

## Flood forecasting: the hard choice

M. Mohssen

*Department of Environmental Management,  
Lincoln University, New Zealand*

### Abstract

A proper forecast of a flood event with a good lead time can result in a flood warning and the issue of effective measures, which can play a major role in mitigating the severe damage that could occur due to this natural disaster. The literature is rich in research for flood forecasting, but due to the complexity and high variability of the driving factors for floods, many of these models fail to accurately forecast floods.

In this research, two models have been developed to forecast the levels for Lake Wakatipu in New Zealand during flood events. The first model is based on a simple hourly mass balance for the lake. However, the choice of the lagged lake level and the cumulative rain in the model formulation were based on a correlation analysis of the available data for the rainfall and lake levels. The second model is a simple regression model that relates the rise of the lake level to the total observed rain. The first model forecasts lake levels 10 hours in advance, while the second model gives 11 hour forecasts. The two models performed well, but the first model is preferred as it is likely to be less biased to the events it was derived from compared to the second one.

On the other hand, non linear optimization has been applied to estimate the parameters of a simple transfer function model, similar to the unit hydrograph concept, for forecasting the Kakanui River flows in Otago, New Zealand. An adjusted form of Philip's equation for infiltration was used to estimate the abstraction of the rainfall event and obtain the effective rainfall that will contribute to the river flow. Weighing factors were assigned to each of the rainfall sites to obtain the best fit between observed and forecasted flows.

*Keywords: flood forecast, flood modelling, rainfall-runoff, lake level, non-linear optimisation.*



## 1 Introduction

Floods are one of the most destructive natural disasters, which cause enormous damage and loss of life every year. Floods can have catastrophic impact on our life, and can cause widespread damage over affected regions. Flood forecasting is an essential tool for flood warning. A proper flood warning could mitigate the impact of a flood event by giving people/authorities enough time to evacuate, take stock or precious items away, or prepare a temporary flood protection scheme.

The Environment Agency of UK has called for a complete review of its flood forecasting and warning system in response to the devastating impact of the Easter 1998 floods of Wales and England, while the strategic plan for the US National Weather Service urges for major investment to develop new forecast models for flood warning [1, 5].

Available mathematical models in the literature can be categorized into two main approaches. The first approach simulates the associated hydrologic processes and utilizes hydraulic or hydrologic routing to estimate river flows. The second approach incorporates the concept of a transfer function (step or simultaneous) to relate effective rainfall to river flow. These techniques are mainly dependent on the simulation of rainfall losses to obtain the effective rainfall, and consequently use hydraulic/hydrologic routing or a response function to estimate the river flow [2–4, 6, 7].

One of the major obstacles for distributed flood modelling for flood forecast is the high variability of spatial and temporal distribution of the rainfall event over the catchment, especially for large catchments [8]. Research is advancing towards the use and utilization of radar data to estimate the spatial distribution of a rainfall event over a catchment [12].

In this paper, lake models are based on a simple hourly hydrologic budget equation, and a simple regression model. The procedure to derive these models was applied to Lake Wakatipu, Otago, New Zealand. Correlation analysis of available rainfall and lake levels data were intensively used to identify lagged parameters for these models. The forecast model for river flows has focused on establishing a proper transfer function between hourly rainfall and flow data. Due to the non-linearity of the objective function to minimize, in order to estimate the model's parameters, a non derivative optimization technique had to be applied. The Powell algorithm for non-linear optimization has been utilized to estimate the model's parameters which describe the infiltration losses, the weighted average between rainfall sites, and the parameters of the transfer function. The Powell algorithm [10, 11], which is an expanded variation of the univariate gradient search, has been widely applied to water resources problems. The Kakanui catchment in North Otago has been selected for the development and application of this model.

## 2 Forecast of lake levels

Rise of lake levels due to flood events is mainly dependent on the cumulative rainfall for the rain storm as the outflow of the lake is dependent on the lake



level regardless of how much inflows to the lake are. Two models have been developed to forecast flood levels for Lake Wakatipu in Otago, New Zealand. Queenstown, a famous town in New Zealand, lies on Lake Wakatipu and is popular for its tourism attractions/activities. The lake area is about 293 km<sup>2</sup>, while the whole catchment area is about 3100 km<sup>2</sup>, see Fig. 1. In November 1999, Queenstown suffered a major flood and the business area beside the lake was completely flooded with devastating losses to the community and the region. Moreover, properties in the flood prone area are facing problems to get insurance after this flood event. The Kawarau River is the main outlet for the lake, and contributes its water to the Clutha River, which is the largest river in New Zealand.

