

## Application of artificial neural networks in flow discharge prediction for the Fitzroy River, Australia

A. Joorabchi<sup>†</sup>, H. Zhang<sup>†</sup> and M. Blumenstein<sup>‡</sup>

<sup>†</sup>Griffith School of Engineering  
Griffith University, Gold Coast Campus  
Queensland 4217, Australia  
A.Joorabchi@griffith.edu.au  
Hong.Zhang@griffith.edu.au

<sup>‡</sup>School of Information And Communication Technology  
Griffith University, Gold Coast Campus  
Queensland 4217, Australia  
M.Blumenstein@griffith.edu.au



### ABSTRACT

JOORABCHI, A. ZHANG, H. and BLUMENSTEIN, M., 2007. Application of artificial neural networks in flow discharge prediction for the Fitzroy River, Australia. *Journal of Coastal Research*, SI 50 (Proceedings of the 9th International Coastal Symposium), 287 – 291. Gold Coast, Australia, ISSN 0749.0208

Prediction of flow discharge, and in particular floods, in rivers is one of the basic and key information in regards to operation and management of the river systems. The Fitzroy River, one of the largest Australian river systems, has a historical recording of heavy floods and there is a concern for the people of that area to have a clear prediction of the stream discharge to avoid damages. In this paper a feed-forward artificial neural network (ANN) model has been used to forecast the daily flow discharge of the Fitzroy River up to four days ahead. The feed-forward neural network uses error Back-propagation learning algorithm. A cross validation method is applied to prevent the over-fitting problem. The network uses multiple inputs including the daily values of discharge. The network output consists of four neurons in respect to the number of forecasted days. A suitable number of inputs for time-series data were selected by trial and error. Two different multi-layer networks were compared to find the optimised network. The results show an accurate forecasting of flow discharge during flood events. However, the neural network overestimates during low discharge with a mean value of 80 (m<sup>3</sup>/s).

**ADDITIONAL INDEX WORDS:** *Flood, Back-propagation, Time-series prediction*

### INTRODUCTION

The application of artificial neural networks (ANNs) has been widely applied to the various areas. Numerous ANN models have been used as alternatives to the traditional numerical models. For example, in water resource engineering a neural network model was developed for river flood forecasting by CAMPOLO, *et al* (1997). FRENCH *et al.* (1992) applied ANNs to forecast the rainfall intensity. Most recently, SAHOO *et al.* (2006) established a neural network to predict the flash flood and attendant water qualities of a mountainous stream.

Fitzroy catchment is the second largest catchment in Australia after the Murray-Darling Basin. It covers twice the size of Tasmania. The Fitzroy River (Figure 1) is one of the main rivers in this catchment that pass the city of Rockhampton. It has a number of dams and weirs to provide the fresh water to the city and its surrounding area. The importance of this area for scientific study is due to its significant loads of sediment and nutrients that are transported through this river and frequent flood events. In the last few years Coastal Cooperative Research Centres (Coastal CRC) studied this area on a number of projects. A one-dimensional hydrodynamic, sediment transport and biochemistry model is available for the Fitzroy estuary (MARGVELASHVILI *et al.*, 2003 and 2005) based on conceptual model (WEBSTER *et al.*, 2003). KELLY and WANG (1996) study the sediment transport in the Fitzroy River during flood events. Nutrient dynamics and

sediment budgets for the estuary during a flood event are examined by FORD *et al.*, (2006).

Prediction of river flow and in particular flood forecasting is an important element of flood control systems. The early prediction helps us to minimise the flood damage. The results of floods in low lying areas are loss of communication and transport system problems. Due to low annual precipitation, runoff and elevated evaporation rates, Australia has a highly variable flow with large peaks and annual floods (WARNER, 1986). The Fitzroy River has had a number of historical floods caused by heavy rains with extended periods of low flow. It usually happens between January and April (Table 1). A flood warning system has been installed by the Bureau of Meteorology (BOM, 2005) which measures the water elevation and has a warning time of up to 60 hours for floods coming from the hinterland to the Rockhampton city.

