the best method currently available (Hosking 1990;Hosking and Wallis 1997;Martins and Stedinger 2000).

Regional

The index flood method is more or less standard and should be used with Q_{mean} as the index variable on the grounds of familiarity for New Zealand users and the ability to relate to previous New Zealand work which has employed Q_{mean} without exception.

3.1.3 Specific methods for estimation of Q_{mean} and q_T

Q_{mean} estimation

At a site without data the main methods of estimating Q_{mean} are from catchment descriptors, by data transfer, from simulation modelling and from channel dimensions. Owing to the limited number of basins having rainfall and flow recording sites in some areas of New Zealand, the data transfer and simulation modelling approaches may be difficult to apply everywhere. Moreover, limited success was achieved in New Zealand by Mosley (1979) using the method of channel dimensions. This leaves catchment descriptors which has been employed by Beable and McKerchar (1982) and Pearson and McKerchar (1991). A technique which is dimensionally correct would be to calculate Q_{mean}/IA at each site with data where I is an index of rainfall intensity and A is basin area. A more sophisticated tack would be to raise I and A to different powers as determined from the data set of annual maxima. Then, to calculate Q_{mean} at a given location without data, the usual approach is to divide the area of concern into so-called homogeneous regions (Hosking and Wallis 1993) and provide a descriptor equation for each region. The user then has to supply relevant values of I and A to find Q_{mean}. Now, there are numerous problems with defining regions - degree of homogeneity to be adopted and complicated boundary issues especially with small catchments. An alternative which is presently gaining some prominence in the literature is to use a region of influence approach and to employ a kriging technique on a stream network where the value of Q_{mean}/IA say is calculated based on weighted values at surrounding sites with data, where the influence of the value at a site is related to distance from the site of interest (Skøien, Merz et al. 2006;Castigloni, Castellarin et al. 2011). HIRDS uses a similar approach but in this case the 'distance' measure takes the river network structure into account.

q_T estimation

The usual method is to define homogeneous regions as with Q_{mean} and then to calculate growth curves of q_T based on site data. To avoid regionalisation kriging could again be employed on the stream network values of q_T where T is specified as with Q_{mean} above.

3.1.4 Other methods considered

A number of statistical techniques were considered for development of a new flood magnitude model:

- ordinary least squares regression (OLS)
- generalised least squares regression (GLS)
- regression tree
- random forests (RF)
- symbolic regression
- cluster analysis.

3.1.5 Recommendations based on review of methods

- For at site analysis and calculation of Q_{mean}, q_T use the GEV frequency distribution fitted by L-moments.
- For prediction of Q_{mean} at a given location use kriging of the parameter Q_{mean}/IA on the stream network
- For prediction of q_T at a given location, use kriging of this parameter for a specified value of T.

Two potential difficulties arise with this approach.

Firstly the assumption that only network and other spatial proximities are relevant when assessing the flood magnitude. Flood response in New Zealand is subject to considerable variation over relatively short distances due to two main factors: steep gradients of rainfall intensity, and large changes in catchment properties especially where the presence of volcanically derived soils and geology are dominant. The first of these complicates the assessment of the rainfall parameter of relevance (see below); the second means discontinuities in properties that are not well handled by kriging techniques.

Secondly the difficulty of finding suitable rainfall intensity parameters. Rainfall intensity estimates from HIRDS are only at-point estimates. There is no accepted method for area reduction factors in New Zealand at present.

For these reasons we chose the simpler method of OLS in order to explore the properties of the flood magnitude dataset, and to estimate mean annual flood.

3.2 Calculation of model error

Model error is measured by the statistic we introduce as the Root Mean Squared Weighted Relative Error (RMSWRE). This is based on the Root Mean Squared Error (RMSE) but with two modifications to reflect the circumstances of the data. Instead of basing the error calculation on model error directly (i.e., $Q_{mod} - Q_{obs}$) we first use the relative error (i.e., $(Q_{mod} - Q_{obs})/Q_{obs}$). This prevents the error calculation from being unduly influenced by catchments with large flows and consequently large floods, while also casting the error in meaningful terms (i.e., 15% error); this is also the basis of the error calculation used in the 1989 study. In addition to this, the sites' relative errors are weighted by record length. This prioritises site statistics that are derived from long records, which are more likely to reflect the actual hydrological variability of a catchment. The resulting expression for RMSWRE is:

$$RMSWRE = \sqrt{\frac{\sum_i \frac{RecordLength_i}{\sum_j RecordLength_j}\left(\frac{Q_{mod,i} - Q_{obs,i}}{Q_{obs,i}}\right)^2}{n}}$$

where Q_{mod} is the modelled mean annual flood (m3/s), Q_{obs} is the mean annual flood calculated from the observational record (m^3/s), and *n* is the number of sites used. Also reported for illustrative purposes, but not used in model selection, is the weighted relative error (WRE):

$$WRE = \frac{\sum_i \frac{RecordLength_i}{\sum_j RecordLength_j}\left(\frac{Q_{mod,i} - Q_{obs,i}}{Q_{obs,i}}\right)}{n}$$

During model selection, competing models are distinguished by the RMSWRE based on all sites, with the objective function being the minimisation of RMSWRE subject to constraints. Following the 1989 study, reporting of the final model error excludes the worst 5% of sites, where these sites are identified based on their individual contribution to the RMSWRE.

4 Trends

Of the 639 flow records being analysed, 613 were of sufficient length to be suitable for trend detection using the Mann-Kendall ranking test. Of these, 63 had p-values below the nominal significance threshold of 5%, suggesting a trend in the data. Because of the nature of the test, however, this does not indicate that the other data sets are trend-free, just that no trend can be detected. Conversely, a low p-value does not guarantee that a trend does indeed exist.

Of the sites with trends, 43 were deemed to exhibit a positive trend and 20 a negative trend. The presence and absence of trends are displayed in Figure 4-1. Visual inspection of these results suggests no spatial clustering of trends, and hence no compelling evidence of a coherent monotonic change in mean annual floods as might be expected as a simple hypothesis under climate change. Furthermore, many of these trends were detected over durations that are too short (e.g., 10 years) to encapsulate a thorough sample of natural variability, including large-scale climatic phenomena such as the Interdecadal Pacific Oscillation (IPO), meaning that some of the detected trends, while they may be statistically robust, should not be interpreted as long-term hydrological trends. No further analysis of potential trends was undertaken.

