SEASONAL PREDICTION OF LAKE INFLOWS AND RAINFALL IN A HYDRO-ELECTRICITY CATCHMENT, WAITAKI RIVER, NEW ZEALAND

J.M. PURDIE[1] and W.E. BARDSLEY[2]

(1) Meridian Energy Ltd, State Highway 8, Twizel 7944, New Zealand

(2) Department of Earth and Ocean Sciences, University of Waikato, Hamilton, New Zealand

Corresponding author address: PO Box 53, Twizel 7944, Canterbury, New Zealand

Corresponding author email: jennifer.purdie@meridianenergy.co.nz

ABSTRACT

The Waitaki River is located in the centre of the South Island of New Zealand, and hydro electricity generated on the river accounts for 35-40% of New Zealand's electricity. Low inflow years in 1992 and 2001 resulted in the threat of power blackouts. Improved seasonal rainfall and inflow forecasts will result in the better management of the water used in hydro generation on a seasonal basis.

Researchers have stated that key directions in the fields of seasonal rainfall and streamflow forecasting are to decrease the spatial scale of forecast products, and to tailor forecast products to end user needs, so as to provide more relevant and targeted forecasts.

Several season ahead lake inflow and rainfall forecast models were calibrated for the Waitaki river catchment, using statistical techniques to quantify relationships between land-ocean-atmosphere state variables and seasonally lagged inflows and rainfall. Techniques included principal components analysis and multiple linear regression, with cross-validation techniques applied to estimate model error, and randomisation techniques used to establish the significance of the skill of the models.

Many of the models calibrated predict rainfall and inflows better than random chance, and better than the long term mean as a predictor. 95% confidence limits around most model predictions offer significant skill when compared with the range of all *probable* inflow seasonal totals (based on the 80 year recording history in the catchment). These models explain up to 19% of the variance in season ahead rainfall and inflows in this catchment.

Seasonal rainfall and inflow forecasting on a single catchment scale, and focussed to end user needs, is possible with some skill in the South Island of New Zealand.

Key words: New Zealand; seasonal forecasting; Principal Components Analysis; Multiple linear regression; cross-validation; randomisation testing; rainfall; inflows

1.   INTRODUCTION

Approximately 60% of New Zealand's electricity is produced from hydro generation. The Waitaki River is located in the centre of the South Island of New Zealand. It is sourced in the highest part of the Southern Alps, at 3764masl, and flows 200km to the sea on the east coast of the South Island (Figure 1). It passes through eight hydro power stations run by Meridian Energy Ltd, which collectively produce 35-40% of New Zealand's electricity. There is significant snow storage in the Waitaki catchment, and this results in a seasonal cycle which produces maximum inflows in spring-summer, with minimum inflows in autumn-winter. The pattern of peak demand for electricity, however, is the converse of this, with peak demands in winter and low demand in summer. Low inflow years in 1992 and 2001 resulted in the threat of power blackouts. In addition to this, New Zealand hydro storage can hold only 17% of total annual inflows (Fitzharris, 1992), and national demand for electricity is currently growing at about 2% a year. It is clear that considerable planning of the use of the water resource is required to prevent water shortages. Improved seasonal rainfall and inflow forecasts will result in the better management of the water used in hydro generation on a seasonal basis.

The unique configuration of the Southern Alps in the South Island of New Zealand, lying perpendicular to the mid-latitude westerly winds results in a strongly regionalised rainfall regime, with very high rainfalls on the west coast and low rainfalls in the east. It has long been recognised that variations and trends in New Zealand regional rainfall respond significantly to circulation changes in the southwest Pacific (Salinger and Mullan 1999, Henderson and Thompson 1999, McKerchar et al. 1996, McKerchar et al. 1998, Wratt et al. 1996), and in particular the low level westerly winds (De Lisle 1956, Salinger and

Mullan 1999). De Lisle (1956) found that variations in rainfall relate to a general polewards displacement of the position of the sub-tropical anti-cyclone tracks.
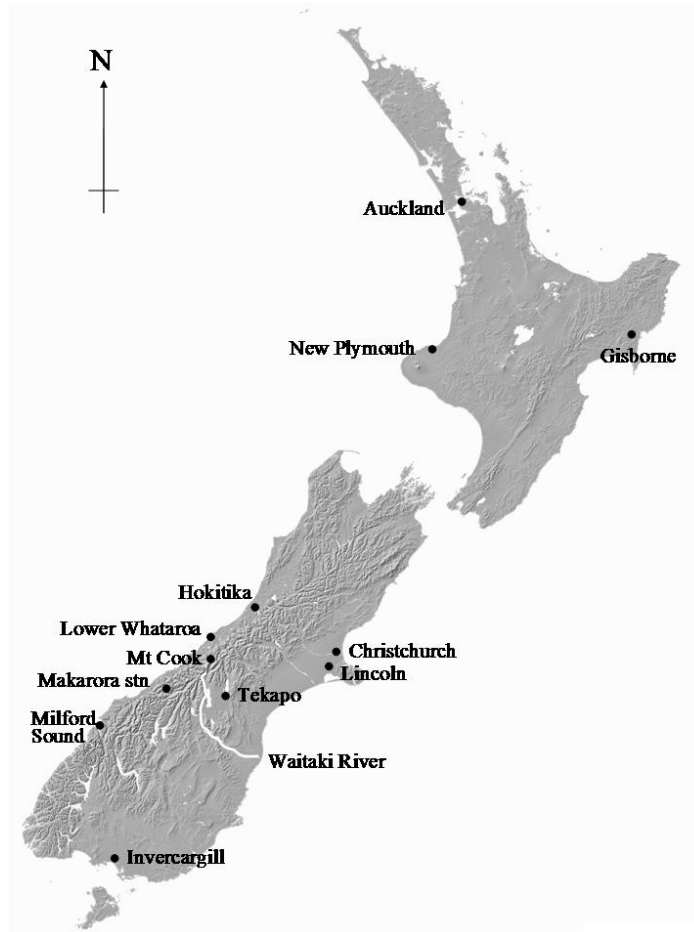
Figure 1: Location of the Waitaki river and sites used in this study

Previous seasonal forecasts of rainfall and inflows in New Zealand have involved statistical techniques, the use of neural networks, and numerical model simulations (Mullan & Renwick 1996, Zheng et al. 2000, Mullan & Thompson 1993, Mullan & Thompson 1995, Mullan & Thompson 2006, McKerchar et al. 1996, Fitzharris et al. 1998). Currently, the National Institute for Water and Atmospheric Research (NIWA) produces a three monthly probabilistic forecast using both statistical techniques and General

Circulation Model (GCM) outputs. This gives the general public a three month probabilistic forecast for each of six regions over New Zealand (http://www.niwascience.co.nz/ncc/cu/archive).

The advent of better computing power and longer, higher quality data records has resulted in a proliferation of research into seasonal to interannual climate forecasting in recent years. These include both statistical (Sharma 2000a, b, Sharma et al. 2000, Zheng et al. 2000, Eshel 2000, Ryuiz et al. 2003, Drosdowsky and Chambers 2001, Francis and Renwick 1998, Krishna Kumar et al. 1995, Greischar and Hastenrath 2000) and dynamical (Colman 2004, Feddersen et al. 1999, Fennessey and Shukla 2000, Widmann 2003) techniques. Recently neural networks have also been used in climatological prediction (Yuval 2000, Chakraborty et al. 1992). Goddard et al. 2001 state that the above forecasting methods are generally competitive with each other in terms of the skill of their forecasts. This is reiterated by Zebiak 2003.

Several reviews of seasonal climate forecasting internationally have been undertaken in the past 15 years (Palmer and Anderson 1994, Goddard et al. 2001). More recently, Zebiak (2003) reviewed research potential for improvements in climate prediction, and outlined several areas for improvement in climate forecasting. Two of these are: greater spatial and temporal resolution in climate forecasts, and tailored forecast products aimed at the user community. These particular areas of focus are addressed by this research.

Where long historical records of variables of interest are available, statistical approaches that relate "at site" hydrology or climate to large scale ocean-atmosphere state variables would seem to provide the best basis for useful seasonal to inter-annual flow and rainfall forecasts. However, as with any real system,

these relationships are likely to be complex. Sharma (2000a) states that multi-variate analyses between groups of possible "predictor variables' would seem to be the way forward in hydro-climatic prediction research.

In the quest to quantify the multi-variate relationships mentioned above, many different statistical techniques have been used in the past. Mullan and Renwick (1996) used empirical orthogonal functions (EOF) analysis, Singular Value Decomposition Analysis, and Canonical Correlation Analysis to forecast monthly and seasonal rainfall and temperature in New Zealand. Shabbar and Barnston (1996) and Landman (1997) also used canonical correlation analysis in forecasting Canadian and South African rainfall (respectively). Statistical analyses of time series were used by Eshel et al. (2000) to forecast Eastern Mediterranean droughts, and Zheng et al. (2000) used analysis of variance techniques in attempting to quantify the potential predictability of seasonal means of monthly time series of meteorological variables in New Zealand.

Of the multi-variate techniques used in climatic and hydrological forecasting, many use multiple regression as a basis for further analysis, or as a further step once data have been detrended or smoothed in some way. Many researchers have used multiple regression in the study and forecasting of rainfall and other climatological variables (Greischar and Hastenrath 2000, Lall et al. 1996, Kidson and Gordon 1986, Hay et al 1998, Kabanda and Jury 1999, McKerchar et al. 1996, Tait and Fitzharris 1998, Zheng and Renwick 2003). Principal components analysis (PCA) or empirical orthogonal functions (EOFs) can be used to identify modes of variability in large hydroclimatic data sets, such as gridded sea level pressures or temperatures. PCA is also used as a data reduction technique, and has been widely used in

hydro-climatic research in both these capacities (Kidson and Barnes 1984, Mutai and Ward 2000, Tait and Fitzharris 1998, Smith 2000, Vuille et al. 2000).

Many researchers have concluded that seasonal *precipitation* anomalies may be forecast by utilising various different techniques and possible predictors. However, the extent to which land surface processes (such as river flow) respond to ocean-atmosphere variability has thus far received "very limited attention" (Wedgbrow et al. 2002). Seasonal streamflow forecasting fits the recommendations of Zebiak (2003), mentioned above. Streamflow forecasting is often aimed at a single catchment or group of proximal catchments, and often aimed at the needs of river users, such as hydro-electricity generators, irrigators, and recreational users.

It is emerging in the research that seasonal streamflow forecasting has advantages over rainfall forecasting, and can result in greater skill, especially for forecasts over smaller spatial areas. There appear to be two main reasons for this. Firstly, river flows are often a "smoothed" representation of rainfall in a catchment, integrating all rainfall over the area of the catchment. This is especially true in mountainous catchments where precipitation variability is high both spatially and temporally, and a point measurement may not well represent the amount of water arriving in the catchment (Benoit et al. 2000). Secondly, many of the users of seasonal rainfall forecasts are actually utilising the water that flows past their land in the form of a river flow, and thus a river flow forecast has more direct applicability to their industry than a precipitation forecast which may be recorded at some distal location (Wedgbrow et al. 2002). Cayan et al. (1999) have shown that ENSO influences on streamflow variations and extremes are proportionately larger than the corresponding precipitation teleconnections. Dettinger et al. (1999) state that in many rivers, significant (weeks to months) delays between precipitation and the release to streams of snowmelt

or groundwater discharge can support even longer term forecasts of streamflow than is possible for precipitation. Benoit et al. (2000) support the above statements by noting that the integrating effect of streamflows filter out some of the spatial and temporal variability that complicate the point by point precipitation verifications that are more commonly used. Wedgbrow et al. (2002) go further, in concluding that investigations of the indices for direct long-range prediction of hydro-climatological variables are "essential to assist the future allocation of finite water resources".

In the search for predictors of inflows in a snowy alpine catchment, some index of winter snow storage in the catchment is desirable as an independent variable in model calibration. However, quantification of the water equivalent stored as snow in a remote catchment with significant orography is a difficult task, and the tools that exist have wide error margins. The relationship between local air temperatures and seasonally lagged river flows has been documented (NIWA 1994), and it would seem to be a reasonable approach to use these variables as predictors in attempting to quantify the empirical relationship between them and seasonally lagged inflows in a snowy catchment, rather than trying to quantify winter snow amount in some other way.

