FLOOD FORECASTING OF LAKE LEVELS:
A NEW CONCEPT

M. MOHSSEN[1] & M. GOLDSMITH[2]
[1]Lincoln University, New Zealand.
[2]Otago Regional Council, New Zealand.

ABSTRACT

A new concept for flood modelling of lake levels has been introduced, analysed, and tested for Lake Wakatipu in the South Island of New Zealand. Lake level response to significant rainfall events is substantially different from rivers. Lakes usually have enormous storage to absorb the peak of the flood event, and their outflows are dependent on their level rather than the inflows. The new concept is based on carrying out a lagged correlation analysis between hourly data for the cumulative rainfall and the total rise of the lake level in order to determine the lagged time series of total lake rise which will provide the best projection for its forecast onto the cumulative rainfall. The focus has been to carry out this analysis on hourly data during significant rainfall events which resulted in significant lake rise. In return, the time frame for such analysis is usually several hours or few days, as is the case for Lake Wakatipu. Thus, some variables which are usually considered for lake analysis, such as evapotranspiration, can be negligible compared with the huge amount of rain causing the flood event. Based on the best lagged correlation, three models have been derived to forecast flood levels of Lake Wakatipu. Correlation analysis of lagged correlations showed that a lag of 11 hours results in the highest correlation between cumulative rainfall and total rise of Lake Wakatipu level, while 14 hours produced the highest correlation between cumulative rainfall and total lake inflows. All three models utilised these optimum lagged hours for the best projection of total lake rise/inflow. The first model is a lag-11 non-linear regression model, while the second model is a lag-11 linear regression model, and the third one is a lag-14 comprehensive hourly mass balance model. The results of model testing showed that the simple linear regression model produced the best forecasts, while the more sophisticated complete mass balance model, in general, was not as good, and the non-linear regression model (while having the highest determination coefficient) was the least performing.

Keywords: flood forecast, flood modelling, lagged correlations, lake level, projection theorem, rainfall-runoff, regression analysis.

1 INTRODUCTION

Floods are one of the most destructive natural disasters causing 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 within flood warning systems. A proper flood warning could mitigate the impact of a flood event by giving people/authorities enough time to evacuate, relocate stock, or move precious items away, or prepare a temporary flood protection scheme.

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 could occur due to this natural disaster.

The Environment Agency of the 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, 2].

There is a wide range of literature and research relating to flood forecasting of lake levels. However, due to the complexity and high variability of the driving factors for floods, many of these models fail to accurately forecast flood events. Available mathematical models in the

literature can be categorised into three main approaches. The first approach simulates the associated hydrologic processes to estimate the runoff from the catchment area due to the rainfall event, and utilises hydraulic or hydrologic modelling/routing or a response function to predict input flow hydrograph to the lake [3–5, 6, 7]. Manual or automatic optimisation techniques [8] are often implemented to estimate the wide range of parameters usually incorporated in the development of these models. After estimating the inflows to the lake, a mass balance is carried out, utilising the usually known rating for the outflows to forecast the rise of the lake levels. The second approach incorporates time series analysis to build a model which can be used to forecast future lake levels based on present and past observed values. Time series models, such as the integrated autoregressive moving average "ARIMA" models, don't incorporate the main driving force in this flooding, which is rain [9]. A Kalman filtering technique can be utilised to include rainfall in the forecast of rising lake levels [10]. The third approach utilises artificial neural network modelling to build a non-linear relationship between input and output [9, 11, 12].

One of the major obstacles for distributed flood modelling for the forecast of lake inflows is the variability of spatial and temporal distribution of rainfall across the catchment, especially for large catchments [13]. Research is advancing towards the use and utilisation of radar data to estimate the spatial distribution of a rainfall event over a catchment [14].

1.1 Lake Wakatipu catchment description

Lake Wakatipu is located immediately to the east of the Southern Alps, within the Queenstown Lakes District (Fig. 1). It is New Zealand's third-largest lake with a surface area of 293 km^2 and a total catchment area of 3,067 km^2. The deep depression in which the lake sits was formed following the retreat of the Wakatipu Glacier approximately 15,000 years ago. The lake itself is relatively long and thin, being over 80 km in length, with a maximum width of 5 km. Average lake level is 310 metres above sea level (masl), and parts of the lake are very deep, its floor being as low as 100 m below sea level.

Following the retreat of the Wakatipu Glacier, a range of landforms were created from streams flowing from the mountain ranges which flank the lake. These features form the underlying geology and shape the surrounding environment of Lake Wakatipu. They consist of alluvial fans, terraces, deltas, and lake shorelines. The main settlements adjacent to the lake are Queenstown, Glenorchy, and Kingston, as shown in Fig. 1. Low-lying parts of these townships are within the range of peak lake levels, and have historically been affected by flooding.

The largest tributaries to Lake Wakatipu are the Dart and Rees Rivers, which lie to the north of Glenorchy. The braided floodplains of these two rivers deliver large quantities of sediment and flow into the lake. Other significant tributaries which flow directly into the lake along its western shoreline include the Greenstone, Caples, Von, and Lochy Rivers. Outflow is via the Kawarau River, which flows out from the lake's Frankton Arm, 8 km east of Queenstown.