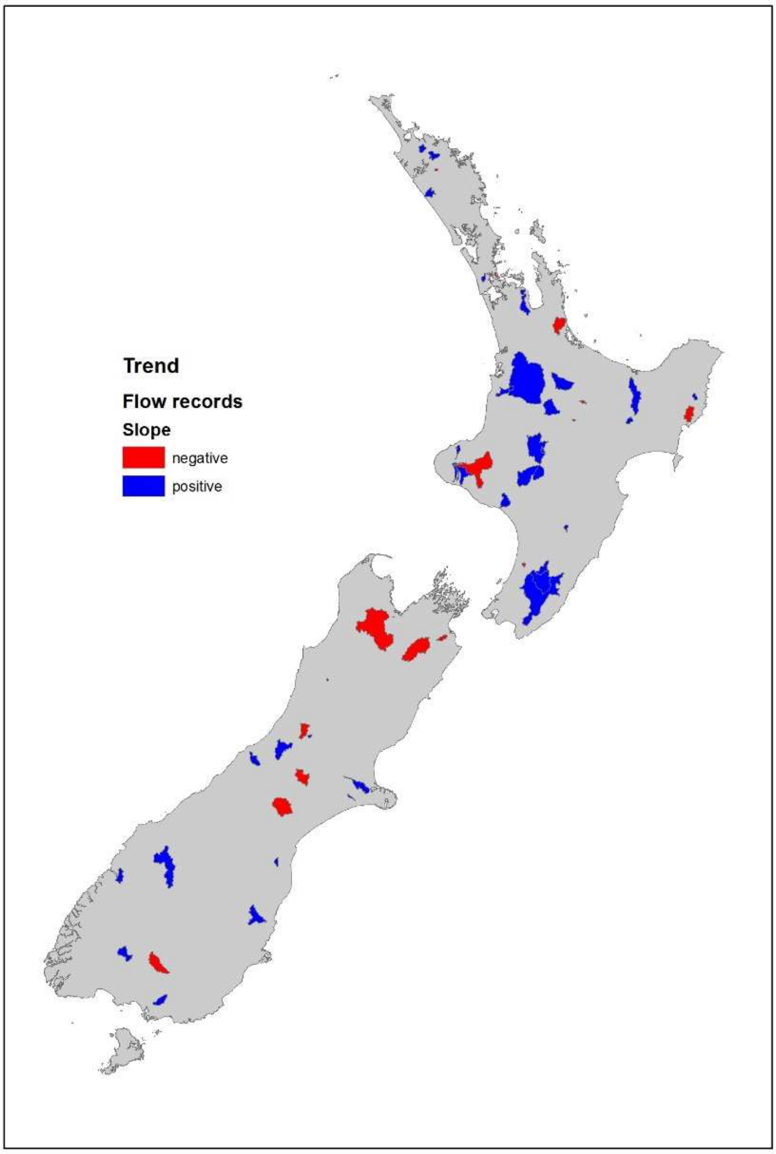

Figure 4-1: Flow records, identified by their upstream catchment, identified as having a monotonic trend (negative or positive).

5 Estimation of mean annual flood

When we turn to developing a model of the mean annual flood, catchment area is the only completely obvious variable of significance. In the 1989 study the dependence of flood size on area was established in the first phase of model development; remaining residual errors were then accounted for by a manually drawn contour map. In the present work, area was retained as a variable in all regressions to ensure it transparent inclusion; if other variables had an areal component, exponents would be correspondingly adjusted by the regression routines. However, it is instructive to examine the area to flood size relationship to see if there is any significant difference between datasets. One hypothesis is that this relationship is inherently more variable in the new enhanced dataset, and this leads to greater difficulty in establishing a good model.

Table 5-1: Bias and RMSRE of a simple area model of mean annual flood. An equation of the form Q_{mean} = mult*Area^{exp} is used.

	Case	N	mult	exp	r^2	bias	Rel. weighted bias	RMSRE	RMSWRE
1989 data	All Data	342	2.147	0.792	0.81	58%	68%	204%	226%
	best 95%	325				23%	25%	100%	99%
Our Data	All Data	648	1.595	0.835	0.82	74%	58%	326%	266%
	best 95%	616				24%	21%	107%	102%

While there is clearly a greater bias and RMSRE in the current full dataset, when outliers (the top 5%) are removed there is little difference in uncertainty despite having double the number of data values. The record length-weighted bias and standard error are slightly smaller than the unweighted errors for the current dataset, and slightly larger for the 1989 dataset.

Following McKerchar and Pearson (1989) we map the mean annual flood values divided by catchment area raised to the power of 0.8 (Figure 5-1). The patterns are similar to those from previous datasets and represent the dominance of rain as a major variable, with a secondary effect of reduced flood intensity evident in the volcanic area of the North Island.

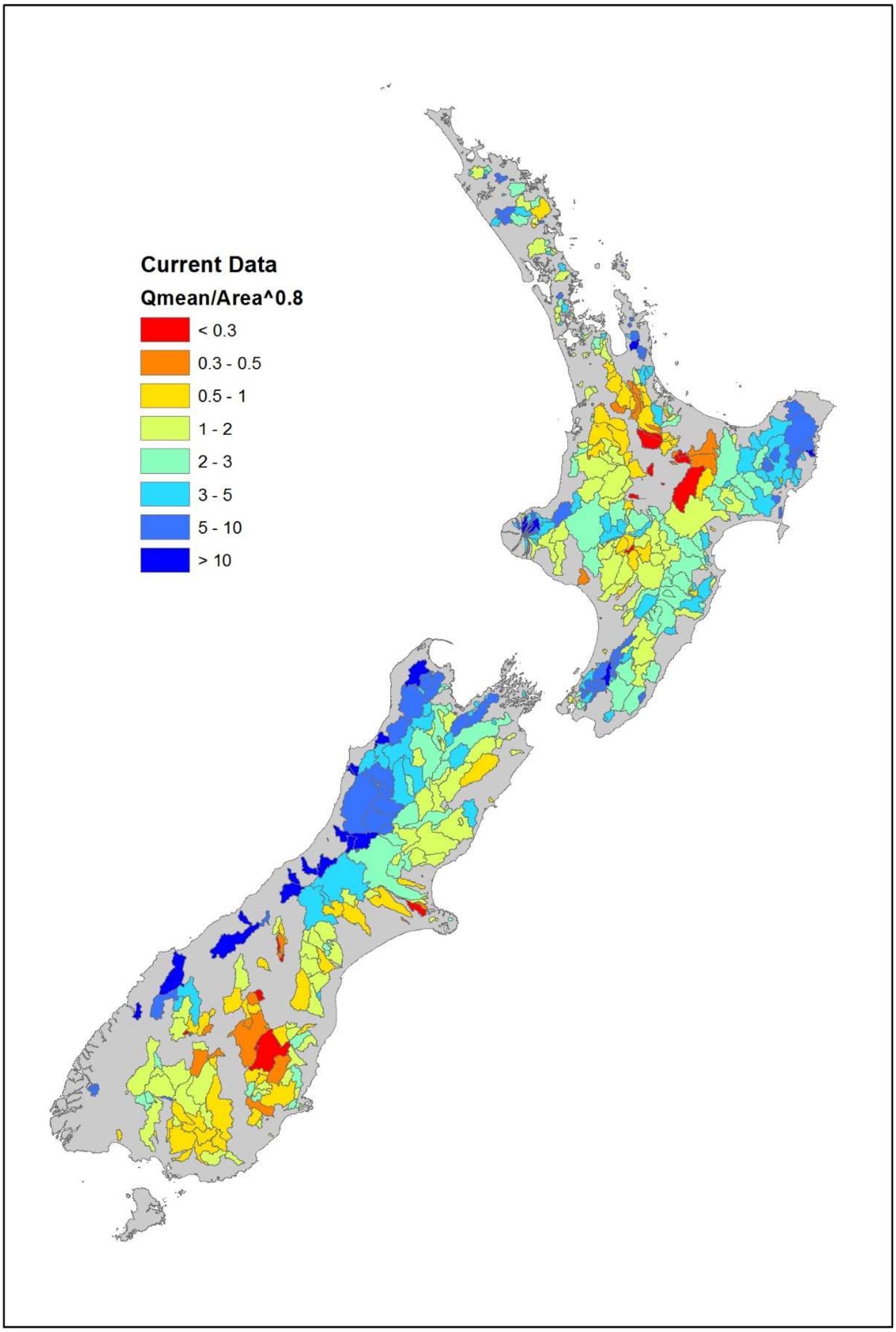

Figure 5-1: Map of mean annual flood divided by catchment area raised to the power of 0.8. Grey areas do not have monitored catchments suitable for this study.