In the New Zealand context, seasonal inflow forecasting attempts have focussed on inflows to hydro electricity reservoirs. An analogue forecast model was formulated to predict monthly inflows to hydro lakes by Mullan and Thompson (1993). Later, Mullan and Thompson (1995) presented an improved version of their analogue forecasting model, which included SST anomalies and lake inflows as predictors. The model predictions were found to be statistically significant when tested on independent data. However, when forecast skill was assessed using rank error criteria, the best case forecasts were found to be not better than a forecast of persistence. Mullan and Thompson (2006) created a further

analogue forecasting model, with statistically significant success in month ahead and season ahead forecasts of rainfall and temperatures in some areas, but not on the west coast of the South Island. McGowan and Sturman (1996) used a multiple linear regression approach to predict one month ahead inflows to Lakes Pukaki and Tekapo from temperature, precipitation, and a synoptic circulation index. Other New Zealand studies have included Bayesian methods to predict summer inflows into South Island hydro lakes (McKerchar et al. 1996) and multi-discriminant analysis to forecast seasonal and annual precipitation, with "limited success" (Peters and Painter 1997).

One of the most crucial and difficult aspects of developing a statistical rainfall or inflow forecasting model is obtaining a realistic estimate of forecast skill (Michaelson, 1987). Tests on the Australian Bureau of Meteorology Seasonal Climate Outlook tercile probabilistic forecasts, show them to be correct about 70% of the time (Sturman and Tapper, 1996). Renwick (2003) discusses how forecast accuracy varies over time, but noted that the skill of seasonal forecasts from the National Institute for Atmosphere and Water (NIWA) are improving year by year. Renwick uses the "hit rate" to score NIWA forecasts. The hit rate is the fraction of times the tercile predicted with the highest probability actually occurs. A random guess as a prediction would therefore result in a hit rate of 33%. In the first year of producing seasonal rainfall forecasts, the hit rate was approximately 33%, or equivalent to the no-skill mark. However, more recently it has consistently been between 40 and 50%. Renwick also states that rainfall is most skillfully forecast over the South Island of New Zealand, and attributes this to the strong dependence of South Island rainfall patterns on the predominant westerly winds and their interactions with the steep topography of the Southern Alps, and that the seasonal changes in the position of the westerly wind belt show some predictability.

Ultimately, the justification for any forecast model is that it hopefully supports better decision making. The usefulness of forecasts in supporting decision making will always depend on their error characteristics, which are defined through verification methods (Wilks, 1995). The enormous pool of potential predictors and limited temporal degrees of freedom of most geophysical data means that statistical models require rigorous testing in operational settings to ensure significance (Goddard et al. 2001).

The purpose of this paper is to predict season ahead lake inflows and rainfall for a single, mountainous, catchment, and to focus the output forecasts to end user needs.

2.   DATA

Variables with some potential lag relationship with the predictands were chosen through review of previous research, to ascertain what predictor variables had been used with any method and any level of skill by other researchers in New Zealand, Australia and to a lesser extent, other international studies. Good quality long term synchronous data records were needed, and particular attention was paid to data records of long duration. Some data records that had been found to be possibly relevant were abandoned due to lack of record duration or quality.

Seasonal averages or totals were formulated for all datasets, and the seasonal cycle was removed for all datasets (except Sunspot number, which has no seasonal cycle) by subtraction of the 1950-00

normal for each season. Seasons are denoted as Summer: Dec – Feb, Autumn: Mar – May, Winter: Jun – Aug, and Spring: Sep – Nov.

The six dependent variables used were: Lake Pukaki total seasonal inflow volume, sum of Lakes Pukaki and Tekapo inflows, number of days of rain at the Hermitage, combination of rain gauges in the Southern Alps, Hokitika rainfall, and Hermitage rainfall.

The independent datasets gathered for the study are listed below:

- Lake Pukaki inflows (m^3)
- Lake Tekapo inflows (m^3)
- Hermitage rainfall (mm)
- Hokitika rainfall (mm)
- Lower Whataroa rainfall (mm)
- Makarora station rainfall (mm)
- Milford Sound rainfall (mm)
- Hermitage rain days (>1mm)
- Hermitage temperature (°C)
- Hokitika temperature (°C)
- Tekapo temperature (°C)
- Lincoln temperature (°C)
- Trenberth indices
- Sunspot number
- QBO of stratospheric winds

- SOI

- PDO

- SST indices

- MSLP on a 5˚ x 5˚grid, lat 5˚N - 75˚S, long 0-355˚ (hPa)

- 700hPa geopotential height on the same grid as MSLP (height above sea level)

Descriptions and sources for these datasets are given below.

Lake Pukaki and Lake Tekapo are the two main hydro storage lakes in the Waitaki catchment, and in New Zealand as a whole. Inflow records for these lakes were obtained from the Meridian Energy Ltd Hydrological database. These lake inflows have been measured since 1925, when the first concepts of a hydro electricity scheme in the catchment were mooted. However the early years have many gaps in the record, and measurement methods are not well documented. Gaps in the record were filled in, but most trust in the record exists post-1940.

Rainfall records for the following climate stations were obtained from Dr Jim Salinger at NIWA or through the NIWA climate database (CLIDB): Hermitage, Hokitika, Lower Whataroa, Makarora station, and Milford Sound. The Hermitage is the rain gauge at Mt Cook village, an alpine village at 785m above sea level in the head of the Waitaki catchment. The last three records were combined with Hermitage rainfall in a weighted combination that represented their contemporary linear relationships with Lake Pukaki inflows. As well as this strong relationship, these stations were chosen for their proximity to the north-south axis of the Southern Alps mountain range. The number of days of rain at the Hermitage

(1mm or more) was also used in this study. The mean air temperature records for several stations were also obtained from the NIWA CLIDB. These stations were Hokitika, Tekapo, and Lincoln.

Trenberth Indices of air pressure are based on Trenberth (1976). These datasets were initially received from Dr Andrew Tait, at NIWA in Wellington. However, when these datasets needed to be updated, pressure data was sourced from the NIWA Clidb, and methods to derive the indices followed Trenberth (1976). Trenberth Indices are the anomalies from normal of the difference in station level pressure between two stations. "Z" indicates zonal flow, "M" indicates meridional flow, and "MZ" is a combination of the two. These are the relevant stations for each index:

- Z1 - Auckland and Christchurch

- Z2 – Christchurch and Campbell Island

- Z3 – Auckland and Invercargill

- Z4 – Raoul and Chatham Islands

- M1 - Hobart and the Chatham Islands

- M2 – Hokitika and Chatham Islands

- M3 - Hobart and Hokitika

- MZ1 – Gisborne and Hokitika

- MZ2 – Gisborne and Invercargill

- MZ3 – New Plymouth and Chatham Islands

Sunspot numbers have been recorded since 1818, and daily data were obtained from the National Geophysical Data Centre website of the National Oceanic and Atmospheric Administration, in the United

States (http://www.ngdc.noaa.gov/stp/SOLAR/SSN/ ssn.html). Seasonal averages were then calculated from these daily data.

The monthly Quasi-biennial oscillation of stratospheric winds at 30hPa over Singapore (Naujokat, 1986) was supplied by Dr Barbara Naujokat of the Free University of Berlin. This dataset is also available on the internet from the website of the Joint Institute for the Study of the Atmosphere and Ocean (JISAO at the NOAA/University of Washington) at: http://tao.atmos.washington.edu/data_sets/qbo/.

There are many derivations of the Southern Oscillation Index available globally. The one chosen for use in this study is from the Bureau of Meteorology, Australia. It is defined as is the Troup SOI (Troup, 1965) which is the standardised anomaly of the mean sea level pressure difference between Tahiti and Darwin. It can be sourced from the Bureau of Meteorology (BoM) website at: http://www.bom.gov.au/climate/current/soihtm1.shtml.

The Interdecadal Pacific Oscillation (IPO) is noted by Renwick (2002) to be the quasi-symmetric Pacific-wide manifestation of the Pacific Decadal Oscillation (PDO) that has been described for the north Pacific by various other authors (Francis & Hare 1994, Mantua et al. 1997). The PDO index is used in this study, and can be sourced at the JISAO website at: http://jisao.washington.edu/pdo/PDO.

A set of sea surface temperature (SST) indices were formulated by Mullan (1998). These indices were derived from a principal components analysis of Southern Hemisphere sea surface temperatures and the contemporary and lag relationship of those principal components with New Zealand temperature and precipitation. Areas of the Southern Hemisphere were chosen where these relationships were strongest,

and SSTs averaged over these areas, and the anomalies from normal calculated. SST areas are defined in table 1.

Table 1: Definition of SST index key areas (Mullan, 1998)

Region	LongW	LonE	LatN	LatS	Location
1	160	190	-30	-45	Northern NZ
2	125	165	-35	-45	Australian Bight and Tasman Sea
3	210	270	5	-5	Nino 3
4	160	210	5	-5	Nino 4
5	150	180	-15	-30	New Caledonia
6	195	215	-20	-40	Pacific NE of NZ
7	75	115	-5	-25	Indian Ocean NW of Australia
8	60	85	0	-15	Equatorial Indian Ocean

Mullan's SST indices were sourced from Dr Brett Mullan, of NIWA, Wellington, originally, and then updated upon his advice by using Optimum Interpolation (V2) SST data from the Climate Data Centre (CDC) at NOAA (National Ocean and Atmospheric Administration) in the United States. (http://www.cdc.noaa.gov/cgi-bin/db_search/SearchMenus.pl).

NCEP/NCAR Reanalysis Gridded Sea Level Pressure (SLP) data and 700hPa geopotential height data on a 5° x 5° grid for much of the Southern Hemisphere were sourced from the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their Web site at http://www.cdc.noaa.gov/. Data for 5°N to 75°S, latitude, 0 – 355° longitude were originally gathered. The grid of data used was subsequently reduced from this initial dataset. See (http://www.cdc.noaa.gov/cgi-bin/db_search/SearchMenus.pl) for data source.

3. METHODOLOGY

3.1   Model calibration

The amount of data gathered as potential predictor variables in this study was considerable. Reducing the number of predictor variables before putting them through a principal components analysis was undertaken to try and eliminate some variables that had potentially less chance of having predictive ability before further analysis began.

Initially, a contemporaneous correlation matrix of all the variables was constructed. The results of this showed high correlations between adjacent gridded mean sea level pressure and 700hPa geopotential height sites over most of the area covered, and this justified the reduction of the gridded datasets from a 5x5 degree grid to a 10x10 degree grid. Backward and Forward Stepwise regression were conducted between all datasets and seasonally lagged lake inflows and rainfall to establish which variables had less significant relationships. Variables highlighted in both the backward and forward regressions as being of low significance to these relationships were eliminated from the model calibration process at this point. This exercise resulted in a reduction of spatial coverage of original data to that of 5° north to 55° south latitude, 75° to 355° longitude. A third method was employed in the data reduction exercise. This was to run all data through a Principal Components Analysis, and eliminate any independent variables that did not feature prominently as contributors to the resulting principal components. This resulted in just a few variables being eliminated, as the orthogonal nature of principal components results in most variables featuring highly in at least one principal component.

16

The original number of potential predictor variables was 1140 for each season. After the variable reduction exercise, the number of variables taken forward to further analysis was 115 for summer, 96 for autumn, 140 for winter, and 120 for spring. Model calibrations were performed on each season separately.

All independent datasets remaining after the intial data exploration were then subjected to an unrotated Principal Components Analysis, using a correlation dispersion. An unrotated PCA was used, because data compression (rather than physical interpretation) was the main aim of the PCA in this instance (Wilks, 1995). The resulting principal components were taken forward and used as the independent variables in a multiple regression with seasonally lagged inflows and rainfall.

The Scree test (Catell 1966), and Jolliffe's (Jolliffe 1972) version of Kaiser's rule (Kaiser 1960) were used in conjunction with each other to determine the number of eigenvalues retained from the PCA. This ranged from 19 to 25 eigenvalues. In this study over 50% of the variance in the data is explained by the first two to three PCs.