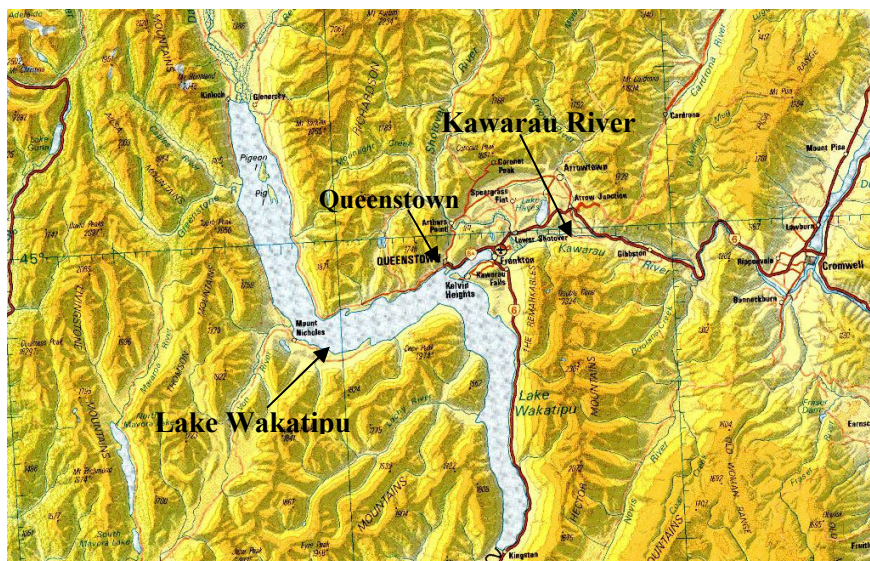


Figure 1: Lake Wakatipu, New Zealand.

## 2.1 Lake model-1 formulation

An hourly hydrologic balance for a lake can be expressed as follows:

$$L_{t+t_f} = L_{t-t_1} - \sum_{j=t-t_1}^{j=t+t_f-1} Q_j + f(R_t) \quad (1)$$

where  $L_t$  is the lake level at present time  $t$ ,  $t_f$  is the forecast time in hours after  $t$ ,  $t_1$  is a lag time before  $t$ ,  $Q_j$  is the lake outflow at time  $j$ ,  $R_t$  is the cumulative rainfall over a period of hours and  $f(R_t)$  is a function of this cumulative rainfall. Of course  $f(R_t)$  will have units of level (mm in our case).

Equation (1) can be re-written as:

$$L_{t+t_f} - L_{t-t_1} + \sum_{j=t-t_1}^{j=t+t_f-1} Q_j = f(R_t) \quad (2)$$

The challenge here is to find the best estimates for  $t_f$ ,  $t_1$  and for the function  $f(R_t)$ . Lake outflows are estimated from the rating of the lake which was

obtained, along with the rainfall and lake levels for the lake for several events, including the flood event of 1999, from the Otago Regional Council, Dunedin, New Zealand. Figure 2 shows the correlation coefficient between cumulative rainfall since the start of the rainfall event and the incremental lake level since the start of the event ( $L_t - L_0$ ), where  $L_t$  is the lake level at time  $t$ , and  $L_0$  is the lake level at the start of the rainfall event ( $t=0$ ). In this case, the initial lake level  $L_0$  at the start of the event is usually used to get the rise of the lake level.

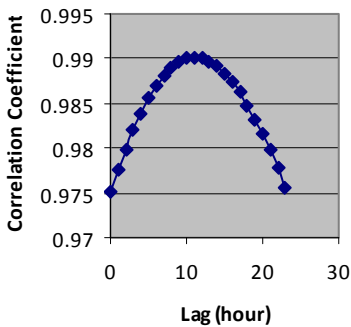


Figure 2: Lagged correlations between cumulative rain and Lake Wakatipu rise.