5.1 Model selection

The method chosen to develop the revised model for mean annual was OLS, exhaustively searching for the most informative spatial variables subject to a set of constraints on the set of variables. These constraints were as follows:

- the variables should have physically plausible relevance to flood generation;
- catchment area must be included (see above);
- at least one precipitation-related variable must be included, but no more than two. If two are included, then one must be a mean annual precipitation variable and the other either storm duration or storm intensity;
- up to three additional variables will be accommodated, beyond area and precipitation.

Using all records, including potential outliers, all possible combinations of the above constraints were regressed against the mean annual flood in log-log space using OLS. This amounted to 474,909 alternative models. The final model was selected from among the combinations that resulted in the lowest RMSWRE, which in reality had negligible difference in error among them, with one last preference: that the additional spatial variables avoided duplicating one another in terms of general physiographic qualities if possible. In other words, a model would be avoided if it had two variables relating to geology, or two to soils, for example, although in practice this may not be possible. More sophisticated searching methods could have been used, such as piece-wise regression or Principle Component Analysis, but the exhaustive approach was adopted for two reasons:

i. The alternative search methods may not result in the globally optimum model fit, only a very good model fit; and

ii. The computational burden of the exhaustive search method was not prohibitively high.

Model selection was initially carried out with the North Island and South Island sites treated together. Following discussions with the regional council stakeholders, models were developed for each island separately. These discussions also led to the elimination of a particular state variable as a predictor in the search algorithm – channel distance to coast. It was initially included as a readily available proxy for storm track, and was consistently present in the best model fits for the whole country. However its physical relevance to flood generation was considered too vague and so the variable was dropped from subsequent model development.

RMSWRE values for each island and for each of the variable combinations are displayed in Figure 5-2. The very first data point on the left is the model using just catchment area. The clusters and step changes along the horizontal axis represent the inclusion or exclusion of other variables. On the whole, a better model fit is possible for the South Island than the North Island, and there is also less volatility in model fitness across model options for the South Island. This supports the decision to separate the two islands, as their climatic and geological characteristics are different enough to lead to different flood generation characteristics.

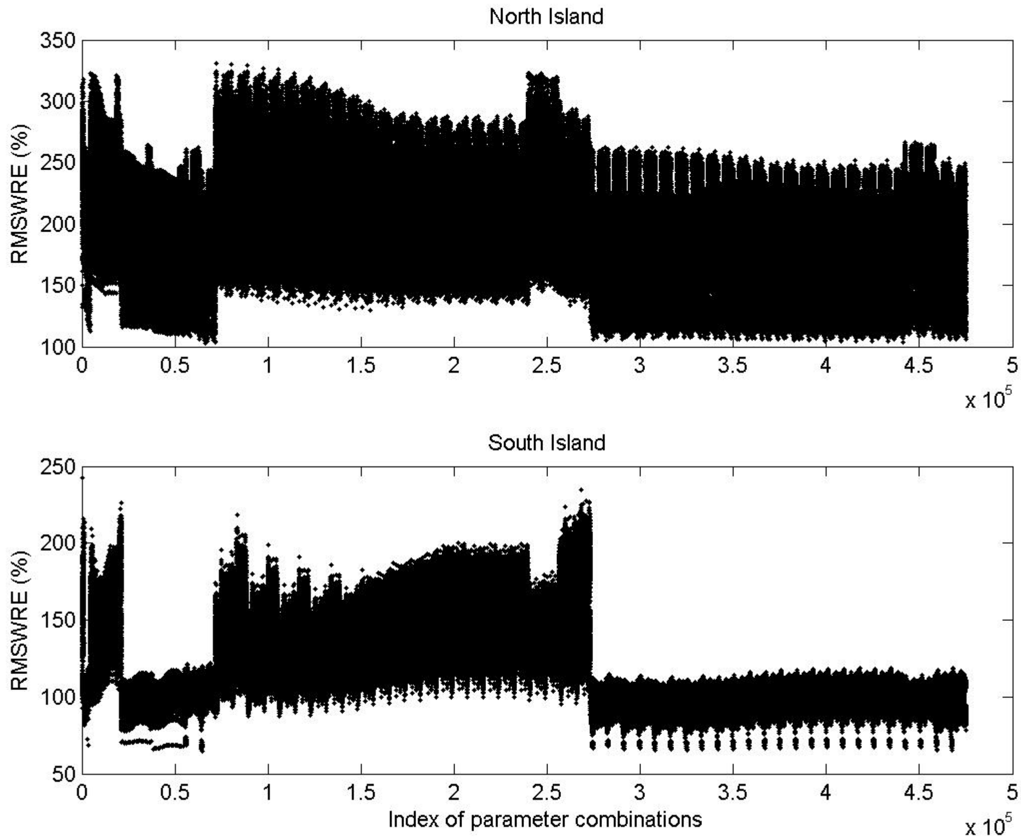

Figure 5-2: RMSWRE values for each variable combination tested. Top figure shows North Island, bottom figure South Island.

5.2 The mean annual flood model

The final model chosen for the North Island using OLS and applied exhaustively to all possible combinations of variables subject to the criteria stated above, is:

$Q = 2.2 \times 10^{-8} \, . \, A^{0.88} \, . \, P^{2.57} \, . \, HI_{4\text{-}5}^{0.14} \, . \, HI_{6\text{-}8}^{-0.25} \, . \, z^{-0.19}$

where A is the catchment area (km^2), P is the FWENZ-based mean annual precipitation (mm), $HI_{4\text{-}5}$ is the catchment fraction associated with Hutchinson's hydrological indices 4-5 (Hutchinson 1990), $HI_{6\text{-}8}$ is the catchment fraction associated with Hutchinson's hydrological indices 6-8, and z is mean catchment elevation (m).

For the South Island the model is:

$Q = 2.4 \times 10^{-4} \, . \, A^{0.88} \, . \, P^{1.41} \, . \, S^{0.40} \, . \, HI_{6\text{-}8}^{-0.13} . \, z^{-0.54}$

where A is the catchment area (km^2), P is the FWENZ-based mean annual precipitation (mm), S is the FWENZ-based slope, $HI_{6\text{-}8}$ is the catchment fraction associated with Hutchinson's hydrological index 6-8, and z is mean catchment elevation (m). These are plotted in Figure 5-3 separately, and in Figure 5-4 (right hand graph) for the country as a whole. It is instructive to see how the multi-variable models improve upon models using just catchment area (Figure 5-4). It is also instructive to see how removal of the 5% worst sites, for each island, produces a better model fit (Figure 5-5).