The SOI seems to be the most widely used predictor in seasonal forecasting in New Zealand. However, there is conflicting evidence in the literature as to whether the SOI has a significant relationship with lagged rainfall in New Zealand (Rhoades et al. 1990, Basher 1998, McKerchar and Pearson 1994, McKerchar et al. 1996, McKerchar et al. 1998, Fitzharris 1992, Fitzharris et al. 1998). To examine this question, these model calibrations were subdivided into calibrations with the SOI and PDO, and calibrations without the SOI and PDO.

Forward stepwise regression was the regression technique employed to regress the independent datasets (the principal components (PCs)) with the lagged dependent dataset (rainfall or inflows), and the F-test (the ratio of explained variances before and after the additions of a new predictor) was used as a stopping criterion, with F-to-enter 4 as the value at which no more variables are added to the regression equation. These model calibrations predicted six different dependent variables. Each of these six analyses was carried out for each of four seasons, resulting in twenty-four model calibrations. Each of these twenty-four model calibrations incorporated cross validation techniques. To this end, each dataset (time series) was divided into five periods, and models were calibrated on each of five forty year periods, and validated on each of five ten year periods.

3.2   Model validation

An approximately 50 season time series was used in the model calibrations, and a 5-fold cross-validation was conducted with a different 20% of the time series left out each time, as is recommended by Breiman and Spector (1992).

Several goodness of fit criteria were evaluated for use in this study. All compared how well the predicted values in the validation dataset "fit" the observed values. The "explained variance" is used widely in seasonal hydro-climatic forecasting both in New Zealand and internationally as a measure of model skill, and was used in this study as it seemed to give the most consistent and logical results.

18

The term "explained variance" has a different definition in least squares regression, and this should not be confused with the application of the term used here. The explained variance of a predictive model, E_v, used in this study is defined as:

$$E_v = 1 - \frac{(\frac{\sum(y-\hat{y})^2}{n-2})}{Var(obs)} \qquad (1)$$

where $(-\infty \leq E_v \leq 1)$

where y and $\hat{y}$ are the observed and model predicted values, and *Var(obs)* is the variance of the observed values.

The explained variance measure describes the proportion of variability within the predictand dataset captured by the model prediction, and is widely used in seasonal climate research (Eshel et al. 2000, Kabanda and Jury 1999, Kidson and Gordon 1986, Tomlinson 1981, Widmann et al. 2003).

3.2.1   Residual analysis

Before model predictions could be applied to error evaluation, some residual analysis needed to be carried out, to ascertain the distribution characteristics of the residuals from the model validations. Trends and bias in the model predictions can be shown up through residual analysis. To apply confidence intervals using the standard error of the residuals, a Gaussian distribution of residuals is required.

For the above reasons, four main residual analyses were undertaken. Tests were carried out to examine residual distributions for a) heteroscedasticity, b) trends in residuals, c) normality and d) mutual independence, and in most cases these criteria were met.

3.2.2 Confidence intervals

Confidence intervals are a convenient way to express average model uncertainty (Goddard et al. 2001), and are used in this study to display uncertainty around model predictions. The formula used to construct the confidence intervals is:

$$\hat{y} \pm t_{1-\alpha/2}(SE) \qquad (2)$$

where SE = standard error of residuals, defined by $s/\sqrt{n}$, where s = sample standard deviation, and n = sample size. Critical t-value tables can be used to find $t_{1-\alpha/2}$ values with df = n-2. For a sample of 50, as in this study, the $t_{1-\alpha/2}$ is 1.96 for a 95% confidence interval, and 2.58 for a 99% confidence interval.

Although sample size of each individual validation period was small (usually 9-10 seasons), cross-validation techniques in effect extended the sample size used for error evaluation to five times this size, as the model standard error used to construct confidence limits was calculated from all approximately 50 residuals from all five cross validation models.

3.2.3 Weighted average of predictions

If the separate models from this cross-validation exercise did display fairly similar equations, a reasonable assumption would be that any of the different period models could be taken forward as the model to predict future, unseen data, and the cross validation exercise is therefore only to extend the sample size to better estimate "true" model error. There is also precedent in the literature for using an average of a range of model predictions as the one prediction to be used (Georgakakos 2001, Mo 2003, Mullan and Renwick 1996, Sharma and Lall 2003, Sivillo et al. 1997, Stephenson et al. 1999). Indeed, ensemble model forecasting, where a mean or some other combination of model outputs is used as the actual prediction, is in regular use as a technique which hopefully minimises the errors associated with each of the individual models. However, the idea that the individual errors are minimised by using an ensemble only holds true if the individual models are independent and unbiased. An examination of dependence and bias in the data is undertaken in section 4.3.

Model skill for each separate cross-validated model was estimated by calculating the explained variance (E_v) for the 9 or 10 predicted-observed pairs in that model's validation data. Model skill varied between different individual period models, and so it was decided in this study to use a weighted average of all five period models, weighted on their skill levels (explained variance, calculated on the 9 or 10 residuals from each validation period), to go forward as the future model. There is some precedent for using the weighted average of model predictions in the literature (Tootle and Piechota 2004, Colman 2004, Zhang and Casey 2000). Tootle and Piechota (2004) found that an ensemble of three forecasts, with different predictors, combined into one weighted combination forecast, had better overall skill than any of the individual forecasts. More weight was given to more skilful forecasts and less weight to less skilful forecasts.

A weighted average of the five period models, then, is used in this study as the prediction for future applications of the model. For future model applications, all five period models would be run to give a prediction. A weighted average of these five predictions is then formulated, using the explained variance of each period (calculated on model validation for that period) as the weighting for that period. So periods where model validation gave a high explained variance meant that more trust was placed in the model, but those with lower validation goodness of fit were not necessarily excluded entirely, but gave a proportionally less weighted contribution to the final future model prediction.

The problem then arises in deciding what is the model error associated with this new ensemble model output. If we could assume independence between different models being averaged, and also that they are unbiased (ie. that their predictions are randomly scattered about the observed without bias), then averaging the models would result in a decrease in model error.

Taking the weighted average of the five individual cross-validated models, in this case, could be reflected in the estimate of the standard error by incorporating the individual weights and the number of components in the mean in the error calculation. Once the requirements of independence and non-bias have been satisfied (where possible), this is reflected in the following equation:

$$SE^{wgt} = \frac{\sqrt{\sum (C_i \times Var(\varepsilon_i))^2}}{\sqrt{n*}} \qquad (3)$$

where:

SE^{wgt} = the standard error associated with a weighted mean of model outputs

C_i = the weights of the individual models in the weighted mean

Var(ε_i) = the variance of the distribution of all the residuals from all the individual model validations

$n*$ = the number of components in the weighted average

The confidence interval could therefore be created about new predictions from the weighted mean of model outputs, by taking the standard error estimate from equation (3), and using it in equation (2).

This methodology would greatly decrease the magnitude of the confidence interval, over that of simply using equation (2) alone. The question remains, however, if the assumption of independence cannot be satisfied, what then is the *true* model error? In reality there is no simple answer to this question.

Due to possible bias and lack of independence of the predictions (see section 4.3 for more on this), it is felt that the confidence limits which take into account weightings and n, the number of components in the model, are probably unrealistically narrow. On the other hand, the confidence limits constructed using the model standard error calculated from the standard deviation of all 48 residuals are probably too wide for the averaged model, and unduly penalise it by attributing too much error to the model predictions. The true error estimate probably lies somewhere between these two confidence intervals. For the scope of this study, and for the risk environment of their likely application setting, the more conservative confidence limits (those constructed using equation 2) would be recommended for operational use.

3.3   Model skill

Models from this study were evaluated for how well the predicted values fit the observed data, as outlined above in relation to confidence intervals, but models were also compared with an already existing forecast, with persistence, with the long term hydro-climatological normal, and with random chance.

Residuals from the mean-as-predictor and from this study's models were compared using a paired t-test to see if one set of residuals was significantly smaller than the other.

To compare this study's prediction models with persistence, serial correlations between seasonal inflows and rainfall at different lags were examined.

Comparing model outputs to a random prediction was undertaken using randomisation testing of the predicted and observed pairs from each validation dataset for each model. Software was written to sample every possible pairing of the predicted and observed, and to test each sampled set of residuals tested for goodness of fit using explained variance, E_v. In this way, the chance of the particular model pairing occurring by random chance was estimated with an associated probability.

4. RESULTS

4.1 Validation test of predictive models

Predictive equations resulted from each of the forward stepwise regressions undertaken on each of the different model calibrations outlined in the previous section. For each dependent variable, for each season, five different predictive equations resulted - one for each of the different calibration/validation periods. The explained variance, E_v, was calculated on validation data for each period, and then averaged over all five periods for each dependent variable and season. These averaged goodness of fit scores were used to compare models predicting different dependent variables and different seasons.

A summary of these averaged goodness of fit scores, can be seen in Table 2. To enable some recognition of "better" and "worse" models, goodness of fit values higher than zero are highlighted. It should be noted that highlighting positive numbers is an arbitrary move, and only gives some initial estimation of which models are more likely to have predictive ability over others.

The main limitation of calculating goodness of fit on 9 or 10 year periods of data is that if a particular 9 or 10 year period does not contain much variability in the observed data, then the model is not truly being "tested". The explained variance measure goes some way to addressing this issue, in that it gives a better fit score for a model that has smaller residuals in relation to the variance contained in the observed values.

Table 2: Explained variance, E_v, as goodness of fit criterion, calculated on validation data and averaged over all five periods for each season and each dependent variable. Values greater than zero are in bold type

	Lake Pukaki inflows	Lakes Pukaki + Tekapo inflows	Hermitage rain days	Combination of rain gauges	Hokitika rainfall	Hermitage rainfall
Summer	**0.09**	**0.04**	**0.07**	-0.60	-0.20	-1.15
Autumn	-0.48	-0.56	**0.19**	**0.22**	**0.02**	**0.18**
Winter	**0.21**	**0.12**	-1.19	-0.10	-0.12	-1.76
Spring	-1.54	-1.57	**0.15**	**0.06**	-0.30	**0.10**

4.1.1   Choosing an initial subset of models for further testing

Further testing was required to ascertain what the predictive ability of the models were, and how well their predictions scored against other prediction models. However, some models seemed to be performing badly under all validation periods, and a decision was made to distinguish between models that were unlikely to offer any predictive ability, based on their E_v, and those which should be retained at this point, for further testing.

A randomisation test was conducted, whereby the explained variance of repeated random orderings of predicted values (matched with observed values) was calculated. All possible pairings of predicted and observed were sampled. A randomisation distribution of explained variance, E_v, is therefore obtained, and the proportion of this distribution that is greater than the calculated explained variance is obtained,

indicating the likelihood of the particular explained variance calculated on validation data from this study's models occurring by random chance.

The explained variances and their associated significance levels (lower values = significantly unlikely to occur by random chance) can be seen in table 3. Those explained variances that are greater than zero and significant at the 95% level are highlighted. An assumption is made at this stage that explained variances that are significant at the 95% level, but less than 0 are unlikely to have forecasting ability, and are therefore disregarded at this stage.

Table 3: Explained variances and their associated significance levels (p-values) for every period model and every dependent variable and season, as obtained from randomisation testing. Those explained variances that are greater than zero and significant at the 95% level are highlighted.