1.2 Flooding in Lake Wakatipu catchment

Flood hazards, in the form of river flooding and lake inundation, pose a threat in the Lake Wakatipu catchment due to the resulting impact on people, property, and infrastructure. Inundation from high lake levels is a risk due to the lake remaining at high levels for prolonged periods, usually days to weeks. The level of Lake Wakatipu has been observed to vary through
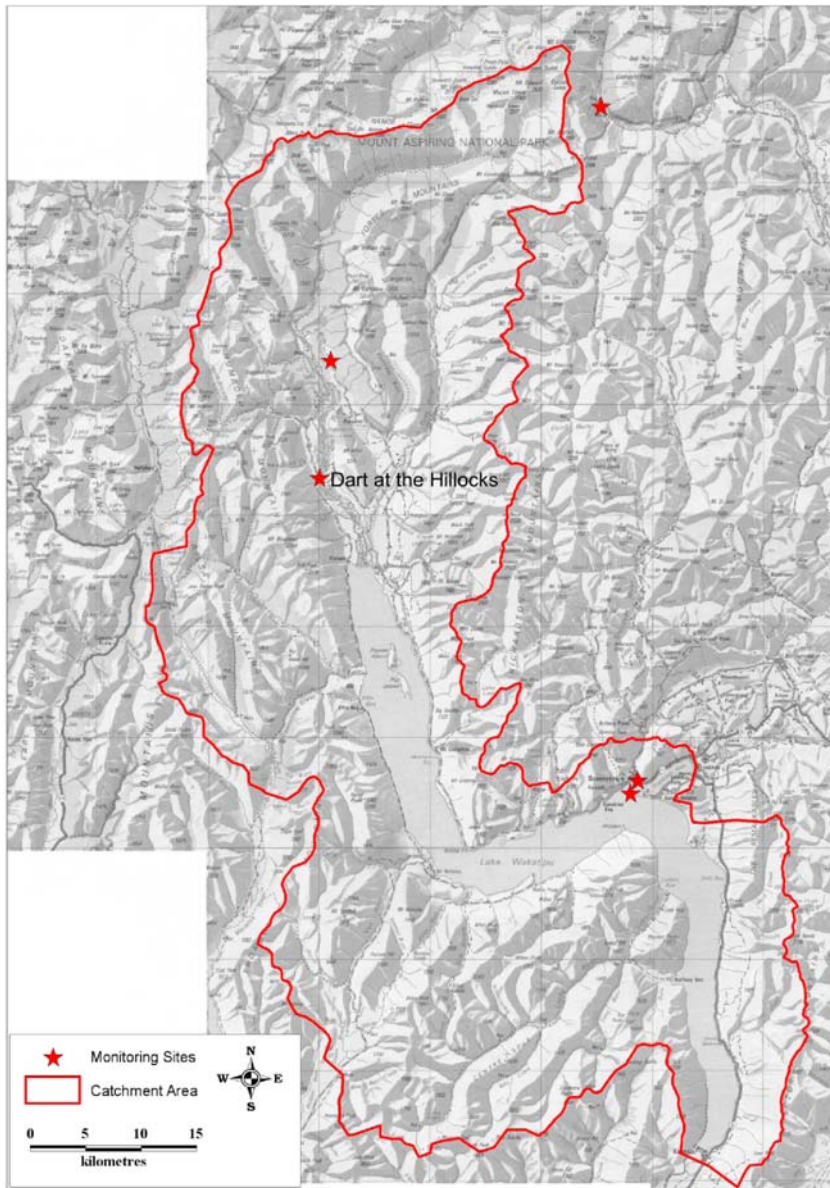
Figure 1: Lake Wakatipu catchment

a range of nearly 3.9 m with a mean level of approximately 310 masl. The flood of November 1999 was the highest lake level on record at 312.77 masl [15] with the second highest level being recorded in September 1878 at 312.60 masl (Fig. 2). During the 1999 flood event, 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 frequency and magnitude of flood events are closely related to the rainfall events from which they are derived. Annual rainfall in the western part of the South Island generally
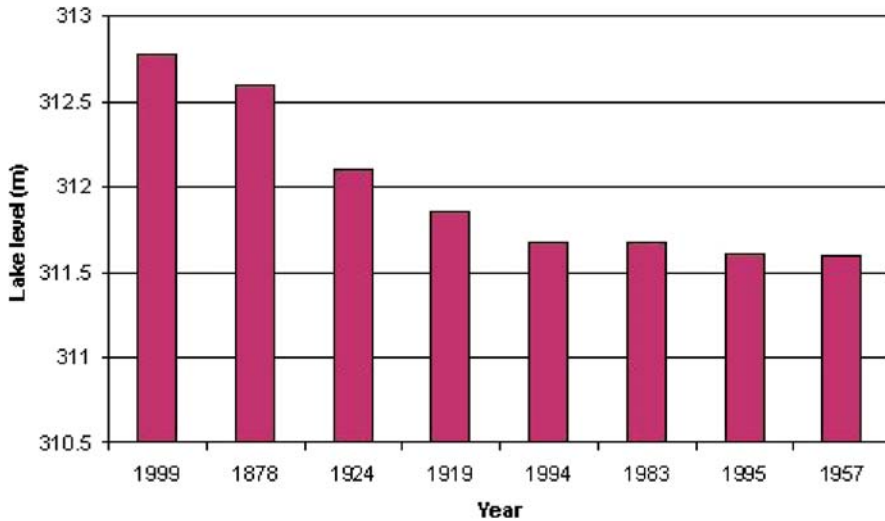
Figure 2: Highest recorded levels of Lake Wakatipu during the period 1878–2007; lower parts of Queenstown are inundated at about 311.5 m.

exhibited an upward trend during the 20th century, with annual rainfall at Queenstown increasing by approximately 160 mm between 1901 and 2003 [16]. The rate of increase appeared to accelerate during the latter part of the 20th century. Since the turn of the century however, annual rainfall totals in north-western Otago have generally declined again, and this trend has also been evident at gauges within the Lake Wakatipu catchment.