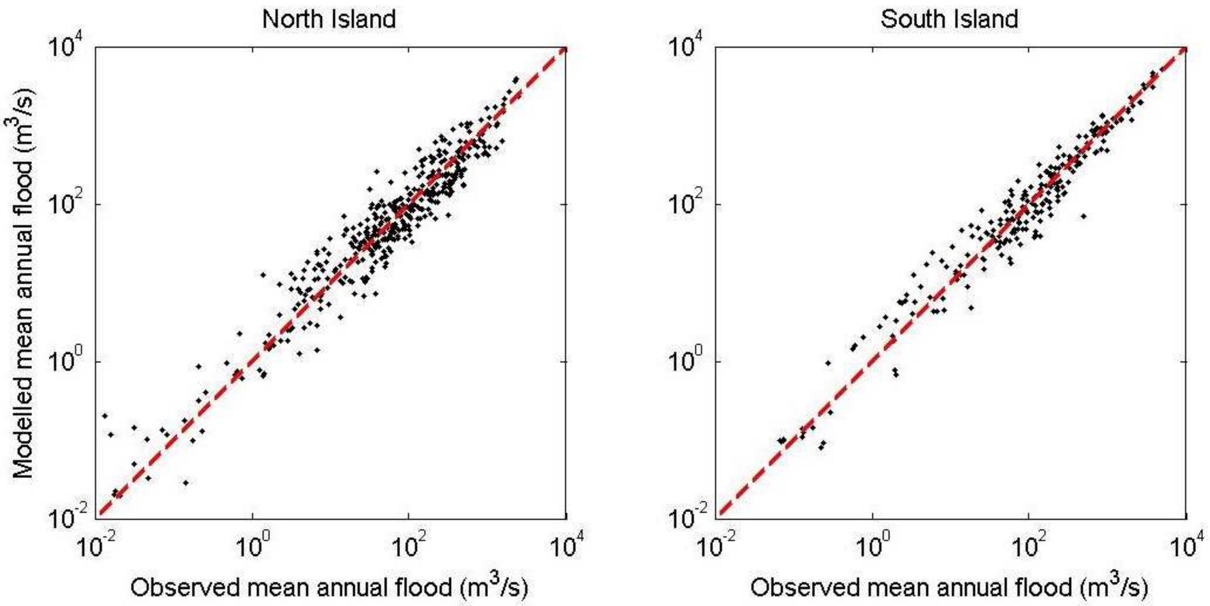

Figure 5-3: Observed mean annual flood vs. modelled mean annual flood by island. North Island (left) and South Island (right).

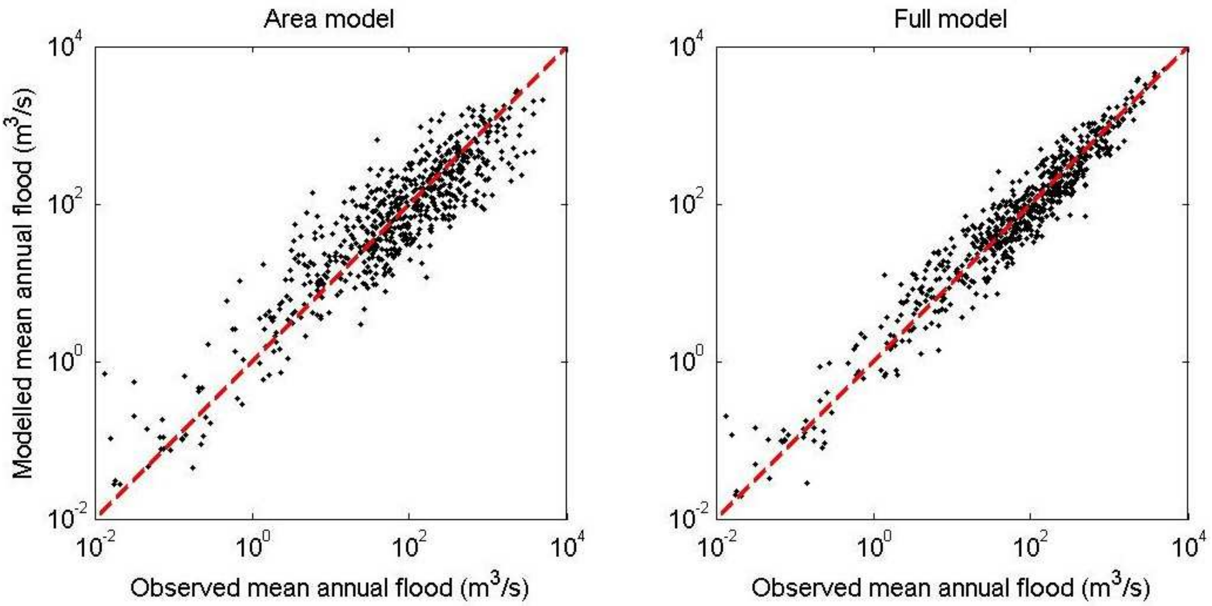

Figure 5-4: Observed mean annual flood vs. modelled mean annual flood. A model only using area (left) and the full model for all New Zealand (right).

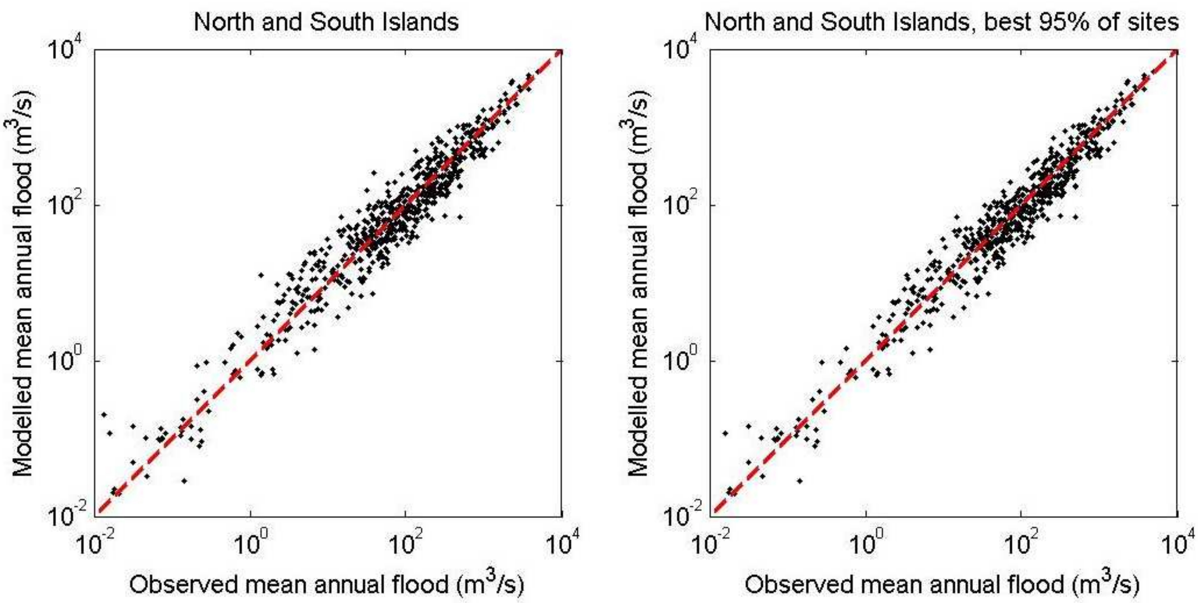

Figure 5-5: Observed mean annual flood vs. modelled mean annual flood showing effect of outliers. All New Zealand, all data (left) and all New Zealand, 95% of data with lowest RMSWRE (right).