	Dependent variable	Period 1		Period 2		Period 3		Period 4		Period 5		Average periods 1-5	
		Expl var	p	Expl var	p	Expl var	p	Expl var	p	Expl var	p	Expl var	p
Summer	Lake Pukaki inflows	0.23	0.10	0.03	0.27	-0.10	0.48	0.00	0.07	**0.55**	**0.00**	**0.14**	**0.00**
	Lakes Pukaki + Tekapo inflows	0.26	0.06	0.04	0.24	-0.07	0.48	-0.51	0.09	**0.41**	**0.02**	**0.03**	**0.00**
	Hermitage rain days	0.54	0.06	**0.24**	**0.05**	0.41	0.06	-0.65	0.05	-0.42	0.40	**0.02**	**0.00**
	Rainfall combination	-0.08	0.12	0.14	0.14	-0.03	0.26	-2.81	0.29	-0.27	0.25	-0.61	0.05
	Hokitika rainfall	-0.15	0.34	**0.38**	**0.02**	-0.29	0.36	**0.13**	**0.05**	-0.23	0.23	-0.03	0.02
	Hermitage rainfall	0.18	0.07	0.04	0.23	0.00	0.42	-10.89	0.43	-0.13	0.08	-2.16	0.13
Autumn	Lake Pukaki inflows	0.13	0.14	-0.48	0.38	-0.16	0.22	-0.92	0.44	-0.06	0.33	-0.30	0.11
	Lakes Pukaki + Tekapo inflows	0.15	0.12	-0.34	0.26	0.14	0.16	-0.32	0.37	-0.51	0.44	-0.18	0.07
	Hermitage rain days	**0.19**	**0.05**	0.08	0.22	0.12	0.17	0.18	0.10	-0.05	0.36	**0.11**	**0.01**
	Rainfall combination	**0.39**	**0.02**	**0.40**	**0.03**	-0.05	0.08	0.15	0.08	-0.03	0.17	**0.17**	**0.00**
	Hokitika rainfall	-0.21	0.31	0.20	0.08	-0.01	0.17	0.00	0.36	0.17	0.11	**0.03**	**0.04**
	Hermitage rainfall	0.15	0.09	0.16	0.11	-0.07	0.11	**0.43**	**0.02**	0.16	0.12	**0.16**	**0.00**
Winter	Lake Pukaki inflows	**0.24**	**0.03**	**0.34**	**0.01**	-0.08	0.57	0.12	0.16	**0.29**	**0.03**	**0.19**	**0.00**
	Lakes Pukaki + Tekapo inflows	-0.46	0.42	**0.65**	**0.02**	**0.55**	**0.01**	-0.44	0.23	0.08	0.14	**0.08**	**0.00**
	Hermitage rain days	0.00	0.10	-2.12	0.05	-0.75	0.09	-1.90	0.51	-0.34	0.23	-1.02	0.01
	Rainfall combination	0.23	0.07	-0.11	0.51	-0.22	0.40	0.08	0.08	-0.26	0.16	-0.05	0.02
	Hokitika rainfall	-0.17	0.35	-0.12	0.15	0.07	0.23	-0.02	0.39	-0.60	0.21	-0.17	0.06
	Hermitage rainfall	-2.35	0.59	-0.43	0.27	0.23	0.06	-0.16	0.32	-0.38	0.36	-0.62	0.28
Spring	Lake Pukaki inflows	-6.15	0.27	-1.08	0.35	0.29	0.07	0.09	0.19	-2.31	0.10	-1.83	0.05
	Lakes Pukaki + Tekapo inflows	-6.07	0.38	**0.35**	**0.02**	0.13	0.15	-0.13	0.24	-0.45	0.49	-1.23	0.19
	Hermitage rain days	-0.54	0.26	**0.18**	**0.04**	**0.21**	**0.03**	0.20	0.17	0.11	0.12	**0.12**	**0.00**
	Rainfall combination	-0.48	0.48	0.10	0.17	0.15	0.13	0.18	0.13	-0.20	0.34	-0.05	0.07
	Hokitika rainfall	-0.37	0.31	**0.07**	**0.02**	0.10	0.18	0.14	0.16	-0.31	0.64	-0.08	0.02
	Hermitage rainfall	-0.11	0.46	0.16	0.14	**0.25**	**0.05**	0.26	0.08	-0.06	0.31	**0.10**	**0.01**

None of the 24 models were significantly better than chance over all five individual period models. However, when an average was taken, and the significance calculated on the larger sample size that this involved, 11 of the models had a positive explained variance that was significantly better than chance at the 95% level.

A subset of 11 models were therefore chosen as those models that had significant (at the 95% level) positive explained variance, and these models were denoted the "initial subset of models", and taken forward for closer examination.

4.2   *SOI and PDO as separate predictors*

As was discussed in section 3.1, two separate continuous format models were calibrated in this study, for each dependent variable (and season). One of these included SOI and PDO as predictors, and one did not. All models in the initial subset of 11 models were re-calibrated, to ascertain if inclusion of these two separate predictors would improve model skill.

Models from the SOI and PDO model calibration process were put through the same cross-validation procedure as the original models, and validated in the same way.

It was found that no model showed consistent improvement from the addition of these variables. Explained variance was invariably worse for each of the models calibrated using these variables as input variables. This result is discussed further the summary and conclusions section.

4.3   Error evaluation

Before further error evaluation was carried out, residual analysis was undertaken that showed that the sets of model residuals were homoscedastic, approximately normally distributed, unbiased, and lacking in significant trends.

Having ascertained a minimal apparent bias in the models and an approximately Gaussian distribution of residuals, the question of independence between the five sets of residuals from the five period models to be used in the weighted average can be addressed. As the five period models were calibrated from overlapping sets of data, it is fairly certain that they will be dependent. This issue is also examined in Table 4, where the correlation matrix between sets of (9 or 10) residuals from two period models is shown.

It can be seen from Table 4 that, although most of the sets of residuals are not related to each other, and therefore satisfy the requirement of independence, there are one or two that have high correlations. These are between period 1 and 2 in the Lake Pukaki winter inflows models, and between periods 1 and 4, 2 and 4, and 1 and 3 in the Rainfall combination autumn models, and this is typical of other sets of residuals tested.

Table 4: Correlation matrices for residuals from each of 5 individual period models that will go into making up the weighted average model to be used for future predictions, for a) Lake Pukaki inflows for winter, and b) Rainfall combination for autumn.

a) Winter Lake Pukaki inflows residual correlations

	Period 1	Period 2	Period 3	Period 4	Period 5
Period 1	1.00	0.52	0.40	-0.41	0.07
Period 2		1.00	-0.12	-0.21	-0.06
Period 3			1.00	-0.08	0.10
Period 4				1.00	0.07
Period 5					1.00

b) Autumn Rainfall combination residual correlations

	Period 1	Period 2	Period 3	Period 4	Period 5
Period 1	1.00	0.16	-0.57	0.72	0.40
Period 2		1.00	-0.13	0.62	-0.18
Period 3			1.00	-0.33	0.02
Period 4				1.00	0.00
Period 5					1.00

Owing to probable lack of independence, it is felt that the confidence limits which take into account weightings and n, the number of components in the model, are probably unrealistically narrow. On the other hand, the confidence limits constructed using the model standard error calculated from the standard deviation of all 48 residuals are probably too wide for the averaged model, and unduly penalise it by attributing too much error to the model predictions. The true error estimate probably lies somewhere between these two confidence intervals. Future work, when more data becomes available, may lead to a more accurate representation of model error. For the scope of this study, however, and

for the risk environment of their likely application setting (their application in the hydro-electric generation management of the river and lake levels means that the implications of non-conservative application of the predictions are significant), the more conservative confidence limits would be recommended for operational use.