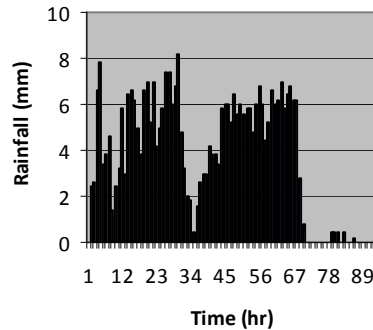


Figure 3: Rainfall event of November 1999 over Lake Wakatipu catchment.

However, as shown in Figure 3, the rainfall event of November 1999 is actually composed of two events, which occurred subsequent to each other. Analyses of the lagged correlation structure for each event separately resulted in the forecast with 10 hours lead time to have the highest correlation, with an average correlation coefficient of 0.93. The maximum correlation between cumulative rainfall and several alternatives for  $(L_{t+10} - L_{t-11})$  is at lag 11, which suggests the use of  $t_1$  to be one hour in Eq (1). Due to the impact of the first event on lake levels during the second event, as water from the first event will still be draining from the catchment during the second event, the results of the first event were given more weight than the second one. Figure 4 shows the relationship between the incremental lake levels ( $L_t - L_{t-11}$ ) lagged 10 hours and the cumulative rainfall. The figure clearly shows that a linear relationship is quite acceptable for  $f(R_t)$ . Thus, Eq. (1) can finally be written as:

$$L_{t+10} = L_{t-1} - \sum_{j=t-1}^{j=t+9} Q_j + 3.16 * (R_t) \quad (3)$$

which can be used to forecast lake levels at time  $(t+10 \text{ hrs})$  based on lake levels at time  $(t-1)$  and cumulative rainfalls at time  $t$  (have in mind that  $t$  is present time). To the contrary, Figure 5 shows the relationship in case the bad choice of using  $L_t - L_{t-1}$  instead of  $L_t - L_{t-11}$  ( $t_f + t_1 = 1$  instead of 11 hours, or incremental 1-hour lake levels, as a normal autoregressive-1 model would do). The results show the importance of carrying this analysis before identifying the best candidates for the model variables.

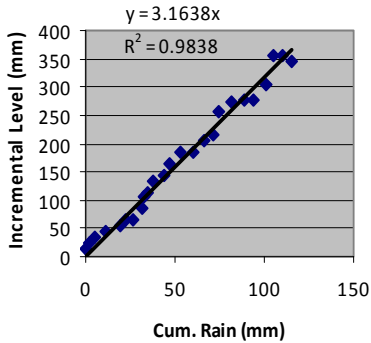


Figure 4: Relation between cumulative rain and lagged 10 hours ( $L_t - L_{t-11}$ ).

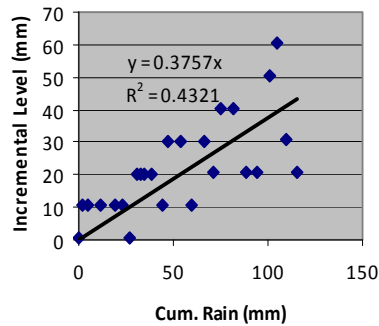


Figure 5: Relation between cumulative rain and lagged 10 hours ( $L_t - L_{t-1}$ ).

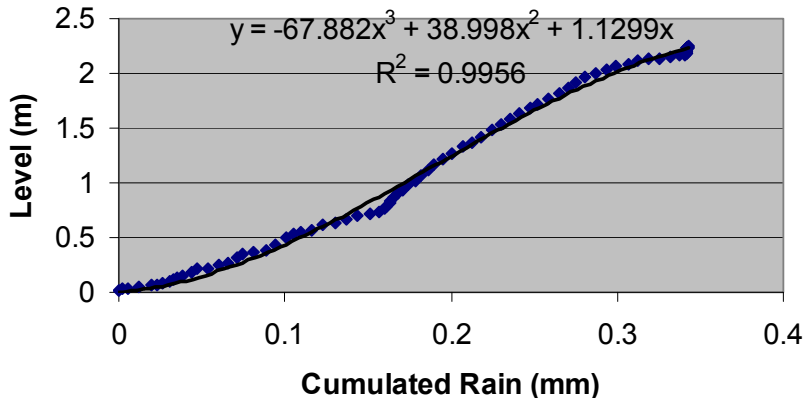


Figure 6: Relation between cumulative rain and lake rise.