There is moderate confidence that average annual rainfall in the north-west of Otago is projected to increase by approximately 12%, with a range of −2 to 34% [17]. Heavy rainfall events are also projected to become more frequent, and may become more intense. Note that these projections might change in the future as our understanding of climate change and its impact improves.

The main cause of high lake levels in Lake Wakatipu is the natural imbalance between the capacity of the lake outlet (Kawarau River) and the magnitude of inflows during heavy rainfall events. Due to the location of the Shotover River confluence near the lake outlet, outflow from the lake can be further impeded by flood and sediment flows [18]. These conditions, in association with a succession of frontal weather systems, can provide circumstances where the lake may stay at high levels for prolonged periods [19].

2 MODELLING OF LAKE WAKATIPU FLOOD LEVELS

2.1 Analysis of the 1999 flood event

This flood event was the worst in recent history that Queenstown had experienced, as shown and mentioned above. The event lasted for about 70 hours, and resulted in a total of 341 mm of rain. The lake level increased by 2.25 m due to this event, and flooded low-lying parts of Queenstown, a major resort area in the South Island of New Zealand. Figure 3 presents a hyetograph for the 1999 rainfall event. The Figure indicates that the rainfall event is actually composed of two consecutive systems of rainfall storms. The cumulative rainfall for this event, as shown in Fig. 4, confirms that this event is actually two events, as the linear trend

flattens at the end of the first event and then picks up again with the second event. In addition, the flood hydrograph of the Dart River which is one of the main tributaries to the lake, as shown in Fig. 5, strengthens the two storm systems theory.

Despite the variability of the hourly rainfall event, the cumulative rainfall shows a good linear trend, which is almost consistent over the two events. Again, cumulative rainfall starts to flatten at the end of the second event as rainfall diminishes.

Figure 5 also indicates that the "bulk" of the inflows to the lake due to this event occurred within 90 hours of the start of the event, about 20 hours after the end of the rainfall event.
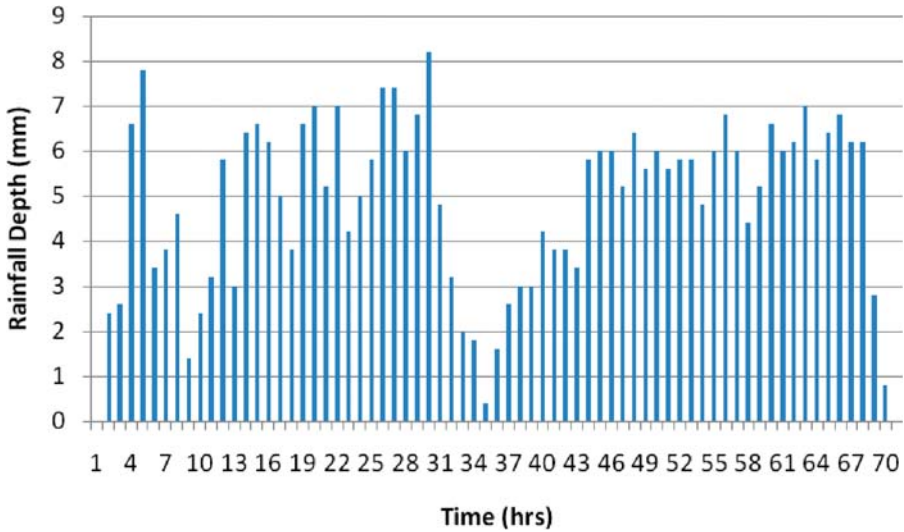

Figure 3: Rainfall hyetograph for the 1999 flood event of Lake Wakatipu.
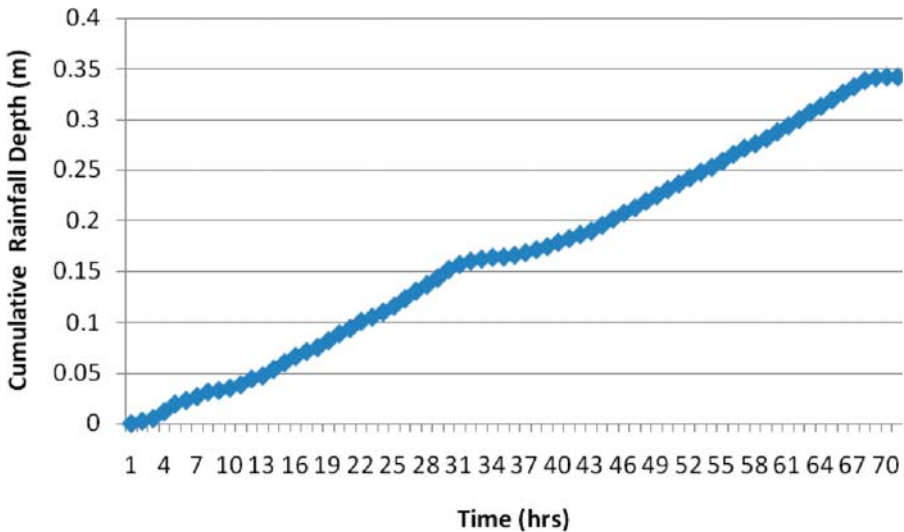