4.4   Model validation results

Examples of validation tests are displayed in figures 2-7 for individual period models for 5 models from the initial subset. It should be noted that, although the validation tests shown in the following graphs form one 9 or 10 year period, the error used to construct the 95% confidence limits is formulated from the standard error from the residuals from validation data from all 5 period models (ie. 48 years). For the model predicting winter Lake Pukaki inflows, all five period model's validation tests are shown on one graph (Figure 2), to give an overall picture of the variability in model predictions from model to model. The other graphs only show one nine or ten year period, to enable clearer inspection.
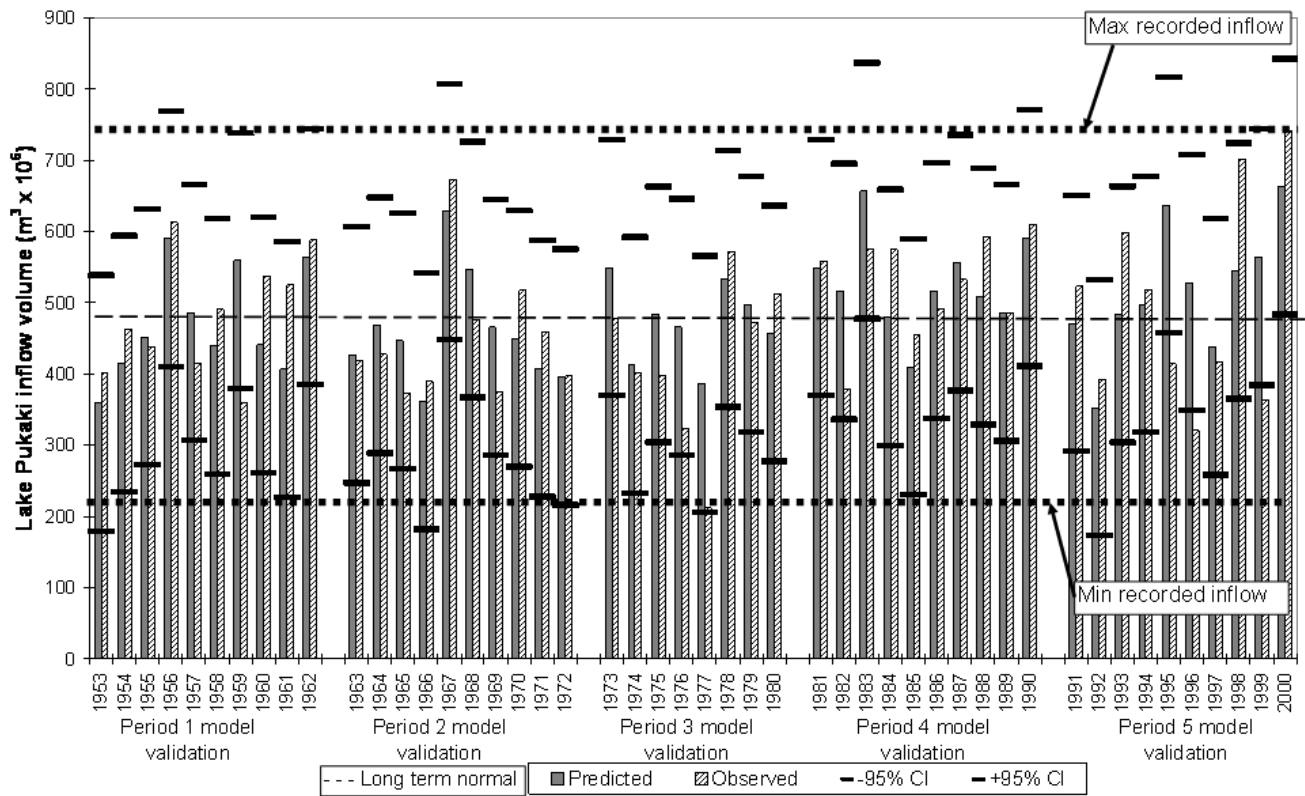
Figure 2: Validation data for all five period models predicting winter inflow total into Lake Pukaki. Predicted inflows with 95% confidence limits shown, and observed inflows, for winter season for the years 1953 to 2000. The shaded area shows all probable inflows for winter from an 80 year history.
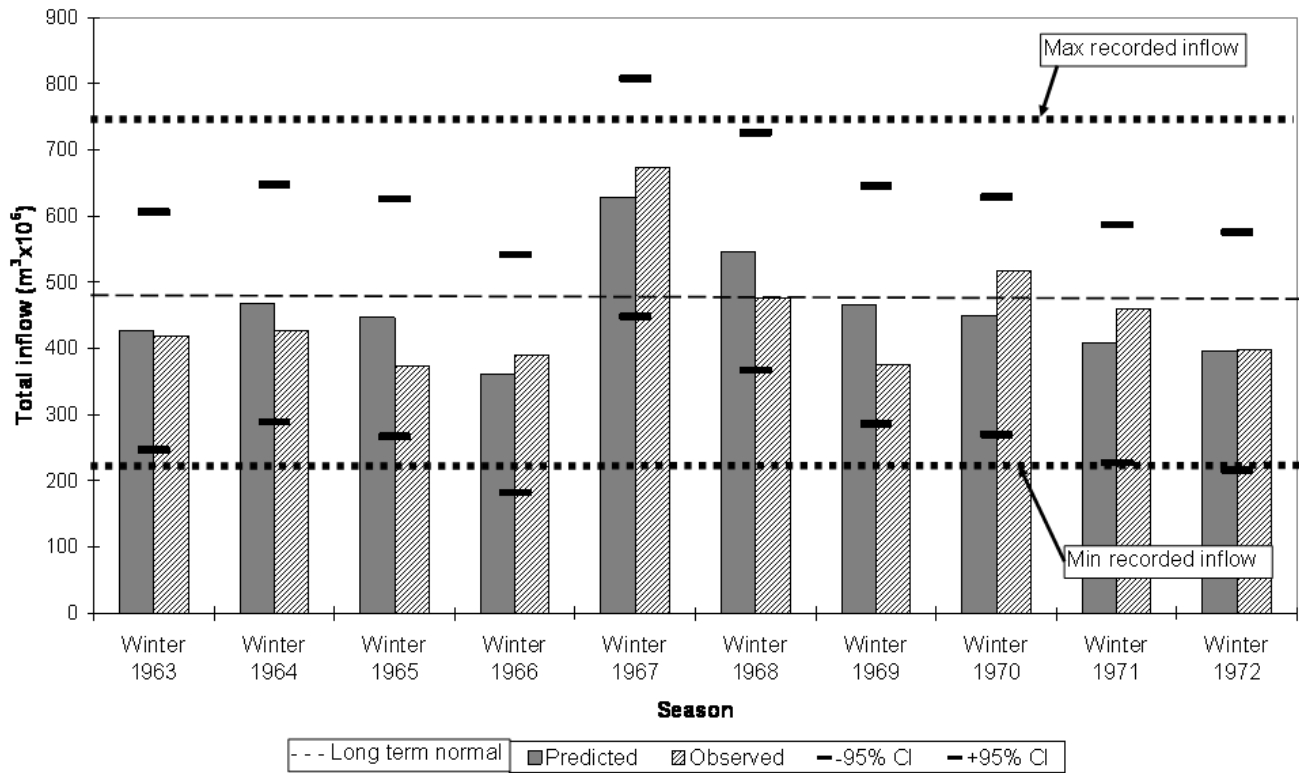
Figure 3: Validation data for period 2 graph for model predicting winter inflow total into Lake Pukaki. Predicted inflows with 95% confidence limits shown, and observed inflows, for winter season for the years 1963 to 1972. The bold dotted lines show all probable inflows for winter from 80 year history.
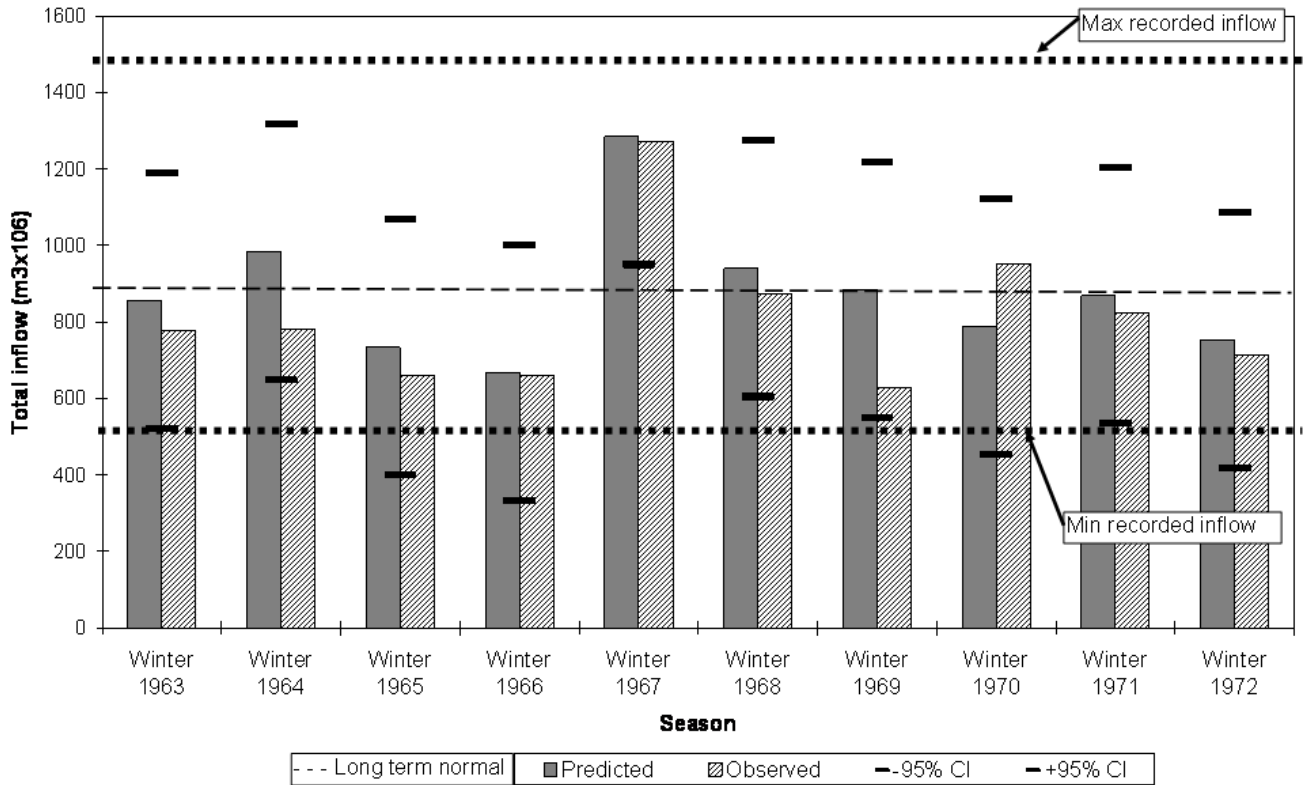
Figure 4: Validation data for period 2 graph for model predicting combined winter inflow total into Lakes Pukaki and Tekapo. Predicted inflows with 95% confidence limits shown, and observed inflows, for winter season for the years 1963 to 1972. The bold dotted lines show all probable inflows for winter from 80 year history.
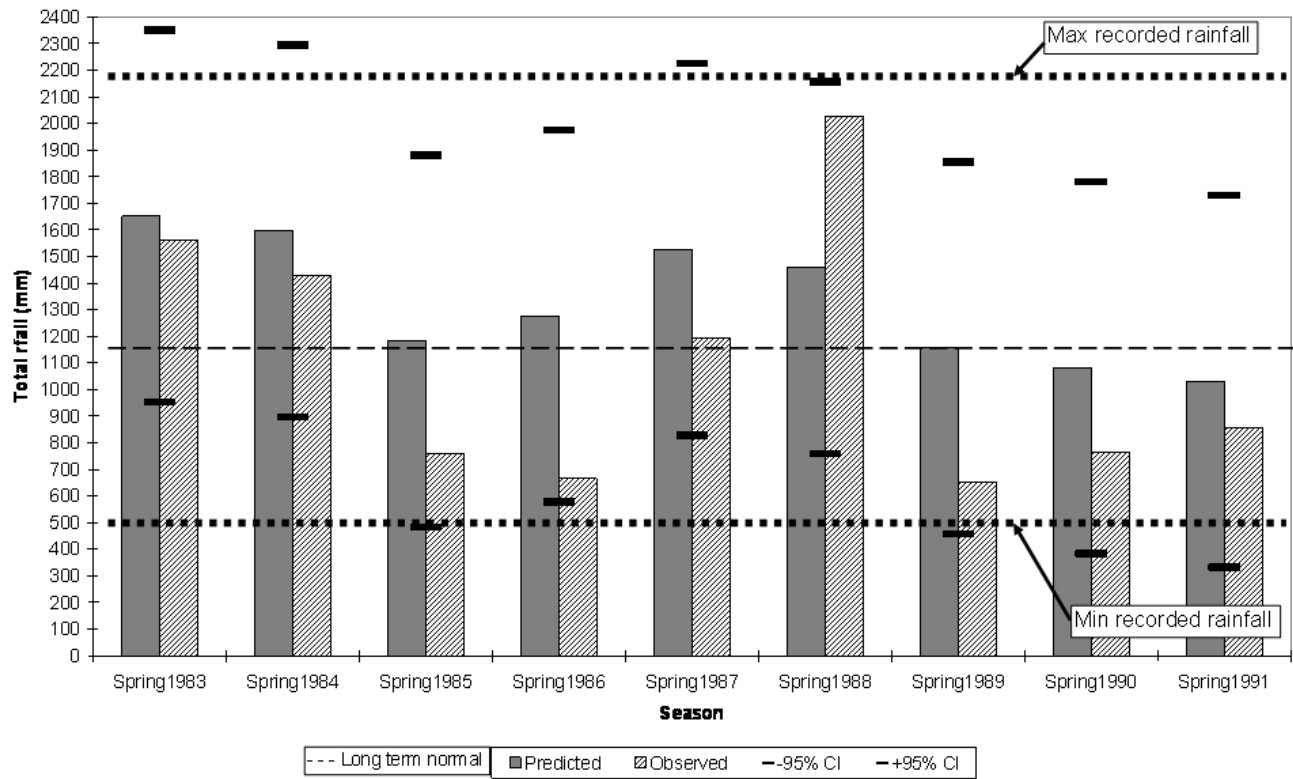
Figure 5: Validation data for period 4 graph for model predicting spring rainfall total at the Hermitage. Predicted rainfall with 95% confidence limits shown, and observed rainfall, for spring season for the years 1983 to 1991. The bold dotted lines show all probable rainfall for spring from 80 year history.
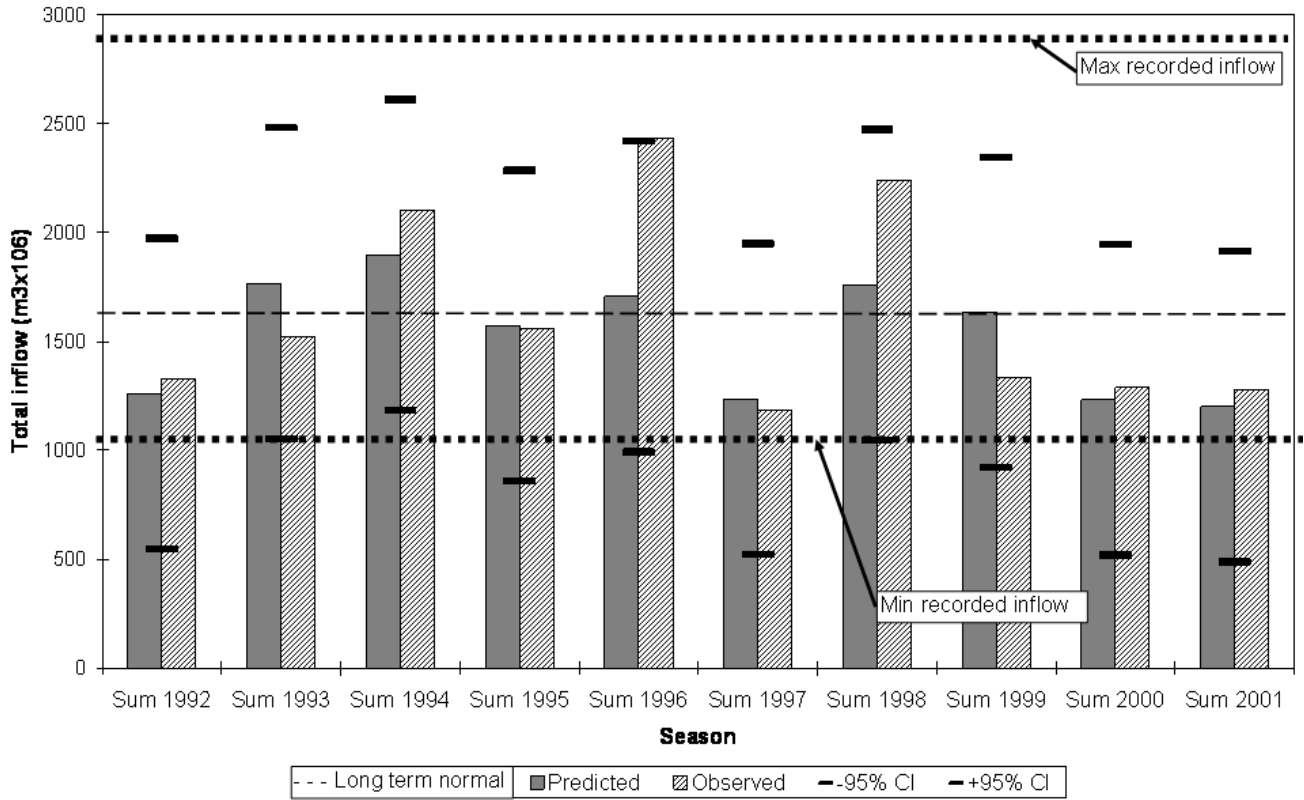
Figure 6: Validation data for period 5 graph for model predicting summer Lake Pukaki total inflows. Predicted inflows with 95% confidence limits shown, and observed inflows, for summer season for the years 1992 to 2001. The bold dotted lines show all probable inflows for summer from 80 year history.
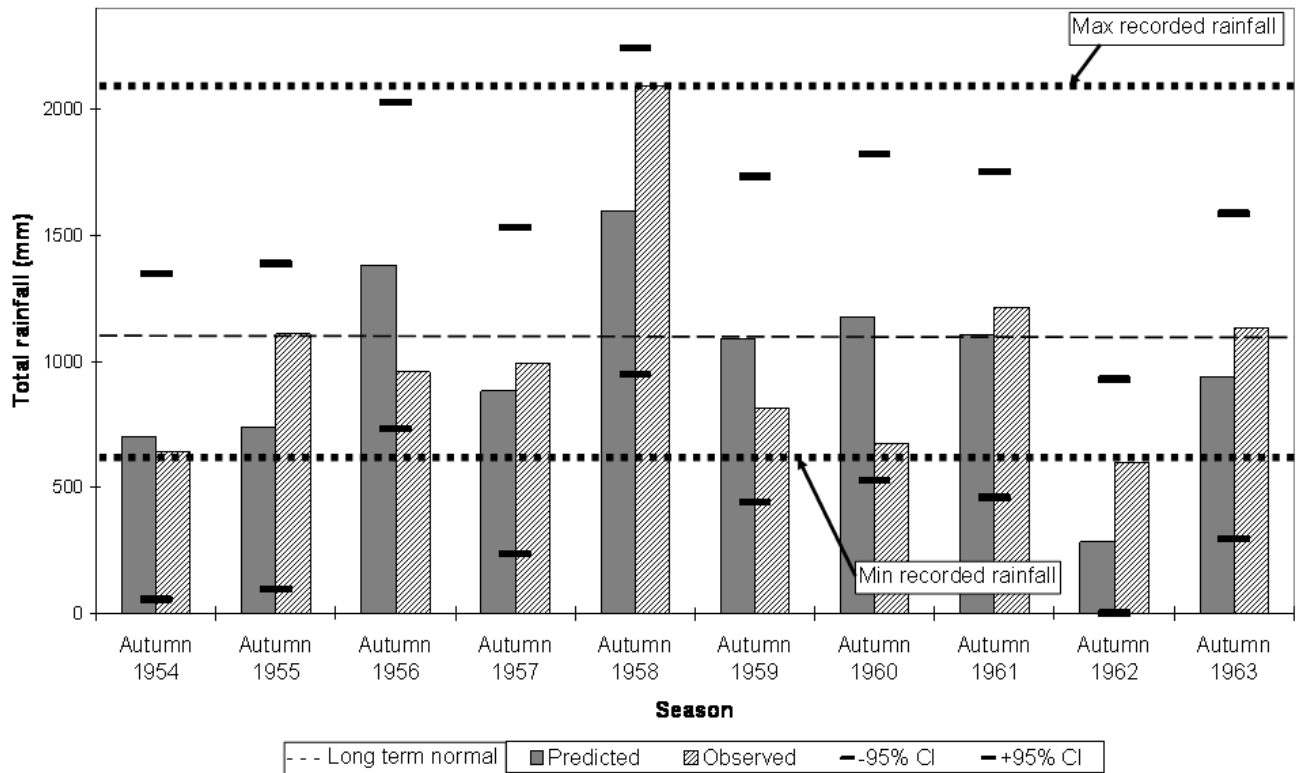
Figure 7: Validation data for period 1 graph for model predicting total rainfall at a combination of Southern Alps rain gauges for autumn. Predicted rainfall with 95% confidence limits shown, and observed rainfall, for autumn season for the years 1954 to 63. The bold dotted lines show all probable rainfall for autumn from 80 year history.