Error statistics for the models are presented in Table 5-2. Excluding the outliers, the RMSWRE for the North Island is 61% and for the South Island 44%, showing improvement in the South Island, while the North Island error is larger than for the combined model.

Table 5-2: Bias and RMSWRE of a model split by island.

	n	Relative weighted bias	RMSWRE
All Data	636	19%	91%
best 95%	604	3%	55%
All North Island	397	21%	102%
best 95%	377	4%	61%
All South Island	239	15%	66%
best 95%	227	3%	44%

Overall, by applying the multiple regression to each island, we have reduced the bias of the best 95% from 24% to 3% and the standard error from 102% to 55%, by comparison with the simple area-based model described in Table 5-1 above. Results for the North Island are slightly worse than the national average with bias 4% and RMSWRE 61%, and the South Island is better with bias 3% and RMSWRE 44%.

Figure 5-6 shows the spatial distribution of relative error in mean annual flood for this model.

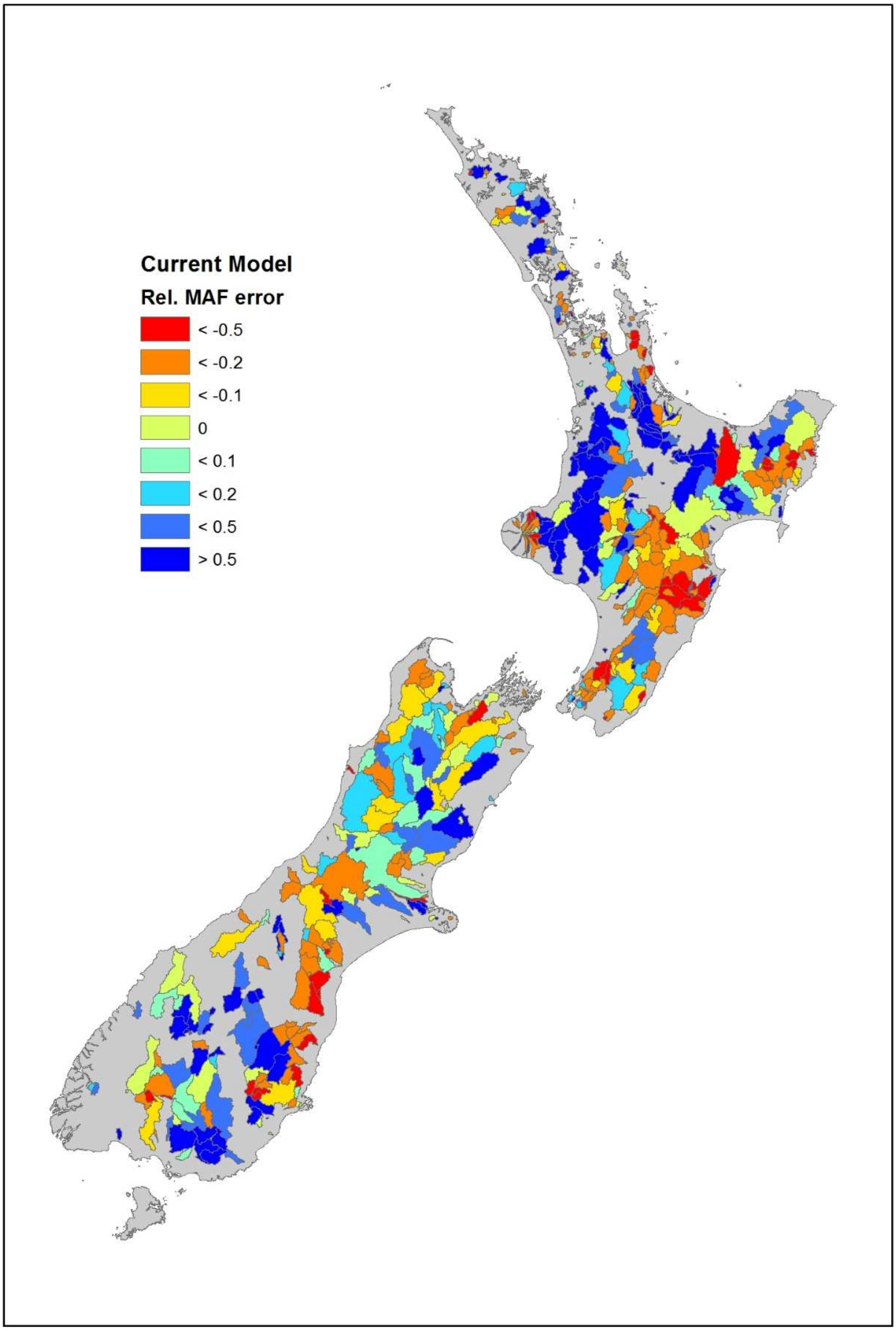

Figure 5-6: Relative errors in mean annual flood for the current model.

5.3 Reassessment of McKerchar and Pearson results

A study (McKerchar and Macky 2001) comparing the regional method to results obtained from use of rainfall-runoff methods used data from six flow records that had not been used in derivation of the procedure. From this limited sample they concluded that the regional method had lower uncertainty than the use of rainfall-runoff methods. We are now in a position to repeat aspects of this analysis using over 600 flow records, with longer data series. This comparison will enable a revised view of the uncertainty of the 1989 method.

From the version of the 1989 flood estimates attached to the digital stream network of New Zealand we derive a regional estimate of Q_{mean} and q_{100} at each flow recorder. From the dataset we take the Q_{mean} and q_{100} values calculated using the Gumbel (EV1) distribution as this was the one used in 1989. We can test three aspects of this dataset: firstly the bias and standard error of the complete set; secondly the bias and standard error of those sites used in the 1989 study, and thirdly the bias and standard error of sites not used in the 1989 study. This latter test is comparable to that of (McKerchar and Macky 2001). Results of these comparisons are detailed in Table 5-3.

Table 5-3: Bias and RMSRE of McKerchar and Pearson estimates of MAF and q_{100} compared to the current dataset. Statistics presented for 95% of the data series, excluding 5% worst behaved.