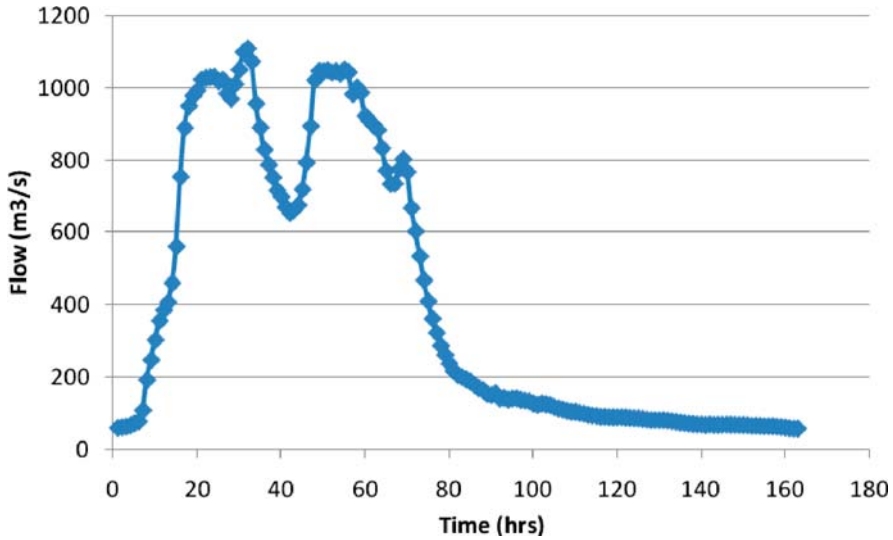Figure 4: Cumulative rainfall depth for the 1999 flood event of Lake Wakatipu.

Figure 5: Flow hydrograph for the Dart River.

Total inflows to the lake, which include surface runoff from the lake's catchment and direct rainfall over the lake, were estimated by using the following simple mass balance equation:

$$Input = \Delta S + Output \qquad (1)$$

Where ΔS is the change in the lake storage, represented by the rise in lake level which is monitored, and the output is the lake's outflow which has been estimated from the rating formula for lake. The rating formula to predict Lake Wakatipu outflow from the lake levels was derived by fitting a non-linear relationship between lake levels and gauged lake outflows as follows:

$$Q_t = 81.11(L_t - 309)^2 + 82.57(L_t - 309) + 18.93 \qquad m^3/s \qquad (2)$$

Where Q_t is the lake outflow at time t in m³/s and L_t is the lake level at that time in m.

The total amount of rain which fell during this event over the whole catchment is not known, as available rainfall sites in the area are scarce, especially in the highest elevations of the catchment near the top of the Southern Alps. Thus, watershed models which estimate runoff from the catchment area have to depend on estimated rainfall values and an intensive calibration process to fit the unknown models' parameters, which would fail to produce a reliable flood forecast.

However, considering a time frame of 90 hours since the start of the event, the estimated total input to the lake due to this event is about 300 mm, by aid of eqn (1) and knowing the areas of the lake and the watershed. This is not far from the observed total rainfall of 341 mm observed in the upper catchment at the Dart at Hillocks site (Fig. 1). Thus, observed rainfall at the Hillocks rainfall site can be utilised to represent rainfall over the whole catchment of Lake Wakatipu. Of course, the pattern and spatial distribution of each rainfall event will have significant impact on this representation. However, most severe rainfall events are usually due to rainfall storm systems which are coming from the west coast of New Zealand and dumping its moisture due to orographic effects.

2.2 Modelling of rainfall – lake rise relationship

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

$$L_{t+t_L} = L_{t-t_1} - \sum_{j=t-t_1}^{j=t+t_L-1} Q_j + f(R_t) \tag{3}$$

Where L_t is the lake level at time t, with $t = 0$ at the time step just before the event starts, t_L is the forecast time L 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 at time t and $f(R_t)$ is a function of this cumulative rainfall. Of course, Q_j and $f(R_t)$ will have units of level (m in this case). Equation (3) can be re-written as:

$$L_{t+t_L} - L_{t-t_1} + \sum_{j=t-t_1}^{j=t+t_L-1} Q_j = f(R_t) \tag{4}$$

The challenge here is to find the best estimates for t_L, t_1, and the function $f(R_t)$. The relationship between cumulative hourly rainfall and incremental hourly increases in lake levels, as shown in Fig. 6, indicates the failure of any attempt to derive a good model capable of producing accurate forecasts, which is fed by these data.

This problem has been significantly resolved when considering total lake rise since the start of the event instead of the incremental lake rise, as shown in Fig. 7. Thus, the best estimate for t_1 of eqn (1) is t, $t - t_1 = 0$, and $= L_0$, where L_0 is the water level of the lake at the time step just before the start of the event. Despite the significant improvement of the variability of the relationship, Fig. 7 suggests a non-linear relationship for the first part of the event (which is the first storm system), and then a linear model for the second storm system. The linear fit shown in the figure is set to start from zero to be more consistent with the physical process (our focus here is during big events, and not the recession period). A lagged correlation analysis has been carried out between total lake rise (lake rise since the start of the event) and cumulative rain, as shown in Fig. 8. The results of this analysis clearly recommend the use of a lag of 11 hours between cumulative rain and total lake rise.