4.5   Comparison of models with previously used forecast tools in the Waitaki catchment

4.5.1   Comparison with long term normal inflows or rainfall

The primary seasonal forecasting tool used in the management of the Waitaki river hydro scheme is the use of the long term normal as a predictor. The "normal" (or long term average) inflows or rainfall that can be expected in the catchment in a season is the most basic model available, and has been the most widely applied model in hydro catchment management in the Waitaki catchment thus far.

Comparisons are made between the long term normal, used as a predictor, and the predictions from the models in this study. In 9 out of the 11 models examined, the standard deviation of model residuals was smaller for the models from this study, indicating that the model error was smaller than for the mean-as-predictor.

The residuals from the models calibrated in this study were compared with the residuals from the mean-as-predictor using a paired t-test. The paired t-test was conducted on sets of residuals for all seasons, as well as for seasons that had been divided into wet, average, and dry portions. This distinction was made because it is acknowledged that being able to predict a normal season is not of much benefit to hydro managers, but being able to predict a dry or a wet season was critical. Paired t-tests, were conducted where the observed inflows (or rainfall/raindays) for the season being predicted were more than 0.5 standard deviations below or above the mean: ie. a dry or a wet season.

The model predicting Lakes Pukaki + Tekapo inflows for winter had residuals that were significantly smaller than the mean-as-predictor. In all other cases the residuals were found to be not significantly

different from each other at the 95% level, indicating that many of the models from this study were not significantly better at forecasting rainfall and inflows than the long term mean, taken over the whole data range.

However 10 of the 12 models were significantly better than the mean-as-predictor in dry or wet years. Significance levels from the paired t-tests in table 5 show that in 10 cases the model residuals from the models from this study were significantly different (always smaller) than the residuals from the mean-as-predictor.

Table 5: Significance levels for paired t-test between residuals from models from this study and the long term mean as predictor, in dry or wet seasons only. Ticks indicate where the residuals from this study's models are significantly different (smaller) than the residuals from the long term mean-as-predictor.

		p-values	Significantly smaller than mean-as-predictor
Summer	Lake Pukaki inflows	0.04	√
	Lakes Pukaki + Tekapo inflows	0.03	√
	Hermitage rain days	0.00	√
Autumn	Hermitage rain days	0.02	√
	Combination of rain gauges	0.02	√
	Hokitika rainfall	0.10	
	Hermitage rainfall	0.01	√
Winter	Lake Pukaki inflows	0.02	√
	Lakes Pukaki + Tekapo inflows	0.01	√
Spring	Hermitage rain days	0.20	
	Combination of rain gauges	0.02	√
	Hermitage rainfall	0.05	√

4.5.2   Persistence – serial correlation in inflows and rainfall

Persistence of inflows and rainfall was examined for its efficacy in prediction at any lag. No combination of adjacent seasons or year to year seasonal serial correlation yielded any significant lags. Persistence, or serial seasonal correlation, in inflows and rainfall was therefore shown to have no significant skill in predicting inflows and rainfall in the Waitaki catchment.

4.5.3   Comparison with random chance

The potential for achieving adequate goodness of fit in validation data through random chance, especially when many model validations are being performed, has to be considered. This was tested for in this study through the use of a randomisation test on the calculated explained variance, E_v, of pairings of predicted and observed for validation data for every time period used in the cross-validation process. The results of this randomisation test can be seen in table 3.

Results for the continuous format models showed that 17 out of a total of 24 models calibrated had pairings of predicted and observed that were significantly unlikely to occur by random chance (at the 95% level, and averaged over all 5 cross-validated periods). Of these, 8 had explained variances greater than 5%. This has been chosen as an arbitrary level in this study as showing some potential predictability. These are summarised in table 6.

4.6   Constituent variables

When the contribution of each constituent ocean-atmosphere state variable to each PC was ascertained, and then the resulting contribution of each PC to the final predictive equation, the result in

most instances was that no predictor variable stood out as contributing significantly more than other independent variables to any of the predictive equations. This is in part due to the orthogonal nature of the principal components, so that the methodology chosen in this study has contributed to the difficulty in finding any real world explanations for the predictability displayed in the model equations. However, a summary of variable contributions to predictions is discussed below.

Table 6: Summary of explained variances and associated p levels for the 8 models with explained variance of greater than 5%, averaged over all 5 cross-validated periods

Dependent variable	Average periods 1-5	
	E_v	p
Summer Lake Pukaki inflows	0.14	0.00
Autumn Hermitage rain days	0.11	0.01
Autumn Rainfall combination	0.17	0.00
Autumn Hermitage rainfall	0.16	0.00
Winter Lake Pukaki inflows	0.19	0.00
Winter Lake Pukaki + Tekapo inflows	0.08	0.00
Spring Hermitage rain days	0.12	0.00
Spring Hermitage rainfall	0.10	0.01

The predictor variable that is chosen most often and ranked most highly in the model calibration process for all models is SST5. This is the Mullan sea surface temperature index region 5, which is SSTs averaged over an area that is north of New Zealand and east of Australia. The predictor variable chosen second most often is Sunspot number, closely followed by sea level pressure at 35 degrees South latitude, 325 degrees longitude, a point off the east coast of South America. Next most important is the QBO, followed by SST8 (the SST index located in the Indian Ocean), and then the sea level pressure at 25 degrees South latitude, 285 degrees longtitude, a point off the West Coast of South America.

Models predicting different seasons of the same dependent variable utilised very different independent variables to predict each different season. The relative importance of ocean-atmosphere state variables as predictors for each season's predictive models are listed in table 7 for the highest ranking 20 variables for every season. For example, a variable that had one 20th place ranking for one model would score 1. A variable that had top ranking for three models would rank 20 points x 3 models = 60.

SST5 comes out clearly as the strongest predictor for winter inflow prediction models. All winter models in the best model set have SST5 as the strongest predictor, except the model predicting Lake Pukaki inflows, which uses it as the fourth strongest predictor. Other important winter inflow predictors include SLP and 700hPa geopotential height locations to the north-east of New Zealand, to the south-west of Australia, and over South America. Also included is the Trenberth indice MZ2, which is the difference in air pressure between Gisborne and Invercargill, representing a north-south pressure gradient over the Waitaki River catchment, or zonal flow across the South Island.

Most summer models utilise the QBO as their most important predictor. This includes all the inflow prediction models in the summer initial models subset. Their next most important predictors are SLP and 700hPa geopotential height gridpoints over the South American land mass. However, the summer Hermitage raindays prediction models in the best models subset, do not include QBO amongst their 20 most important predictors. They use Hermitage rain days from last season (ie. persistence) and the 700hPa geopotential height at 25 degrees latitude and 305 degrees longitude (a point over Brazil) as their two most relevant predictors.

Table 7: Sum of individual rankings (1-20) of the top 20 constituent variables for each model in the initial models subset, divided into seasons. eg. A variable that had one 20th place ranking for one model would score 1. A variable that had top ranking for three models would rank 20 points x 3 models = 60.