## 2.2 Lake model-2 formulation

By using a forecast time of 11 hours, figure (6) shows the relation between cumulative rain from the start of the rain process, and the lake rise, compared to the start of the event. Model 2 is written as:

$$\Delta L_{t+11} = -67.9(R_t)^3 + 39(R_t)^2 + 1.13R_t \quad (4)$$

where  $\Delta L_{t+11}$  is the rise of lake level since the start of the rainfall event and up to 11 hours after present time  $t$ , and  $R_t$  is the cumulative rainfall since the start of the rainfall event up to the present time  $t$ . The equation is really simple, and the trick was to get the suitable lag time for forecast. However, it is expected that a non-linear equation like this will be more biased towards the event from which it was estimated (November 1999 flood event) compared to the first model.

2.3 Model testing

Both model-1 and model-2 were verified by applying them to events which were not included in their calibration process to estimate their parameters. Figures 7 and 8 show that both models performed well. However, model-1, in general, overestimated lake levels more than model-2. It is quite interesting to note that Model-2 is performing well despite its simplicity.

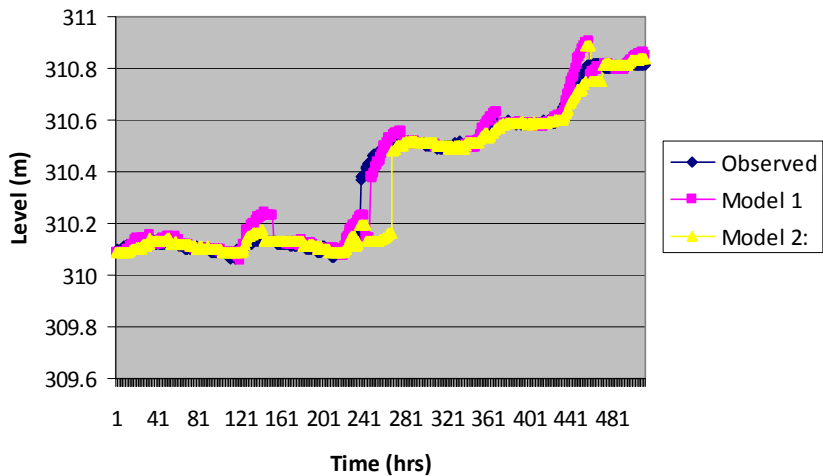


Figure 7: Model testing for rainfall events during November 1997.

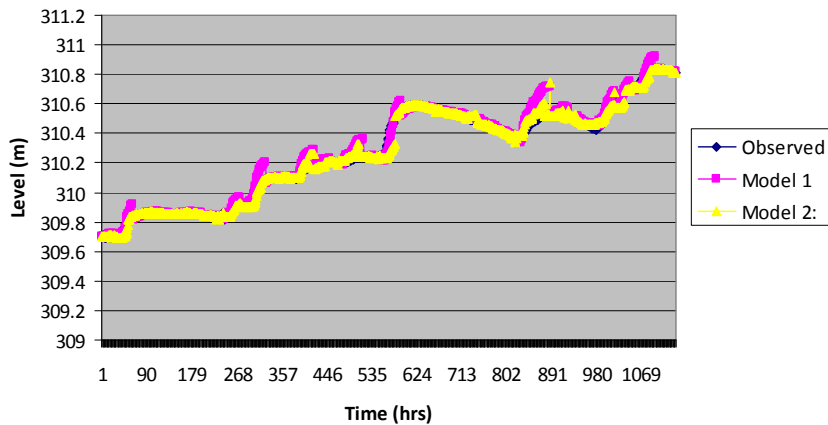


Figure 8: Lake models testing for the rainfall events during September–October 1998.





### 3 Forecast of river flows

A forecast model for the flows of the Kakanui River in North Otago has been developed by using non-linear optimization to estimate the parameters of a transfer function to transfer rainfall into river flows, similar to the concept of the Unit Hydrograph [4, 7].

The Kakanui River flows into the Pacific Ocean 10km south of Oamaru, New Zealand. With a catchment of 894km<sup>2</sup>, the catchment consists of about 35% river valley and 40% of rolling hills of less than 600m elevation. The remaining 25% of the catchment is mountainous, reaching heights of some 1640m [9]. Figure 9 shows the Kakanui catchment with rainfall and flow sites, and with the location for the samples for the infiltration test carried out for this research.