The relationship between cumulative rain and lag-11 total lake rise significantly improved, as shown in Fig. 9.
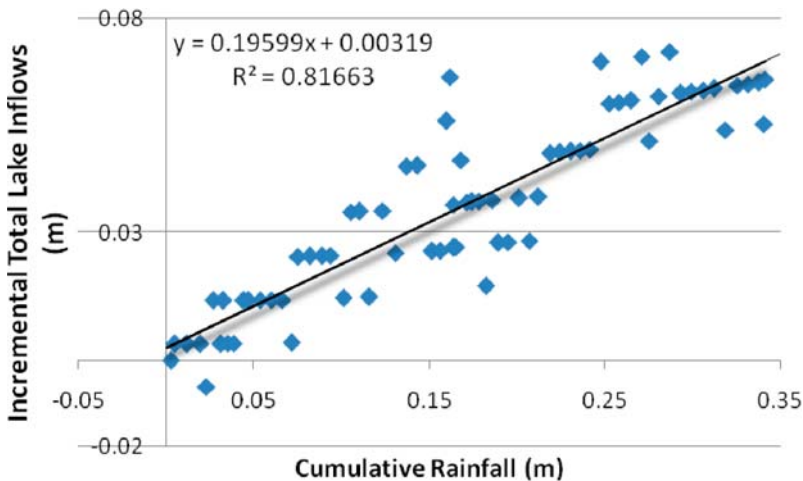

Figure 6: Cumulative rainfall vs incremental Lake Wakatipu inflows for the 1999 flood event.
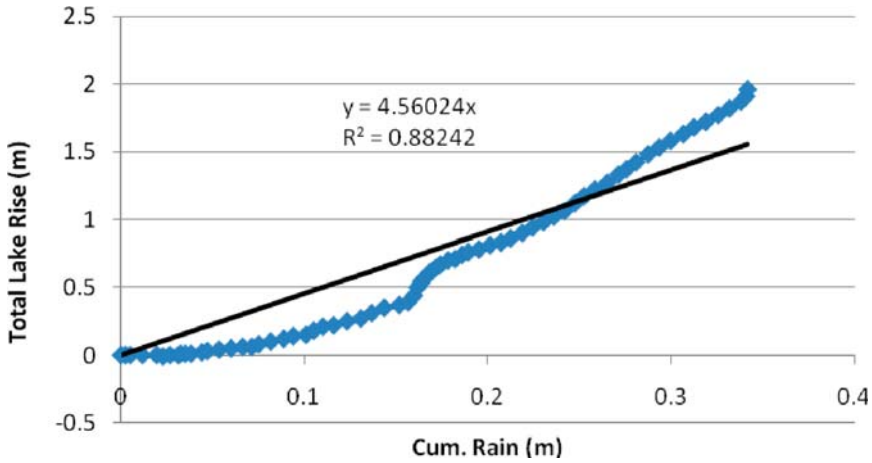
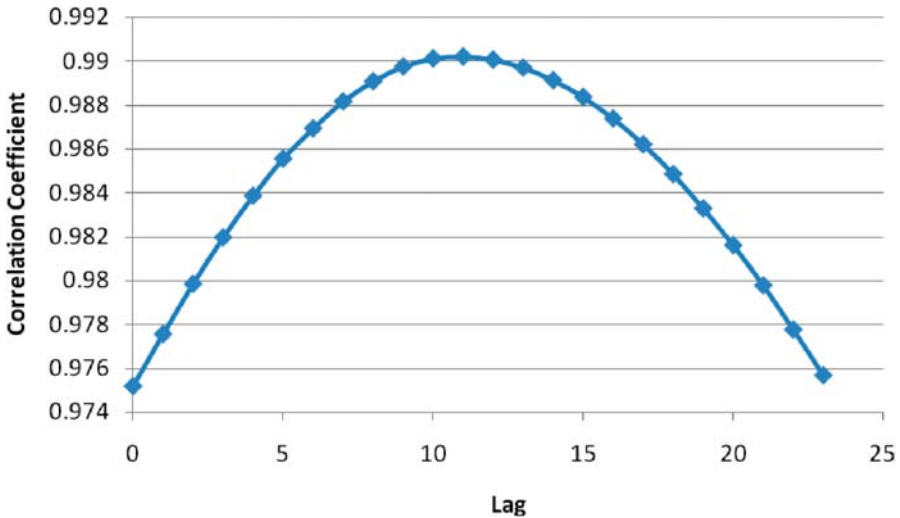Figure 7: Cumulative rainfall vs total lake rise since the start of the 1999 event.

Figure 8: Lagged correlations between cumulative rain and total Lake Wakatipu rise (1999 event).

2.2.1 Non-linear regression model

The fitting of a non-linear regression model for this relationship would result in a determination coefficient of 0.997, which is significantly higher than the linear fit. However, one has to be careful with the application of such non-linear model, as it will definitely work perfect with the calibrated 1999 flood event from which it was derived, but its application to other events could be quite unsatisfactory, as will be shown later. In addition, extrapolation of non-linear models outside the domain they were tested for can be quite risky.

The first model for Lake Wakatipu forecast is a lag-11 non-linear regression model as follows:

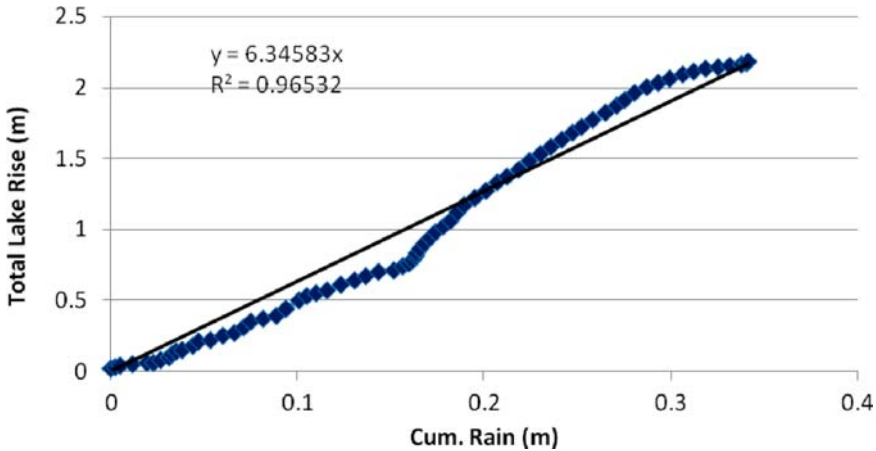$$L_{t+11} = L_0 - 65.7R_t^3 + 38.15R_t^2 + 1.2R_t \tag{5}$$

Figure 9: Lag-11 relation between cumulative rain and total Lake Wakatipu rise (1999 event).