The Department of Natural Resources and Water operates a number of stations along the Fitzroy River to control a few parameters including stream water level and discharge (NRW, 2006). In this study a neural network model was developed and trained based on the 64 years of daily discharge measured since 1964 up to the end of 2005 in *The Gap* station (23° 5' 18" S and 150° 6' 28" E). Figure 2 shows the maximum measured value of discharge happened each year (from 1964 to 2006) at *The Gap* station. The presented neural network model, due to independency of the physical parameters such as boundary conditions, initial condition and bathymetry as well as reliable results and real-time prediction of flow discharge, could outper-

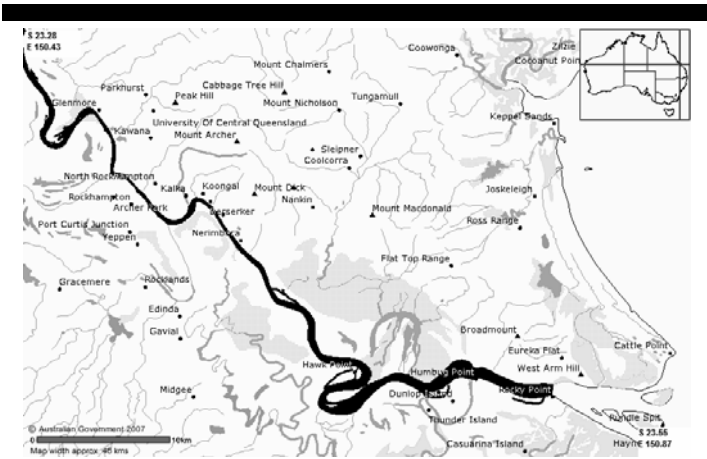


Figure 1. The Fitzroy River, Central Queensland, Australia

Table 1: Historical flood on Rockhampton City (from Bureau of Meteorology)

Date	Flood gauge height (m)
Jan 1918	10.11
Feb 1954	9.40
Jan/Feb 1978	8.15
May 1983	8.25
Mar 1988	8.40
Jan 1991	9.30

form the conventional statistical and numerical models. In general, although training and optimising a neural network for a long period of data is time consuming, simulation of a new data is very fast and can be used for real time forecasting.

**METHODS**

ANNs use a parallel computing system with interconnections similar to biological neural networks. Since the principles of ANN models has been documented in the literature, for example refer to (HAYKIN, 1994) and (WASSERMAN, 1989), the general concept of Back-propagation neural network is summarised here. In particular, application of neural networks in hydrology is described by ASCE, (2000). Back-propagation developed by RUMELHART *et al.* (1986) is one of the most used methods of ANN training in Engineering problems. In this study, two different ANN models with Back-propagation algorithm are applied by trial and error to find an optimum network structure, number of inputs, outputs and hidden layers (Figure 3). The input layer includes 15 inputs, the hidden layers have 10 neurons and the output layer consists of 4 neurons as well as interconnection weights and biases are also depicted. The input layer receives the inputs from the training data and the hidden layer and output layer receives it from the interconnections. Neurons use Log-sigmoid and hyperbolic tangent sigmoid transfer functions to produce their output. In the Back-propagation algorithm, interconnection weights are adjusted according to the error convergence technique to obtain the required output for the given inputs. The interconnection weights are adjusted based on the following equation

$$\Delta w_{ij}(n) = -\epsilon * \frac{\partial E}{\partial w_{ij}} + \alpha * \Delta w_{ij}(n-1) \quad (1)$$

where  $\epsilon$  and  $\alpha$  are the learning rate and the momentum, respectively.  $\Delta w_{ij}(n)$  and  $\Delta w_{ij}(n-1)$  in equation (1) are the increment of weights between nodes  $i$  and  $j$  for the  $n$ -th and  $(n-1)$ -th iteration.