Statistic		Sites	Bias	RMSRE
Q_{mean}	1989 study	343 sites	-0.9%	22%
	Current study	All sites – 648	3%	49%
		Sites used in 1989 – 301	-5%	32%
		Sites not used in 1989 - 347	14%	70%
q_{100}	1989 study	275 sites	0.3%	17%
	Current study	All sites – 648	1%	18%
		Sites used in 1989 – 301	1%	14%
		Sites not used in 1989 - 347	2%	21%

For mean annual flood overall the standard error is larger (less so for sites used in the original study, more so for new sites). The overall New Zealand-wide error is ±49%. Bias is also larger but still small relative to the standard error. The overall New Zealand-wide bias is 18%. For the growth factor q_{100} results show little change in uncertainty. Maps of the error in each statistic are presented to illustrate the spatial distribution of errors.

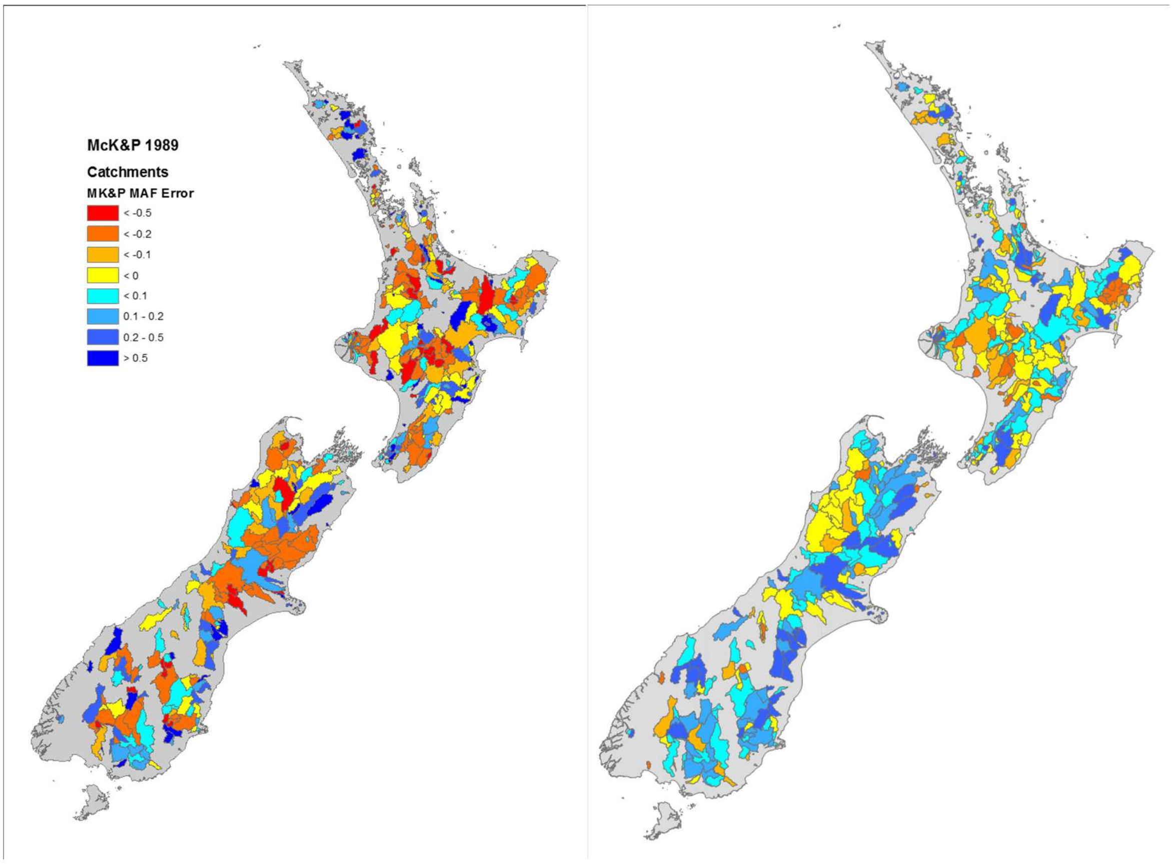

Figure 5-7: Maps of the error term for MAF (mean annual flood) (left) and q_{100} (right) assessed against the 1989 method. Colour scales are the same for both maps.

6 Estimation of flood growth curves

The second aspect of flood frequency after estimation of the magnitude of the mean annual flood as a measure of the catchment flood yield, is an estimate of the growth curve so that floods of other return periods than the mean annual can be assessed. In McKerchar and Pearson (1989) this was approached by provision of contours that allowed estimation of q_{100}, the ratio of the 100-year return period flood, to the mean annual flood. Because the EV1 or Gumbel distribution was adopted after testing, this single number allowed estimation of floods of all return periods within the defined range, set at 2 to 100 years (or 50% to 1% annual exceedance probability). This linear feature of the EV1 made the estimation method easy to apply in practice.

We repeated the tests used in McKerchar and Pearson (1989) to assess the degree to which the new dataset conformed to EV1 assumptions. The statistical test derived by (Hosking, Wallis et al. 1985) uses a test statistic $z = k^*(n/0.5633)^{0.5}$, which is asymptotically distributed standard normal. 'k' is the third parameter of the GEV distribution. If z, and hence k, is significantly positive (negative), the EV1 is rejected in favour of the EV3 (EV2). If not, then EV1 is acceptable.

Results of application of the Hosking et al. (1985) Z-test are best seen on a linear moment skew-kurtosis diagram, as shown in Figure 6-1.

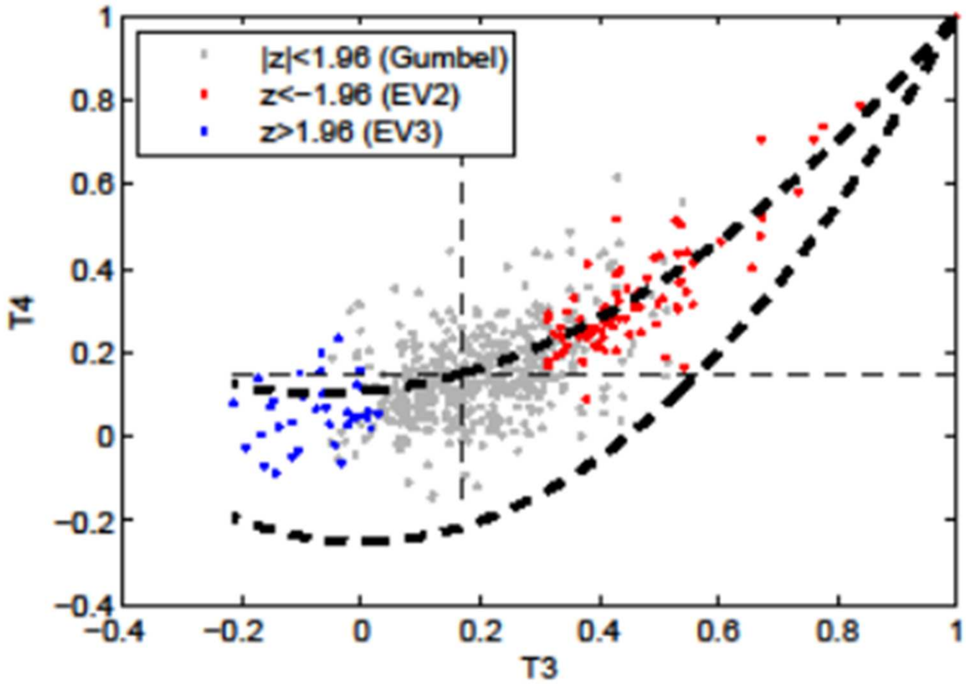

Figure 6-1: Linear moment skew vs. kurtosis for annual maxima flood series. Blue dots are significantly EV3; red dots are significantly EV2; grey dots are suitably Gumbel (EV1). The intersection of dashed lines indicate the EV1 data point (fixed skew and kurtosis), and the heavy dashed curve through that point is the GEV distribution. The lower heavy curve is the theoretical envelope.