Where t is time at present, R_t is cumulative rain (in meters) from the start of the event until present time t, and L is the lake level in meters. This model is capable of forecasting lake levels 11 hours later based on current cumulative rainfall at the hillock rainfall site. However, one can use forecasted rainfall to have longer forecast.

2.2.2 Linear regression model

The projection theorem can be applied to obtain the best forecast of the total lake rise based on cumulative rain [20]. Assuming that R and ΔL are two variables in Hilbert Space, denote $\widetilde{\Delta L}$ as the projection of ΔL on R, where $\widetilde{\Delta L} - \alpha R$. Thus, $\Delta L - \alpha R$ is orthogonal to R and their inner product in Hilbert space is:

$$< \Delta L - \propto R, R > = 0 \tag{6}$$

starting from equation (6), note that is ΔL with a hat above it.

Solving eqn (6) gives the following formula to estimate:

$$\propto = \frac{E[\Delta L R]}{E[R^2]} \tag{7}$$

Equation 7 would produce the same value obtained from the fitted regression line, as shown in Fig. 9. The projection theorem guarantees that the obtained projection is the best forecast (linear) for lake rise based on cumulative rain. Thus, our second model for the flood forecast of Lake Wakatipu is a lag-11 linear regression model with a determination coefficient of 0.96 as follows:

$$L_{t+11} = L_0 + 6.35 R_t \tag{8}$$

It should be noted that eqn (8), implicitly, incorporates lake outflows, as the total lake rise $(L_{t+11} - L_0)$ is the result of the net gain to the lake (inflows – outflows). The weak point for this approach is that this implicit incorporation of the outflows is based on the outflows during the 1999 event, which is dependent on the lake levels during that event. Lake levels during other events could be different, and in turn, lake outflows would be different. However, the significance of this will be investigated during the model's validation process.

2.2.3 Linear regression model with explicit outflows

So far, outflows of the lake have not been explicitly incorporated into the first and second forecast models, eqns (5) and (8), respectively. Carrying out another lagged correlation analysis between total inflows to the lake and cumulative rain indicates that the use of 14 hours lead lag time would produce the best projection of total lake inflows onto cumulative rain, as shown in Fig. 10. This gives 3 hours more forecast time compared with the first and the second models, and the shape of the lagged correlations is not as bell shaped as that in Fig. 8. Moreover, the peak correlation at lag-14 is slightly higher than that of Fig. 8 at lag-11. Figure 11 presents the linear relationship between the two sides of eqn (4), the total inflows to the lake (left hand side), and the cumulative rainfall (right hand side), starting from $t = 0$.

Thus, the third forecast model for Lake Wakatipu flood levels is lag-14 mass balance model derived from eqn (3) as:

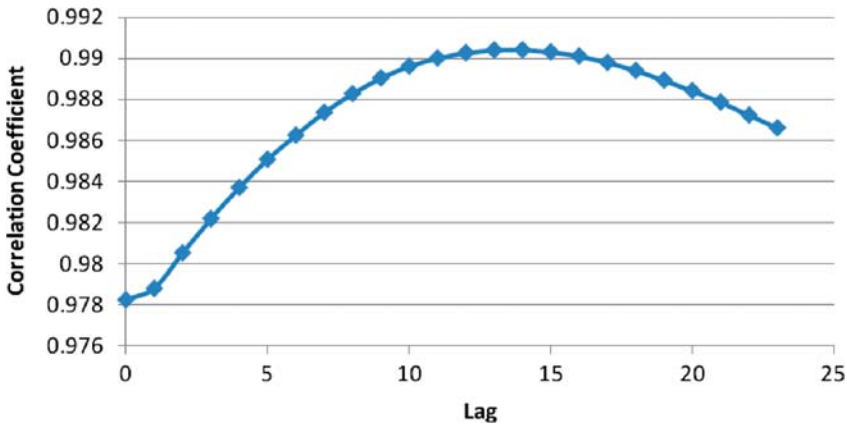$$L_{t+14} = L_0 - \sum_{j=0}^{j=t+13} Q_j + 8.76*(R_t) \tag{9}$$

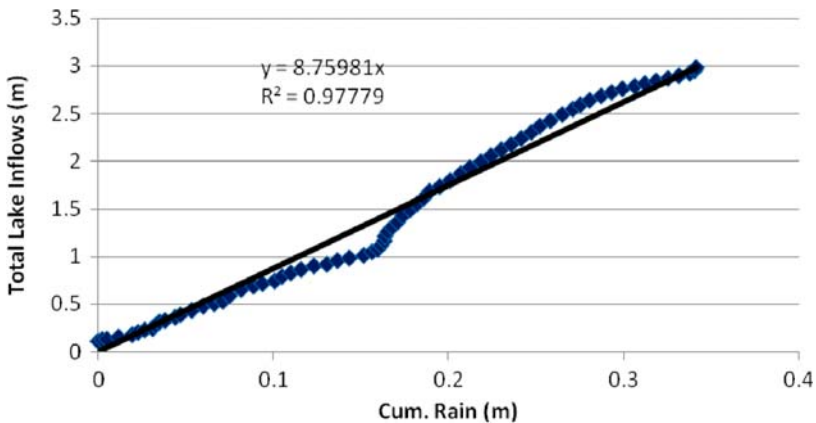Figure 10: Lagged correlations between cumulative rain and total Lake Wakatipu inflows (1999 event).