Training of a neural network is controlled by supervised or unsupervised learning algorithms. The difference between these two is that in supervised learning both input and output data are required for calculation of the error of the network based on the difference between the calculated output and given output. However, unsupervised training uses the input data only. In this model, a supervised learning algorithm is used. Selection of the number of hidden layers and number of neurons in each layer is one of the most important factors in the application of the neural network. In general, there is not a certain rule to estimate

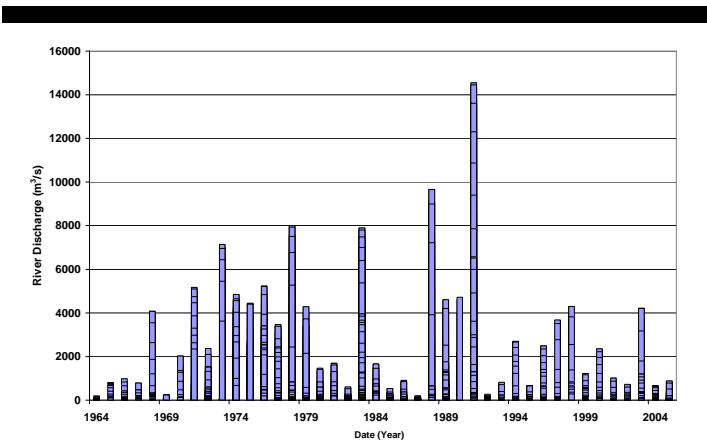


Figure 2. Maximum flow discharge happened each year at Fitzroy River (The Gap station) from 1964 to 2006. (Data source: NRW, 2006)

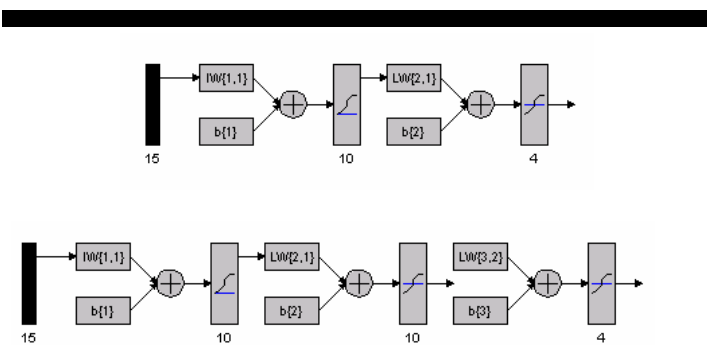


Figure 3. Structure of the proposed artificial neural networks with 1 hidden layer with 10 neurons (top) and 2 hidden layers with 10 neurons in each layer (bottom). IW and LW indicate interconnection weights and b denotes biases for each layer.

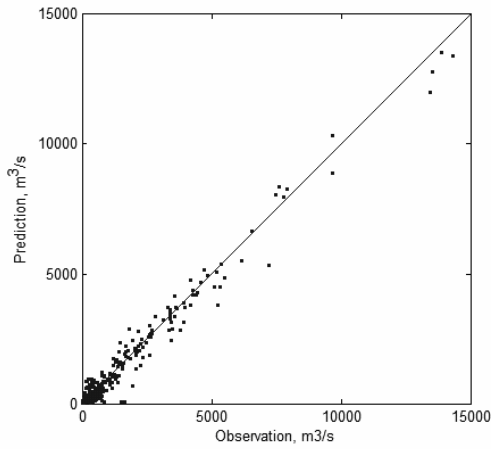


Figure 4a. First day predictions ( $R=0.99$ ,  $RMSE=139 \text{ m}^3/\text{s}$  and  $SI=0.77$ )