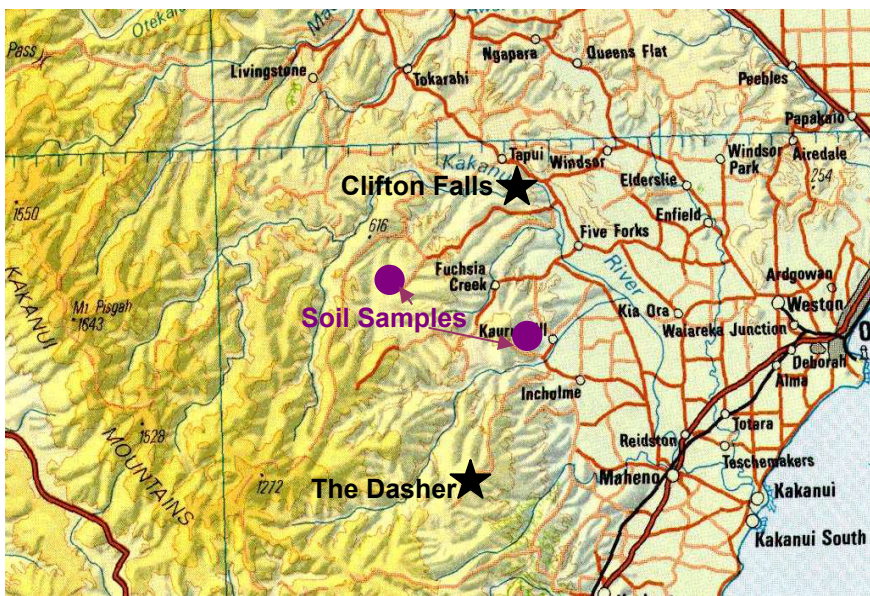


Figure 9: Location of soil sampling sites, rain-gauges, and flow monitoring sites for the Kakanui catchment.

#### 3.1 River forecast model

The concept behind this model is that each catchment should have its own “unique” hydrologic characteristics which will impact on its response to a rainfall event. This transfer function is assumed to be invariant with time, and the case which is presented here considers a linear transfer function. The parameters of the transfer function have been estimated by using Powell algorithm to minimize the objective function which describes the relation between observed and modelled flows. The objective function  $F_x$  is defined as the sum of squared errors as follows:

$$F_x = \sum_{j=1}^m \sum_{i=1}^{n_j} (Q_{i,j} - \hat{Q}_{i,j})^2 \quad (5)$$

where  $m$  is the number of events included in the calibration process,  $n_j$  is the number of intervals of event  $j$ ,  $Q_{i,j}$  is the observed flow at interval  $i$  of event  $j$ , and  $\hat{Q}_{i,j}$  is the forecasted flow for event  $j$  at time interval  $i$ . The forecasted flow  $\hat{Q}_{i,j}$  is calculated from the transfer function  $\underline{X}$  as follows:

$$\hat{Q}_{i,j} = \sum_{k=1}^{i \leq L} R_k X_{i-k+1} \quad (6)$$

where  $L$  is the rainfall intervals of event  $j$ ,  $R_k$  is the rainfall depth at time interval  $k$ , and  $X_i$  is the  $i^{\text{th}}$  parameter of the transfer function. It should be noted that the objection function  $F_x$  is highly nonlinear in so many parameters, and thus a numerical non-linear optimization technique had to be carried out.

### 3.2 Hydrologic abstractions

Philip's equation to simulate infiltration through a soil [7] was used to estimate the hydrologic abstractions from a rainfall event in the model. The infiltration rate is defined as:

$$f_t = a + \frac{b}{2\sqrt{t}} \text{ mm/} \quad (7)$$

while the cumulative infiltration is:

$$F_t = \int_0^t f_t = at + b\sqrt{t} \quad (8)$$

It is the cumulative infiltration after a time interval "t" which we are interested in for this model. However, this model assumes ponding conditions apply since the start of the event, which is not the case. To account for this, cumulative infiltration every time step has been calculated and compared to the cumulative rainfall up to this time step, then the estimated infiltration depth is taken as the smaller value of the cumulative observed rain or the cumulative infiltration depth from equation (8). The same procedure was carried out to define the ponding time, after which equation (8) is applicable. The parameters for the infiltration model were estimated also through the optimization process by Powell algorithm. It should be noted that equation (8) was used to estimate the abstraction losses, and in turn the effective rainfall of the event.