Figure 11: Lag-14 relation between cumulative rain and total Lake Wakatipu rise (1999 event).

3 MODEL TESTING

3.1 Application to the 1999 flood event

The results of this simulation process were expected to be reasonably good as all the parameters of these models were estimated to fit well with the observed 1999 flood event. Figure 12 confirms these speculations, and all the models performed very well in forecasting the 1999 flood event, from which they were calibrated.

The first model (Model-1, lag-11 non-linear regression), with the highest determination coefficient of 0.997, performed the best. Both the second model (Model-2, lag-11 linear regression) and the third model (Model-3, lag-14 mass balance) performed almost the same, despite the fact that Model-3 has higher determination coefficient (0.978 compared with 0.965 for Model-2).

3.2 Application to an event before the 1999 flood event

The application of the three derived models to an event which occurred during 15–16 May 1999, before the 1999 flood event, showed that Model-2, in general, performed the best, while Model-3 was capable of catching up at the end, and in fact was capable of forecasting the peak lake level better than model-2, as shown in Fig. 13. Model-1, which had the highest determination coefficient for the 1999 flood event, performed poorly and produced unreliable forecasts.

3.3 Application to an event after the 1999 flood event

Another event, which occurred during the period 18–20 September 2002, was simulated by the three models to forecast Lake Wakatipu levels based on cumulative rainfall at the Hillocks rainfall station.

The results of this simulation, as shown in Fig. 14, strengthen the position of Model-2 as the best among the derived three models, as its capability to accurately forecast the lake levels for this event outperformed Model-1 and Model-3. It is worth noting that Model-2 is the simplest model among the three, being linear and doesn't explicitly incorporate lake outflows in its structure. However, complexity of models does not necessarily lead to more accurate results as they usually require estimation of more parameters, which adds to the variability of its results as more uncertain parameters have to be incorporated.
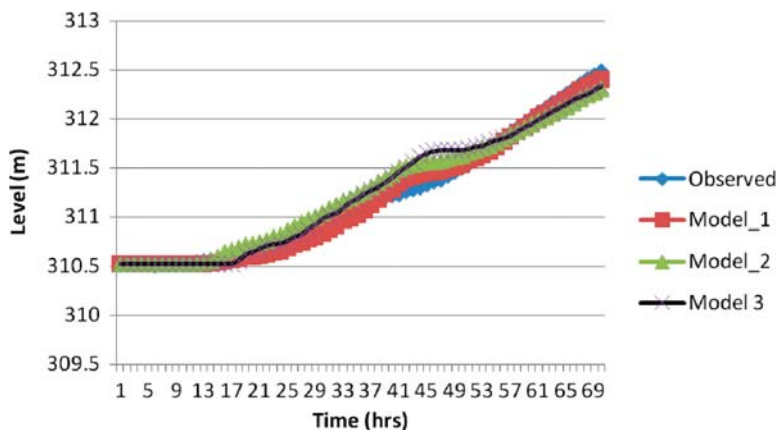

Figure 12: Application of the three models to the 14–17 November 1999 flood event.
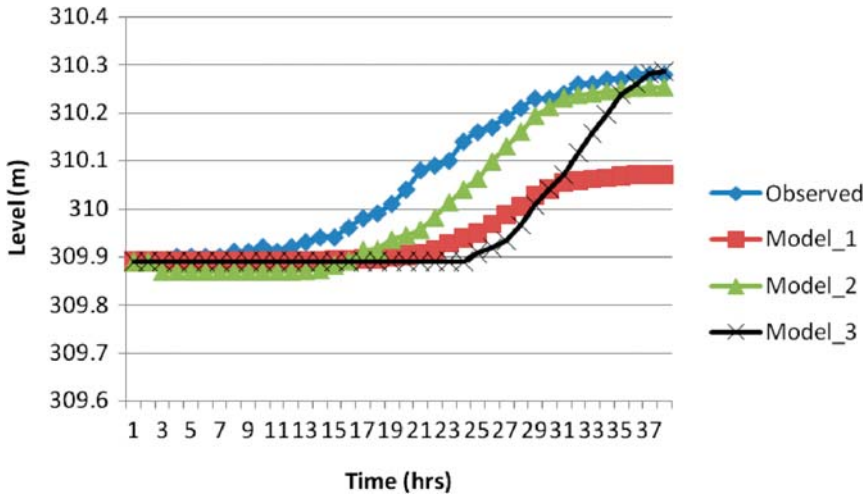
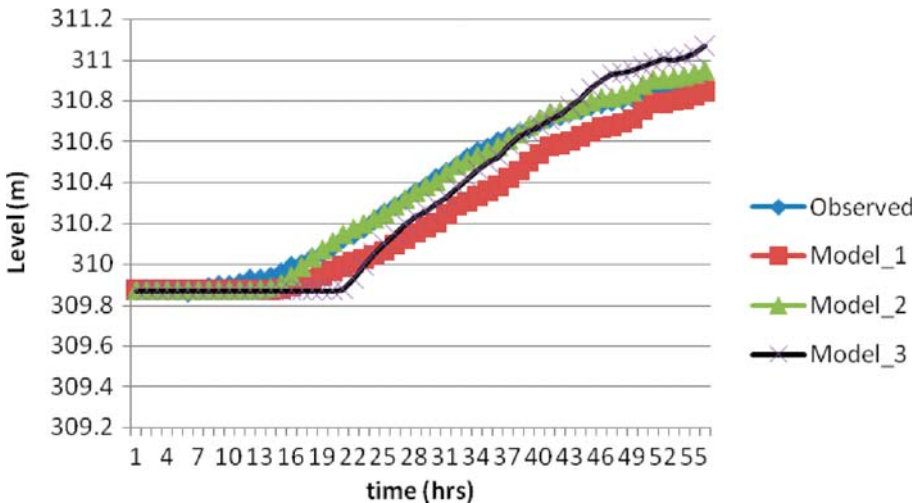Figure 13: Application of the three models to the 15–16 May 1999 flood event.