Figure 6-2 shows the spatial distribution of the k parameter of the GEV distribution and the z-statistic of Hosking et al. (1985). In the North Island patterns are difficult to discern, but in the South Island there is a distinct area in the south and east that shows distinct EV2 tendencies. This is echoed by the distribution of the GEV k parameter from the maps used in the HIRDS package, but there the picture is far more muted, and tests show that a national regression or even a South Island only regression is a poor predictor of GEV k for floods.

The catchment physical and climate/weather related variables available from the development of the mean annual flood model do not assist in predicting this distribution of GEV-k either.

We have tested the McKerchar and Pearson (1989) contours of q_{100} as a predictor of q_{100} from the current dataset. This derivation assumes that the 1% aep flood (Q_{100}) has been assessed with the assumption of a Gumbel distribution, sometimes applied to biennial maxima as per McKerchar and Pearson (1989).

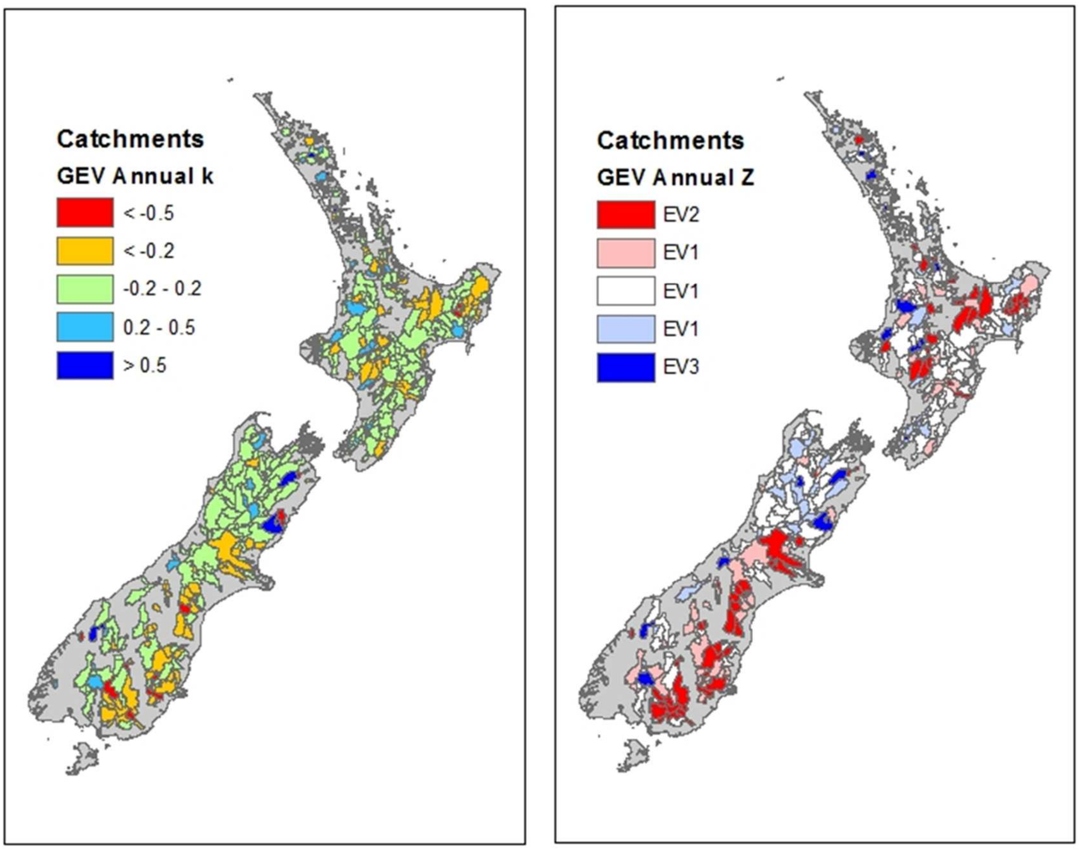

Figure 6-2: Maps of GEV k parameter and Hosking et al. z statistic.

As discussed in section 4.3 above, the uncertainties of the estimates of q_{100} from the current implementation of McKerchar and Pearson (1989) are similar to their published uncertainties (see Table 5-3). This suggests that for flood growth curves at least, the existing method could still be used with some degree of certainty.

The concentration of under-estimation of q_{100} in the south-east of the South Island suggests that further investigation is warranted. It may be that a combination of revised contours and/or a region approach would provide a better result here.

7 Discussion

Testing of the new dataset shows that when a simple area model is considered, the standard errors are comparable to those for the McKerchar and Pearson (1989) dataset. Thus we must look elsewhere for an explanation of the greater uncertainty of the current model when compared to the previous contoured solution.

The longer time span of the current dataset covers several IPO periods. The previous dataset finished in the mid-1980s, so many of the data series covered only one IPO period. The IPO period 1978 to 1999 is anomalously westerly. This may have added complexity to the model. It is possible that adjustment of flood series according to their time span and location could mitigate some if the uncertainty. Various analyses of IPO effects in New Zealand (Woods, Henderson et al. 2011, McKerchar and Henderson 2003; Griffiths, Pearson et al. 2009) suggest this as a fruitful area of enquiry.

Mean annual precipitation is the second most selected variable in the regression analysis after catchment area. However it is not directly related to flood production, but rather a surrogate for rain intensity. Thus we are clearly not capturing the true drivers of floods. In general the use of variables derived from HIRDS did not add appreciably to the fit of models. This aspect needs to be the subject of further research.

We have several avenues that could be explored. There are new maps of New Zealand's catchment physical properties (SMAP and QMAP). These may contribute better numeric variables for the understanding of hydrological processes, including floods and low flows. Output from the national hydrological model has shown some utility in flood estimation (unpublished data). Hand in hand with further development of this aspect is the need for a revised rainfall map of New Zealand. Inclusion of regional council rain data into both annual rain and event rain estimation methods has shown considerable improvement in estimation of catchment rain.

A finely detailed contour map of error corrections or a highly detailed random forest are both able to perfectly model the training data, but they need to be able to fit verification data too. These ideas have not been tested to date.

The benchmark assessments were computed by applying McKerchar and Pearson to the modern dataset. The assessed model error is ±49%.

Users have thus far been comfortable with McKerchar and Pearson, and its stated uncertainties. However our re-assessment of this uncertainty indicates that it may be almost twice that previously understood. In this light the error of the new model, at ±55%, is similar to the assessed error of the previous method for all New Zealand, of ±49%.