Summer	PKIinflows	log(PKIinflows)	PKI+TEKinflows	log(PKI+TEKinflows)	Hermraindays	log(Hermraindays)	Summer total
QBOanom	20	20	20	20			80
MSLP-35.325	17	13	19	19			68
MSLP-45.235	9	8	16		10	17	60
SST8	13	16	17	12			58
SST7	8	18	13	13			52
700hpa-35.345		19	8	18			45
700hpa-15.315		15	12	17			44
700hpa-35.275		14	15	14			43
Sunspotnoanom	18	6	18				42
MSLP-5.315	3	9		1	14	15	42
MSLP-45.255	5	7	11	10	6	3	42
MSLP-55.155	10	12	4	15			41
700hpa-55.145		17	7	16			40
Hermraindayanoms					19	20	39
700hpa-25.305					20	19	39
MSLP-5.185	14	11	1	11			37
700hpa-25.265					18	18	36
MSLP-35.255					17	16	33
MSLP-45.95	6	3	14	9			32
MSLP-25.245					15	14	29

Autumn	Hermraindays	log(Hermraindays)	Rfallcombo#1	log(Rfallcombo#1)	Hokitikarfall	log(Hokitikarfall)	log(Hokrfall)+SOI&IP...	Hermitagerfall	log(Hermrfall)	Autumn total
MSLP-25.285		5	19	18	20		7	17	18	104
SST7	19		17	19	14			14	15	98
SST5			18	16		20		20	20	94
SST6	20	10	11	12		19		7	7	86
MSLP-55.335			12	3	12	17		18	17	79
700hpa-25.275			15	15	9	8		16	16	79
SST1			20	20		11		12	14	77
700hpa-25.295			16	17	8			13	13	67
Sunspotnoanom		13	9	5	11			6	9	53
SST2	15	12					12	8	6	53
MSLP5.295			7	8	17	12		3	4	51
700hpa-15.195			14	14				11	12	51
SST8		19	2	11			17			49
700hpa-15.215			13	7				15	11	46
Hokitikatempanom		1	6					19	19	45
MSLP-5.255			10	13			11	5	5	44
MSLP-25.205		3	3	1	15			9	8	39
MSLP-45.205		7	8	2				10	10	37
MSLP5.305					19	15	1			35
MSLP-5.235	18							15		33

Winter	PKIinflows	log(PKIinflows)	PKI+TEKinflows	log(PKI+TEKinflows)	PKIinflowsminusfloods	log(PKIminuslfoods)	Winter total
SST5	17	20	20	20	20	20	117
MSLP-35.335		19	18	19	15	19	90
700hpa-55.75	5	14	19	18	10	14	80
MSLP-35.325		16	9	16	18	16	75
Sunspotnoanom		17	5	17	13	17	69
700hpa-15.305		18		15	16	18	67
700hpa-25.175	14		14	9	19		56
MSLP-55.285	7	6	11	10	9	6	49
Hokseasrfallanom	11	2	6	5	17	2	43
MSLP-5.225		12		14		13	39
MSLP-25.205	18		16	3			37
MZ2	19		17				36
700hpa-35.345		5		12	14	5	36
MSLP-35.245		13	1		8	12	34
700hpa-25.195	20		12		1		33
MSLP-25.245		10		6	6	10	32
700hpa-25.185	15		3	2	12		32
QBOanom		9		13		9	31
MSLP-45.205		15				15	30
700hpa-55.215	3		13	11			27

Spring	Herm rain days	Hermitage rainfall	log(Hermrfall)	Spring total
700hpa-35.295	15	20		49
700hpa-25.285	20	18		38
700hpa-35.165		19		38
700hpa-35.285	17	17		34
MSLP-25.175		12		32
MSLP-35.165		13		30
MSLP-35.295	19	11		30
700hpa-45.345	2	16	12	30
700hpa-25.115		14		29
700hpa-25.175		10		28
700hpa-35.325	13	15		28
700hpa-25.125		9		25
700hpa-25.275	18	3		21
MSLP-45.235			20	20
MSLP-45.245			19	19
700hpa-45.225			18	18
PKIsiminflowanom			17	17
700hpa-25.315	16	1		17
MSLP5.185			16	16
Sunspotnoanom			15	15

All seasons	Total total
SST5	211
Sunspotnoanom	179
SST7	150
MSLP-35.325	143
QBOanom	126
SST8	114
MSLP-25.285	114
700hpa-55.75	103
700hpa-35.345	102
700hpa-25.275	100
MSLP-35.335	90
MSLP-25.205	87
SST6	86
MSLP-45.235	85
700hpa-25.175	84
Hokitikatempanom	79
MSLP-55.335	79
SST1	77
MSLP-45.205	67
700hpa-15.305	67

All the autumn models in the initial subset are rainfall or rain days prediction models. There are no Autumn inflow prediction models in the initial models subset. The best predictor variables utilised by these rainfall and rain day prediction models are sea level pressure and 700hPa geopotential height locations over South America and at high latitudes to the south-west of Australia. Hokitika rainfall and temperature (from last season) also feature as important predictors. Sea surface temperature indices also featured strongly as predictors of Autumn rainfall and raindays. Important SST indices were SST1, SST5, SST6, SST7, and SST8.

The predictor variables which appear to be important for predicting spring rainfall and rain days are all sea level pressure and 700hpa geopotential height locations in the mid-latitudes. Longitudinally, these locations are found just north of New Zealand, in the mid-Pacific, and over South America.

5.   SUMMARY AND CONCLUSIONS

This research has focussed on creating season ahead forecasts of lake inflows and rainfall, predominantly in a single river catchment, the Waitaki river. Although seasonal hydro-climatic research internationally is moving toward greater spatial resolution of forecast models and tailoring forecast products more to end user needs, this area of research is still in its infancy and has not been successfully applied in New Zealand before.

Forecast models were created to be easily applied operationally in the management of a hydro-electricity generation setting, and this has been achieved. The models calibrated in this study are

currently being tentatively applied in management of the hydro-electricity scheme in the Waitaki river catchment. A number of predictive models could be said to predict better than random chance and better than the long term mean as a predictor. One model, that predicting winter Lakes Pukaki + Lake Tekapo combined inflows, predicts all magnitudes of the predictand better than random chance and better than the long term mean as a predictor. Confidence limits around most model predictions in this study offer some skill when compared with the range of all probable inflows or rainfall (based on the 80 year recording history in the catchment).

Estimation of model error was a significant part of this study, and is an important part of the creation of tools to be used in an environment where the consequences of relying on poor predictions are large. Software was developed to apply randomisation testing to observed-predicted pairs, and this testing allowed the conclusion that many of the models had significantly more skill than random chance. Attention was given to the variability in predictive skill between cross-validated models, and methods for defining the model, from the five cross-validated models developed, to be used operationally against future data. Probabilistic tercile formats were intentionally avoided, as being not operationally applicable. The use of a continuous predictand with confidence limits applied is unusual in seasonal climate forecasting, but there is beginning to be some precedent for this format, especially from researchers seeking to create tools from which some operational decisions can be made.

The winter Lake Pukaki inflow models can be said to be those with the strongest predictive relationships in this study. Winter was the season predicted with most skill, with models predicting autumn dependent variables yielding models with almost as much skill as winter models. Spring and

summer predictions generally were less skilful than those for winter and autumn. Inflows could be predicted with some skill in winter and summer, but not rainfall. The reverse was true for autumn and spring, where rainfall could be predicted with some skill, but not inflows. If rainfall can be forecast with skill over inflows in some seasons, there is merit in using these rainfall forecasts as a proxy for inflows in hydro storage management in an all year round hydro management scheme. Implicit in these differences was the issue of snowmelt in a large snow-covered catchment. However, there is temporal and spatial variation in the predictability, and therefore using forecasts of different dependent variables at different times of the year is a potential means to improve the information available to hydro managers on the Waitaki river.

A significant reason for comparing model outcomes with random chance and the mean-as-predictor is that comparison with previous international forecasts is difficult due to varying spatial scale of forecasts, different predictands, prediction format, and techniques of model validation. The amount of variance in the rainfall record that is potentially predictable, given perfect models, is stated by various researchers to be 30% or less (Madden et al., 1999), or approximately 50% (Renwick, 2003), in New Zealand. The rest of the variance in the rainfall record is deemed to be chaotic, and therefore unpredictable. Mullan and Renwick (1996) captured up to 6% of the variance in autumn and winter rainfall in the north of the South Island on independent trials, but found forecast models had little skill in the central and lower South Island. Researchers have explained over 50% of the variance in seasonal precipitation over eastern Australia (Cordery, 1999), and 35% of the variance in seasonal rainfall models in the Pacific Islands (He and Barnston, 1996). The same figure for studies in South America is up to 70% (Hastenrath and Greischar, 1993), 30% in South Africa (Hastenrath et al., 1995), and 15 to 25% in the eastern Mediterranean. Techniques vary between methods of seasonal

forecasting of rainfall and validation techniques, so it is difficult to directly compare these explained variance figures.

It cannot be said that all models consistently predict either wet or dry seasons with more skill – ie it cannot be concluded from this study that it is easier to predict a low inflow season or a high inflow season, or a low rainfall season or a high inflow season.

The lack of stability of relationships found between land-ocean-atmosphere variables and seasonally lagged inflows and rainfall meant that cross-validation model skill varied considerably. The highest scoring models calibrated in this study are also the ones that display the lowest variability amongst goodness of fit scores for all five period models. The conclusion is that these models with low variability of skill between period models (and high skill scores over all) will demonstrate a far greater robustness in the predictive relationships they define.

A major theme in this research has been to attempt to predict seasonal inflows and rainfall on a finer spatial scale than has been done before in New Zealand. It was not generally found that predicting larger macro-catchments provided any increase or decrease in skill over predicting the dependent variables over a finer spatial area. The only possible exception to this was that of the model predicting Lakes Pukaki + Tekapo inflows for winter. The application here of statistical seasonal climate forecasting methods to river flows and to a single catchment, with some success, would lead to the conclusion that seasonal climate forecasting on single catchment scale, and focussed to end user needs, is possible with some skill.

Models predicting inflows and rainfall for different seasons in this study use very different sets of predictor variables to accomplish their seasonal predictability. This leads to the conclusion that inflows and rainfall, at least in a catchment with significant snow in the South Island of New Zealand, should be studied in the future either separately, or with the knowledge that their predictive mechanisms are likely to be very different. Similarly, predicting the same dependent variable but for different seasons led to different contributing variables. This also leads to the conclusion that different wider physical causative mechanisms are behind the predictability in different seasons, and they too, should be studied separately in any future research in this area. SST5 was found to have more relevance than any other predictor in predicting Waitaki river inflows and rainfall in any season.

The models calibrated with SOI and IPO included as predictor variables were invariably worse in their predictive skill than those without. In addition to this, SST3 and SST4 (those in the equatorial area of the Pacific ocean most influenced by the El Nino-Southern Oscillation) were the only SST areas that did not rank in the top 20 constituent variables for any of the initial subset of models, and MSLP and 700hPa geopotential height locations in the top 20 ranking of predictive variables for all models are all located in the mid-latitudes. There are no equatorial locations amongst them. The conclusion from these findings is that equatorial ocean-atmosphere state variables had a much smaller part to play in the predictability of these models than mid-latitude and local variables. ENSO related features were repeatedly poorly represented in important contributing variables to the models.

ACKNOWLEDGEMENTS

REFERENCES

Basher R, 1998. The 1997/98 El Nino Event: Impacts, Responses, and Outlook for New Zealand. In: *Ministry of Research, Science, and Technology*; **28**.

Benoit R, Pellerin P, Kouwen N, Ritchie H, Donaldson N, Joe P, Soulis ED, 2000. Toward the Use of Coupled Atmospheric and Hydrologic Models at Regional Scale. *Monthly Weather Review* **128**:1681-1706.

Breiman L, Spector P, 1992. Submodel selection and evaluation in regression: The X-random case. *International Statistical Review* **60**:291-319.

Cattell RB, 1966. The scree test for the number of factors. *Multivariate Behavioral Research* **1**:245-76.

Cayan DR, Redmond KT, Riddle LG, 1999. ENSO and hydrologic extremes in the western United States. *Journal of Climate* **12**:2881-2893.

Chakraborty K, Mehrotra K, Mohan CK, Ranka S, 1992. Forecasting the behaviour of multivariate time series using neural networks. *Neural Networks* **5**:961-970.

Colman AW, 2004. Combined statistical/dynamical forecast of 2004 season rainfall in the Sahel and other regions of tropical North Africa using observed ocean and atmosphere information from up to June 2004. *Experimental long lead forecast bulletin* **13**.

Cordery I, 1999. Long Range Forecasting of Low Rainfall. *International Journal of Climatology* **19**:463-470.

De Lisle JF, 1956. Secular Variations of West Coast Rainfall in New Zealand and their Relation to Circulation Changes. *New Zealand Journal of Science and Technology*:700-715.