Moreover, an infiltration test, using the double ring infiltrometer, was carried out at two sites in the catchment to investigate the infiltration capacity of the Kakanui catchment. The locations of these two sites are at the upper of Clifton Falls, and at the Dasher, as illustrated in Fig. 9. The average "stable" observed infiltration rates at upper Clifton Falls and the Dasher were 40 mm/hr and 150 mm/hr, respectively. This is a very high value, and will exceed the rate of any rainfall event. However, these infiltration tests were carried out for a flat area,



while the majority of the catchment is hilly. Moreover, a rainfall runoff experiment was carried out in the field (also for almost a flat area) in the upper Clifton Falls catchment, and confirmed that the “steady” infiltration capacity of the “flat” soil is higher than 30mm/hr.

In addition to at site field experiments, two soil samples of the Kakanui catchment at upper Clifton Falls and the Dasher have been obtained for further experiment in the hydraulic lab of Lincoln University. Each sample is 700x700x200 mm, and a sprinkler was used to simulate rainfall events over these samples. Surface runoff and through flows were collected and measured for several simulated rainfall events. Again, the results confirmed that the infiltration capacity of the “flat” soil is high. The above analysis leads to the fact that the infiltration model is used to estimate the temporal hydrologic abstractions from a rainfall event, rather than the actual infiltration to the soil, as a component of this infiltrated water could contribute to the river flow as a through flow. The optimized values for parameters  $a$  and  $b$  are 0.5 mm/hr and  $1.5\text{mm/hr}^{0.5}$ , respectively.

### 3.3 Model calibration

Nine observed flood events at Clifton Falls, along with their corresponding rainfall events at Clifton Falls and the Dasher, were used for model calibration and the estimation of the model parameters. Table 1 presents a summary of these flood events.

The selected floods for model calibration cover a wide variety of flood events with regard to the duration, total rainfall, peak flow and the total runoff volume, as shown in table 1. Moreover, it is obvious from the table that usually the Dasher receives more rain than Clifton Falls. Unfortunately, the ratio between the rainfalls of the two sites is not the same for all rainfall events, and in some

Table 1: Flood events for model calibration.

Event	Date	Duration (hrs)	Total Rain (mm)		Peak Flow( $\text{m}^3/\text{s}$ )	Total Runoff ( $10^6$ $\text{m}^3$ )
			Dasher	Clifton falls		Clifton Falls
1	13/06/1995	85	128.0	37.5	77.1	9.0
2	20/11/1996	43	31.0	37.0	82.1	3.1
3	4/02/1997	56	73.8	22.5	44.3	3.4
4	31/08/2000	129	217.0	67.0	148.0	19.0
5	18/07/2001	50	40.0	15.0	12.0	0.8
6	20/07/2001	72	149.0	77.0	96.0	7.9
7	12/01/2002	63	265.0	62.5	256.0	16.7
8	10/01/2002	68	182.5	58.5	256.6	11.9
9	12/01/2002	63	265.0	62.5	256.0	16.7



cases, such as event 2, Clifton Falls received more rain than the Dasher. Of course this is dependent on the rainfall event, its direction, and its spatial distribution as it hits the catchment. Such discrepancy is expected to confuse the model. Figure 10 presents a good match between the forecasted and observed flows using the developed model for events 8 and 9 “January 2002”. This could be attributed to the fact that Clifton Fall’s rain was less than 1/3 the rainfall at the Dasher for events 8 and 9, which is the case for most of the events. The good performance of the model for the flood events of January 2002 indicates the ability of the model of reliably forecasting the flows of flood events which preserve the pattern of the significantly lower rainfall at Clifton Falls than the Dasher.

3.4 Model validation

The fitted forecast model for the Kakanui River was applied to several flood events which were not included in its derivation. Figure 11 presents the application of the transfer function model to the observed rainfall events at Clifton Falls and the Dasher, then comparing the forecasted flows versus the observed ones. The model performed well in forecasting the flood flows of the event 11 February 1997, and did well in forecasting the peak flow for the most recent event on 30 July 2007. It is noted that the Dasher rain gauge stopped sending signals for its rainfall after about 20 hours from the start of the event, and estimated values for the Dasher rainfall were estimated based on Clifton Falls. This could have an impact on the hydrograph for the forecasted flows.