The question of whether this scale of uncertainty is acceptable, and how practitioners will respond to it, remains to be resolved.

7.1 Future work suggestions – a potential plan for what next

Possible avenues for further work include:

- Investigate the contoured residual dataset. Objective contouring using a spline-based method, akin to the VCSN, should result in smaller model error for the data set.

- Get greater clarity about the potential effect of top-kriging; does it allow the effect of geology or vegetation, for example, to be explicitly incorporated before being applied?
- Develop a new rainfall surface for New Zealand to act as a better interpolator for both annual and event rain estimation.
- Develop areal reduction factors for HIRDS as input to catchment estimation of flood size. However, there are still issues around large catchments across climate and weather zones. Is there a natural limit to the catchment area that these methods should be applied to (both upper and lower potentially)?
- Continue to explore the national TopNet output as a potential substitute.
- Refine how knowledge of what aspects of the climate and weather systems are responsible for floods by studying the circumstances of historical floods. This may allow us to move away from mean annual statistics to more appropriate time-scales.
- Implement the accepted models (with provision of flood estimates and uncertainty) on a publicly available web-site such as stream-explorer.niwa.co.nz.

These should be the subject of discussion with stakeholders, and lead to the development of a future plan of work.

8 Conclusions

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

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. No spatially coherent temporal trends were detected 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.

The uncertainty of the previous model for estimating mean annual flood is larger than originally stated, but the uncertainty of the growth curve parameter (q_{100}) is very similar. Derivation of a new model for mean annual flood proved more difficult than expected, at least to the extent that the eventual uncertainty is more than twice the previously published estimate of ~±22%.

Assessment of the higher order moments of the new dataset confirm previous work that shows some areas of New Zealand may have EV2 tendencies. However with the climate and physiographic variables at our disposal, as well as limited scientific knowledge of the hydro-climatic processes that contribute to higher order moments, development of a suitable model to allow estimation of these factors everywhere proved elusive. The previous model remains as the best option for flood growth estimation.

Non-instrumental historic events have been included in the dataset, extracted from the published data list in McKerchar and Pearson (1989) and Beable and McKerchar (1982). However, few of these were returned as part of the data provided for the study. It is possible that these data are still held by recording authorities, and it would be useful to have confirmation of their current level of acceptance of these early data.

Isolated historic events remain an issue. Very few examples were provided (only six flow records) whereas Beable and McKerchar (1982) document over 50 sites where this knowledge existed. The analysis techniques for these remain subjective and we recommend that these be used for at-site estimation rather than as components of the regional method.

9 Acknowledgements

The work described in this report has been funded by MBIE through an EnviroLink tools grant C01X1308, and NIWA Core funding under the programme ”Impacts and risks of weather related hazards”.

We thank regional council and NIWA staff who have collected and performed quality assurance on river flow data over many years, to provide the archives on which this study is founded. Collection of river flow data, and especially flow gaugings during floods, is often carried out in arduous circumstances.

Alistair McKerchar and George Griffiths provided advice and many useful discussions.

We also thank the regional stakeholder representatives for the EnviroLink funding, and especially Martin Doyle and Gary Clode who have provided valuable feedback along the way.

10 References

Acreman, M.C., Sinclair, C.D. (1986) Classification of drainage basins according to their physical characteristics: An application for flood frequency analysis in Scotland. *Journal of Hydrology*, 84: 365-380.

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

Bocchiola, D., De Michele, C., Rosso, R. (2003) Review of recent advances in index flood estimation. *Hydrology and Earth System Sciences*, 7(3): 283-296.

Burn, D.H. (1990) Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research*, 26: 2257-2265.

Castigloni, S., Castellarin, A., Montaanari, A., Skøien, J.O., Laaha, G., Blöschl, G. (2011) Smooth regional estimation of low-flow indices: physiographical space based interpolation and top kriging. *Hydrology and Earth System Sciences*, 15: 715-727.

Gabriele, S., Arnell, N. (1991) A hierarchical approach to regional flood frequency analysis. *Water Resources Research*, 27: 1281-1291.

Griffiths, G.A., Pearson, C.P., McKerchar, A.I. (2009) Climate variability and the design flood problem *J HYd (NZ)*, 48(1): 29-38.

Helsel, D.R. and Hirsch R.M. (2002) Statistical Methods in Water Resources Techniques of Water Resources Investigations, Book 4, chapter A3. U.S. Geological Survey. 522.

Hosking, J.R.M. (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society (Series B)*, 52(1): 105-124.

Hosking, J.R.M., Wallis, J.R. (1993) Some statistics useful in regional frequency analysis. *Water Resources Research*, 29(2): 271-281.

Hosking, J.R.M., Wallis, J.R. (1997) *Regional frequency analysis*. Cambridge University Press: 224.

Hosking, J.R.M., Wallis, J.R., Wood, E.F. (1985) Estimation of the General Extreme Value by the method of probability weighted moments. *Technometrics*, 27(3): 251-261.

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

Institution of Engineers Australia (2012) Australian Rainfall and Runoff. *Revision Projects*, Project 5, Regional flood methods, Stage 2 report P5/S2/015.

Martins, E.S., Stedinger, J.R. (2000) Generalized maximum likelihood GEV quantile estimators for hydrologic data. *Water Resources Research*, 36(34): 737-744.

McKerchar, A.I., Henderson, R.D. (2003) Shifts in flood and low-flow regimes in New Zealand due to interdecadal climate variations. *Hydrological Sciences Journal*, 48(4): 637-654.

McKerchar, A.I., Macky, G. (2001) Comparison of a regional method for estimating design floods with two rainfall-based methods. *Journal of Hydrology (NZ)*, 40(2): 129-138.

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

Mosley, M.P. (1979). Prediction of hydrologic variables from channel morphology, South Island rivers. *Journal of Hydrology (NZ)* 18(2): 109-120.

Robson, A., Reed, D. (1999) Statistical procedures for flood frequency estimation, Flood estimation handbook, Volume 3.

Rodgers, M.W., Thompson, S.M. (1992) Tideda Reference Manual. *Publication of the Hydrology Centre*, No. 24.

Skøien, J.O., Merz, R., Blöschl, G. (2006) Top-kriging – geostatistics on stream networks. *Hydrology and Earth System Sciences*, 10: 277-287.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E. (1993) Frequency analysis of extreme events. In: D.R. Maidment (Ed). *Handbook of Applied Hydrology*. McGraw-Hill, New York.

US Water Resources Council (1981) Guidelines for determining flood flow frequency, Bulletin 17B.

Walter, K.M. (2000) Index to Hydrological Recording Sites in New Zealand. *NIWA Technical Report*, No. 73: 216.

Wiltshire, S.E. (1986) Identification of homogeneous regions for flood frequency analysis. *Journal of Hydrology*, 84: 287-302.

Woods, R.A., Henderson, R.D., McKerchar, A.I. (2011) Recent Impacts of Climate Variability on Streamflow. *50th Jubilee New Zealand Hydrological Society Conference: Learning from the Past: Creating the Future*, Wellington, N.Z., December 2011.