Figure 14: Application of the 3 models to the 18–20 Sept. 2002 event.

4 CONCLUSIONS

This research provides a new concept for flood modelling of lake levels. The new concept is based on an extensive analysis of the rainfall and lake level data to obtain the best set of lagged series to be fed for the modelling process. Applications to Lake Wakatipu in the South Island of New Zealand indicate that using the most suitable lagged series of total lake rise since the start of the rainfall event produces the best projection of lake rise, based on the cumulative rainfall of that event. A simple linear regression model has been produced, which outperformed the non-linear model and the more sophisticated hourly mass balance model, which also used the optimum lagged series in their derivation. Another major output of this research is that the use of more complex models doesn't necessarily lead to more accurate forecasts.

REFERENCES

[1] Bye, P. & Horner, M., *Easter 1998 floods report by the independent review team to the board of the environmental agency*, vol.**1**, Environmental Agency: Bristol, 1998.

[2] Demeritt, D., Cloke, H., Pappenberger, F., Thielen, J., Bartholmes, J. & Maria-Helena R., Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards*, **7**, pp. 115–127, 2007. doi:http://dx.doi.org/10.1016/j.envhaz.2007.05.001

[3] Chang-Fi John, Lin-YongHuang, Chang-LiChiu & Chang-FiJohn, Enforced self-organizing map neural networks for river flood forecasting. *Journal of Hydrological Processes*, **21(6)**, pp. 741–749, 2007. doi:http://dx.doi.org/10.1002/hyp.6262

[4] Chen-ShenHsien, Lin-YongHuang, Chang-LiChiu & Chang-FiJohn, The strategy of building a flood forecast model by neuro-fuzzy network. *Journal of Hydrological Processes*, **20(7)**, pp. 1525–1540, 2006. doi:http://dx.doi.org/10.1002/hyp.5942

[5] Chow, V.T., Maidment, D.R. & Mays, L.W., *Applied Hydrology*, McGraw Hill, Inc.: New York, pp. 201–342, 1988.

[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] Reklaitis, G.V., Ravindran, A. & Ragsdell, K., *Engineering Optimization*, John Wiley & Sons, Inc.: USA, pp. 87–98, 1983.

[9] Vaziri, M., Predicting Caspian Sea Surface water level by ANN and ARIMA models, *Journal of Waterway, Port, Coast, Ocean Engineering*, **123(4)**, pp. 158–162, 1997. doi:http://dx.doi.org/10.1061/(ASCE)0733-950X(1997)123:4(158)

[10] Brockwell, P.J. & Davis, R.A., *Time Series: Theory and Methods*, Springer-Verlag New York Inc., pp. 451–454, 1991. doi:http://dx.doi.org/10.1007/978-1-4419-0320-4

[11] Coulibaly, P., Reservoir computing approach to great lakes water level forecasting. *Journal of Hydrology*, **381**, pp. 76–88, 2010. doi:http://dx.doi.org/10.1016/j.jhydrol.2009.11.027

[12] Talebizadeh, M. & Moridnejad, A., Uncertainty analysis for the forecast of lake level fluctuations using ensembles of ANN and ANFIS models. *Expert Systems with Applications*. doi:http://dx.doi.org/10.1016/j.eswa.2010.09.075

[13] Moore, R., *et al.*, Forecasting for flood warning. *C.R. Geoscience*, **337**, pp. 203–217, 2005. doi:http://dx.doi.org/10.1016/j.crte.2004.10.017

[14] Smith, J.A., *et al.*, Radar rainfall estimation for flash flood forecasting in small urban watersheds. *Advances in Water Resources*, **30**, pp. 2087–2097, 2007. doi:http://dx.doi.org/10.1016/j.advwatres.2006.09.007

[15] URS New Zealand Ltd, *Flood return period analysis for Lake Wakatipu and Lake Wanaka*, 2003.

[16] Mojzisek, J., *Precipitation variability in the South Island of New Zealand. A thesis submitted for the degree of Doctor of Philosophy at the University of Otago, Dunedin, New Zealand*, 2005.

[17] Ministry for the Environment (MfE), *Preparing for climate change. A guide for local government*, 2008.

[18] Queenstown Lakes District Council and Otago Regional Council, *Learning to Live with Flooding: A flood risk management strategy for the communities of Lakes Wakatipu and Wanaka*, 2006.

[19] Barnett & MacMurray Ltd, *Kawarau and Shotover Rivers sedimentation investigation, client report prepared for Otago Regional Council*, 2006.

[20] Brockwell, P.J. & Davis, R.A., *Time Series: Theory and Methods*, Springer-Verlag New York Inc., pp. 46–51, 1991. doi:http://dx.doi.org/10.1007/978-1-4419-0320-4