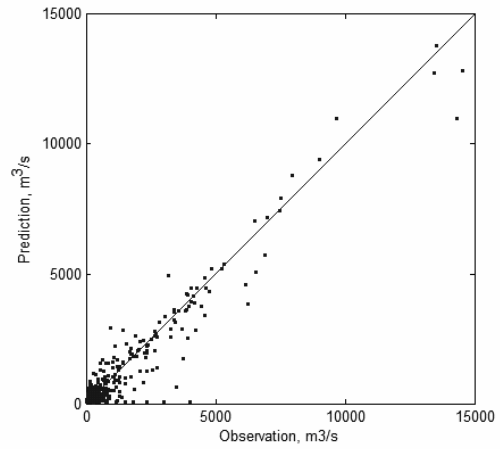


Figure 4b. Second day predictions ( $R=0.97$ ,  $RMSE=256 \text{ m}^3/\text{s}$  and  $SI=1.25$ )

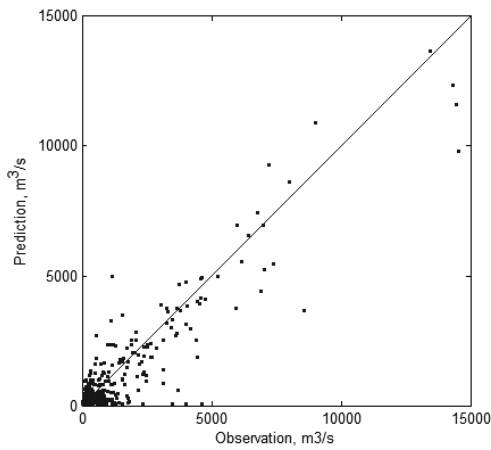


Figure 4c. Third day predictions ( $R=0.93$ ,  $RMSE=315 \text{ m}^3/\text{s}$  and  $SI=1.76$ )

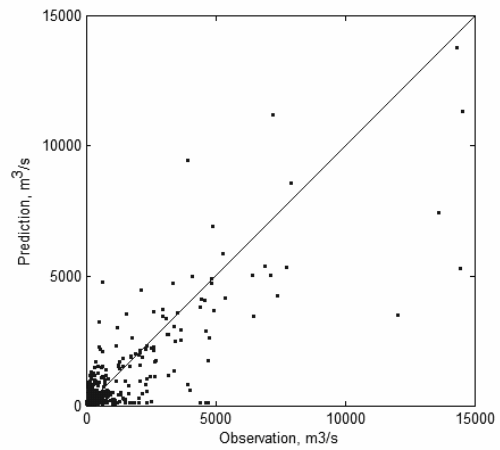


Figure 4d. Fourth day predictions ( $R=0.85$ ,  $RMSE=442 \text{ m}^3/\text{s}$  and  $SI=2.52$ )

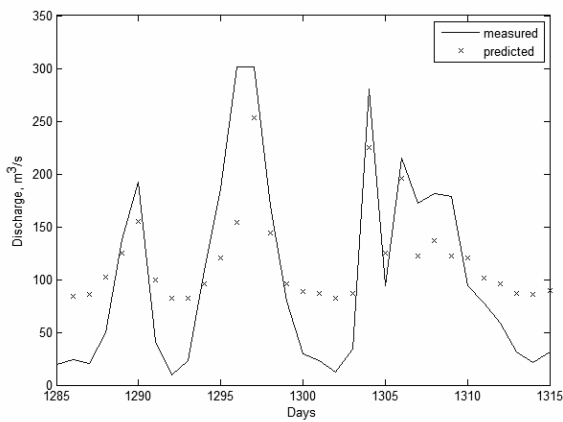


Figure 5a. Time series of measured and predicted flow discharge in the first day of prediction for low discharge values.

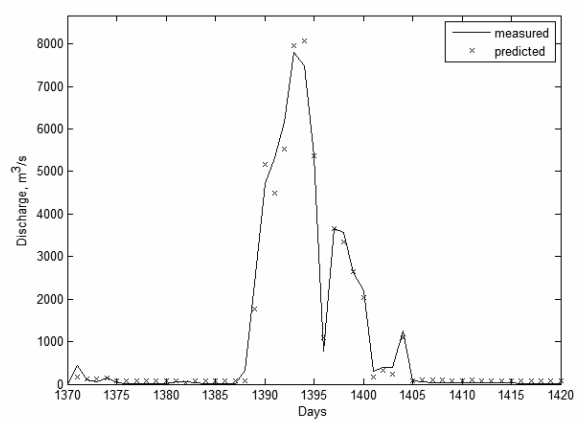


Figure 5b. Time series of measured and predicted flow discharge in the first day of prediction for high discharge values.