4 Conclusions

Two simple models for the forecast of lake levels have been derived and applied for Lake Wakatipu, New Zealand. Both models were capable of properly forecasting the lake level, and, in general, performed well. Intensive analysis of the correlation structure between incremental lake levels with different lags and

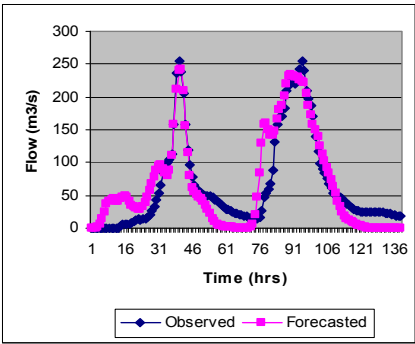


Figure 10: Flood event January 2002 (events 8 and 9).

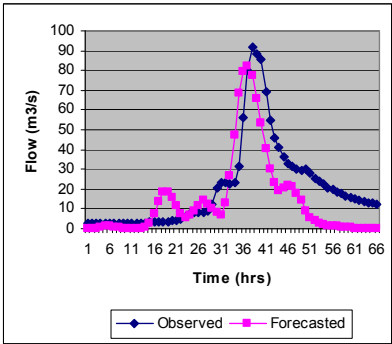


Figure 11: Flood event 11 February 1997.



cumulative rainfalls was carried out to identify key parameters for the model. The approach has shown the importance of properly carrying out the correlation analyses, otherwise a major error in regard to the use of a wrong lag between the model's variables would result in a useless model.

The derived model for river flow forecast was capable of forecasting the Kakanui River flows satisfactory. However, the high variability of rainfall events, and the difficulty of representing the spatial variability of a rainfall event makes it difficult to exactly estimate the perfect transfer function. Moreover, the derived transfer function was linear, while adding non linearity could improve the performance of the derived model. In addition, it is recommended to include the direction or the type of the rainfall event as this could incorporate the spatial variability in the modelling process

## References

- [1] Bye, P. and M. Horner, Easter 1998 Floods Report by the Independent Review Team to the Board of the Environmental Agency, vol.1, Environmental Agency, Bristol, 1998.
- [2] Chang-Fi John; Lin-Yong Huang; Chang-Li Chiu; Chang-Fi John, Enforced self-organizing map neural networks for river flood forecasting. *Journal of Hydrological Processes*, 21(6), pp.741-749, 2007
- [3] Chen-Shen Hsien; Lin-Yong Huang; Chang-Li Chiu; Chang-Fi John, The strategy of building a flood forecast model by neuro-fuzzy network. *Journal of Hydrological Processes*, 20(7), pp.1525-1540, 2006.
- [4] Chow, V.T., Maidment, D. R., and L.W. Mays, *Applied Hydrology*. McGraw Hill, Inc, NY, USA, pp. 201-342, 1988.
- [5] Demeritt D., et al., Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting, *Environmental Hazards*, vol 7, pp 115-127, 2007.
- [6] Jowitt, P.W., A Conceptual System Model of Rainfall-Runoff on the Haast River. *Journal of Hydrology, New Zealand*, 38(1), pp. 121-144, 1999.
- [7] Maidment, D. R., *Handbook of Hydrology*. McGraw-Hill, Inc, Chapter 26 "Hydrologic Forecasting", 1993.
- [8] Moore, R. et al., Forecasting for flood warning, *C.R. Geoscience* vol 337, pp 203-217, 2005.
- [9] Otago Regional Council, North and coastal Otago river catchments monitoring report, *Otago Regional Council Miscellaneous Publication ISBN No.1-877265-05-5*, 2, 2000.
- [10] Press, W.H.; B. Flannery; S.A. Teukolsky; W.T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, USA, pp. 294-301, 1989.
- [11] Reklaitis, G.V.; A. Ravindran and K. Ragsdell, *Engineering Optimization*, John Wiley & Sons, Inc., USA, pp87-98, 1983.
- [12] Smith, J. A., et al., Radar rainfall estimation for flash flood forecasting in small urban watersheds, *Advances in Water Resources*, vol 30, pp 2087-2097, 2007.