Dettinger MD, Cayan DR, Redmond KT, 1999. United States streamflow probabilities based on forecasted La Nina, winter-spring 2000. *Experimental long-lead forecast bulletin* **8**:57-61.

Drosdowsky W, Chambers LE, 2001. Near-Global sea surface temperature anomalies as predictors of Australian seasonal rainfall. *Journal of Climate* **14**:1677-1687.

Eshel G, Cane MA, Farrell BF, 2000. Forecasting Eastern Mediterranean Droughts. *Monthly Weather Review* **128**:3618-3630.

Feddersen H, Navarra A, Ward MN, 1999. Reduction of model systematic error by statistical correction for dynamical seasonal predictions. *Journal of Climate* **12**:1974-1989.

Fennessy MJ, Shukla J, 2000. Seasonal Prediction over North America with a Regional Model Nested in a Global Model. *Journal of Climate* **13**:2605-2627.

Fitzharris BB, 1992. The 1992 Electricity Crisis and the Role of Climate and Hydrology. *New Zealand Geographer* **48**:79-83.

Fitzharris BB, McAlevey B, Richards K, 1998. Improving Forecasts of Seasonal Inflows into Waitaki Hydro Lakes - Stage 1: A review and proposal using Neural Networks and Information about Seasonal Snow. In: *University of Otago Consulting Group*; **45**.

Francis RC, Hare SR, 1994. Decadal-scale regime shifts in the large marine ecosystems of the Northeast Pacific: a case for historical science. *Geophysical Research Letters* **27**:731-734.

Francis RICC, Renwick JA, 1998. A regression-based assessment of the predictability of New Zealand Climate anomalies. *Theoretical and Applied Climatology* **60**:21-36.

Georgakakos A, 2001. Probabilistic and ensemble forecasting. *Journal of Hydrology* **249**:1.

Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA, 2001. Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology* **21**:1111-1152.

Greischar L, Hastenrath S, 2000. The Rainy Seasons of the 1990s in Northeast Brazil: Real-Time Forecasts and Verification. *Journal of Climate* **13**:3821-3826.

Hastenrath S, Greischar L, 1993. Further work on the prediction of northeast Brazil rainfall anomalies. *Journal of Climate* **6**:743-758.

Hastenrath S, Greischar L, van Heerden J, 1995. Prediction of the summer rainfall over South Africa. *Journal of Climate* **8**:1511-1518.

Hay LE, Viger R, McCabe GJ, 1998. Precipitation interpolation in mountainous regions using multiple linear regression. In: *Headwater '98* (Kovar K, Tappeiner U, Peters NE, Craig RG, eds). Merano, Italy.

He Y, Barnston A, 1996. Long-lead forecasts of seasonal precipitation in the tropical Pacific islands using CCA. *Journal of climate* **9**:2020-2035.

Henderson RD, Thompson SM, 1999. Extreme rainfalls in the Southern Alps of New Zealand. *Journal of Hydrology (NZ)* **38**:309-330.

Jolliffe IT, 1972. Discarding variables in a principal component analysis. I. Artificial data. *Applied Statistics* **21**:160-173.

Kabanda TA, Jury MR, 1999. Inter-annual variability of short rains over northern Tanzania. *Climate Research* **13**:231-241.

Kidson JW, Barnes BG, 1984. Indices of the Atmospheric Circulation in the Australasian Region. *New Zealand Journal of Science* **27**:355-364.

Kidson JW, Gordon ND, 1986. Interannual variations in New Zealand temperature and precipitation patterns. *N.Z. Journal of Geology and Geophysics* **29**:363-375.

Krishna Kumar K, soman MK, Rupa Kumar K, 1995. Seasonal forecasting of Indian summer monsoon rainfall: a review. *Weather* **50**:449-466.

Lall U, Sangoyomi T, Abarbanel HDI, 1996. Nonlinear dynamics of Great Salt Lake: Nonparametric Short-term Forecasting. *Water Resources Research* **32**:975-985.

Landman WA, 1997. A canonical correlation analysis model to predict South African 1997-98 summer rainfall and temperature. *Experimental long lead forecast bulletin* **12**.

Madden RA, Shea DJ, Katz RW, Kidson JW, 1999. The potential long-range predictability of precipitation over New Zealand. *International Journal of Climatology* **19**.

Mantua NJ, Hare SR, Zhang Y, Wallace JM, Francis RC, 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society* **78**:1069-1079.

McKerchar AI, Pearson CP, 1994. Forecasts of Seasonal River Flows using Southern Oscillation Index. *Journal of Hydrology (NZ)* **32**:16-29.

McKerchar AI, Pearson CP, Moss ME, 1996. Prediction of Summer Inflows to Lakes in the Southern

Alps, NZ, Using the Spring Southern Oscillation Index. *Journal of Hydrology* **184**:175-187.

McKerchar AI, Pearson CP, Fitzharris BB, 1998. Dependency of Summer Lake Inflows and Precipitation on Spring SOI. *Journal of Hydrology* **205**:66-80.

Michaelsen J, 1987. Cross-validation in statistical climate forecast models. *Journal of Climate and Applied Meteorology* **26**:1589-1600.

Mo KC, 2003. Ensemble canonical correlation prediction of surface temperature over the United States. *Journal of Climate* **16**:1665-1683.

Mullan AB, 1998. The science of climate variation and prediction. In: *Workshop on 1997/98 El Nino: lessons and opportunities*. Wellington, New Zealand: Royal Society of New Zealand.

Mullan AB, 1998. Southern Hemisphere Sea-surface Temperatures and their Contemporary and Lag Association with New Zealand Temperature and Precipitation. *International Journal of Climatology* **18**:817-840.

Mullan AB, Renwick JA, 1996. Predictability of New Zealand Climate on Monthly and Seasonal Timescales. In. *Wellington: National Institute of Water and Atmosphere*; 52pp.

Mullan AB, Thompson CS, 1993. Hindcast testing of analogue forecast model. In. Wellington: *National Institute of Water and Atmospheric Research*.

Mullan AB, Thompson CS, 1995. Analogue forecasting of low inflows to Lake Pukaki. In. Wellington: *National Institute of Water and Atmospheric Research.*

Mullan AB, Thompson CS, 2006. Analogue forecasting of New Zealand climate anomalies. *International Journal of Climatology* **26**:485-504.

Mutai CC, Ward NM, 2000. East African Rainfall and the Tropical Circulation/Convection on Intraseasonal to Interannual Timescales. *Journal of Climate* **13**:3915-3939.

Naujokat B, 1986. An update of the observed quasi-biennial oscillation of the stratospheric winds over the tropics. *Journal of Atmospheric Science* **43**:1873-1877.

Palmer TN, Anderson DLT, 1994. The prospects for seasonal forecasting - A review paper. *Quarterly Journal of the Royal Meteorological Society* **120**:755 - 793.

Peters E, Painter D, 1997. Reservoir Inflow Forecasting. In. *Christchurch, New Zealand*; **9**.

Renwick JA, 2002. Southern Hemisphere Circulation and relations with sea ice and sea surface temperature. *Journal of Climate* **15**:3058-3068.

Renwick JA, 2003. Scoring National Climate Centre Seasonal Outlooks. In. The Climate Update, Number 51, 15 September 2003, Niwa, Wellington, New Zealand. See: http://www.niwa.co.nz/__data/assets/pdf_file/0006/32847/cu-2003-09.pdf

Rhoades DA, Barr H, Woods RA, 1990. Improved medium term anticipation of flows into New Zealand Hydro electric reservoirs. In: *DSIR*.

Ruiz JE, Sharma A, Cordery I, 2003. Upper Ocean Temperatures as Predictors of Australian Rainfall. In: MODSIM 2003 (Post DA, ed). Townsville, Australia; 106-111.

Salinger JM, Mullan AB, 1999. New Zealand climate: temperature and precipitation variations and their links with atmospheric circulation. *International Journal of Climatology* **19**:1049-1071.

Shabbar A, Barnston A, 1996. Skill of seasonal climate forecasts in Canada using canonical correlation analysis. *Monthly weather review* **124**:2370-2385.

Sharma A, 2000a. Seasonal to Interannual Rainfall Probalistic Forecasts for Improved Water Supply Management: Part 1 - A Strategy for System Predictor Identification. *Journal of Hydrology* **239**:232-239.

Sharma A, 2000b. Seasonal to Interannual Rainfall Probalistic Forecasts for Improved Water Supply Management: Part 3 - A Non-parametric Probalistic Forecast Model. *Journal of Hydrology* **239**:249-258.

Sharma A, Luk KC, Cordery I, Lall U, 2000. Seasonal to Interannual Rainfall Probalistic Forecasts for Improved Water Supply Management: Part 2 - Predictor Identification of Quarterly Rainfall Using Ocean-Atmosphere Information. *Journal of Hydrology* **239**:240-248.

Sharma A, Lall U, 2003. Using model averaged probabilistic forecasts for water resources management. In: *Modsim 2003, Conference of the Modelling and Simulation Society of Australia and New Zealand* (Post DA, ed). Townsville, Australia; 118-123.

Sivillo JK, Ahlquist JE, Toth Z, 1997. An ensemble forecasting primer. *Weather Forecasting* **12**:809-818.

Smith TM, 2000. Tropical Pacific Sea Level Variations (1948-98). *Journal of Climate* **13**:2757-2769.

Sturman AP, Tapper NJ, 1996. *The Weather and Climate of Australia and New Zealand*. Melbourne: Oxford University Press Australia.

Stephenson DB, Kumar KR, Doblas-Reyes FJ, Royer J-F, Chauvin F, Pezzuli S, 1999. Extreme daily rainfall events and their impact on ensemble forecasts of the Indian Monsoon. *Monthly Weather Review* **127**:1954-1966.

Tomlinson AI, 1981. Some variations of rainfall and river flow in New Zealand. *New Zealand Journal of Science* **24**:103-110.

Tootle GA, Piechota TC, 2004. Suwannee river long range streamflow forecasts based on seasonal climate predictors. *Journal of the American water resources association* **40**:523-532.

Trenberth KE, 1976. Fluctuations and Trends in Indices of the Southern Hemispheric Circulation. *Quarterly Journal of the Royal Meteorological Society* **102**:65-75.

Troup AJ, 1965. The Southern Oscillation. *Quarterly Journal of the Royal Meteorological Society* **91**:490-506.

Vuille M, Bradley RS, Keimig F, 2000. Climate variability in the Andes of Ecuador and its relation to tropical Pacific and Atlantic sea surface temperature anomalies. *Journal of Climate* **13**:2520-2535.

Wedgbrow CS, Wilby RL, Fox HR, O'Hare G, 2002. Prospects for seasonal forecasting of summer drought and low river flow anomalies in England and Wales. *International Journal of Climatology* **22**:217-236.

Widmann M, Bretherton CS, Salathe EP, 2003. Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. *Journal of Climate* **16**:799-816.

Wilks DS, 1995. *Statistical Methods in the Atmospheric Sciences*. San Diego: Academic Press.

Wratt DS, Ridley RN, Sinclair MR, Larsen H, Thompson SM, Henderson RD, Austin GL, Bradley SG, Auer A, Sturman AP, Owens IF, Fitzharris BB, 1996. The New Zealand Southern Alps Experiment. *Bulletin of the American Meteorological Society* **77**:683-692.

Yuval, 2000. Neural Network Training for Prediction of Climatological Time Series, Regularized by Minimization of the Generalized Cross-Validation Function. *Monthly Weather Review* **128**:1456-1473.

Zebiak SE, 2003. Research Potential for Improvements in Climate Prediction. *Bulletin of the American Meteorological Society* **84**:1692-1696.

Zhang H, Casey T, 2000. Verification of categorical probability forecasts. *Weather and Forecasting* **15**:80-89.

Zheng X, Nakamura H, Renwick JA, 2000. Potential Predictability of Seasonal Means based on Monthly Time Series of Meteorological Variables. *Journal of Climate* **13**:2591-2604.

Zheng X, Renwick JA, 2003. A Regression-based scheme for seasonal forecasting of New Zealand temperature. *Journal of Climate* **16**:1843-1853.