Table 2: Verification statistics of Fitzroy River flow discharge for different predicted day. Here,  $R$  is correlation coefficient,  $RMSE$  is root mean square error and  $SI$  is scatter index.

Network	Days	$R$	$RMSE$ ( $m^3/s$ )	$SI$
ANN with 1 hidden layer	1	0.97	149	0.94
	2	0.93	253	1.59
	3	0.86	358	2.23
	4	0.68	531	3.28
ANN with 2 hidden layer	1	0.99	139	0.77
	2	0.97	256	1.25
	3	0.93	315	1.76
	4	0.85	442	2.52

the number of neurons for hidden layers and the optimum topology is obtained by trial and error. Some methods such as cascade-correlation (FAHLMAN and LEBIERE, 1990) are developed to find the optimum hidden units. In some cases with an extensive number of inputs and data sets, the higher number of neurons in the hidden layers and number of hidden layers prevents the neural network from generating poor results and increases the accuracy of forecasting. However, a greater number of hidden layers and neurons causes more computation time and for some particular cases it may result in over-fitting. In general, over-fitting is one of the problems that happens during training of a neural network. Since the network has memorised the training pattern it can not generalise to new data. In this case the error for a training set is very small but the network results into large error for new data sets. Cross validation method was used to improve the generalisation. The existing data set consisted of 14960 training pairs with 15 daily river flow discharge values in cubic metres per second and 4 outputs representing the river discharge in the next four days. The data was divided into three subsets include training, validation and testing sets. The validation error is monitored during the training process. The errors on the validation and training set usually decrease during the initial phase of training and as soon as the network begins to overfit the data, the validation error increases for a number of iterations (here it is five iterations), the training is stopped.

## RESULTS AND DISCUSSIONS

The ANN model was validated in terms of a correlation coefficient ( $R$ ), root mean square error ( $RMSE$ ) and scatter index ( $SI$ ).

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - x_i)^2}{N}} \quad (2)$$

$$R = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (3)$$

$$SI = \frac{RMSE}{\bar{x}} \quad (4)$$

where  $x_i$  is the observed values at the  $i$ -th time step,  $y_i$  is the simulated values,  $N$  is the number of time increments,  $\bar{x}$  and  $\bar{y}$  are the mean value of observations and simulations, respectively.

Simulation results for testing set (Table 2) show that the neural network with two hidden layers could outperform the network with one hidden layer. The difference between two proposed network structures was increasing with more number of flow discharge predictions. The proposed neural network with two hidden layers resulted in 20% more accuracy for prediction of the flow discharge in the fourth day. Therefore, the multi-layer neural network with two hidden layers was selected to simulate and forecast the flow discharge time-series.

After 28 iterations, training is stopped by cross validation method to avoid the over-fitting. When short prediction intervals of one to three days are considered, the flow discharge simulated by this network has a high accuracy. Figures 4a to 4d illustrate the comparison between the measured and the predicted flow discharge with 2992 number of testing pairs (equal to 20% of total data) for the first to fourth days of prediction. The  $RMSE$  are equal to 139, 256 and 315 ( $m^3/s$ ) for the first to the third day flow discharge prediction respectively. The correlation coefficients that indicate the strength of the relationship between observed and predicted data are higher than 0.9 (maximum scale is 1) for the first three days. The best result was predicted for the first day prediction with a correlation coefficient equals 0.99. The scatter indexes for the first three day predictions are 0.77, 1.25 and 1.76 respectively. The prediction of flow discharge is less reliable for the fourth day. However, it has reasonable values equal to 442  $m^3/s$ , 0.85 and 2.52 for  $RMSE$ ,  $R$  and  $SI$  respectively.

The neural network overestimated the discharge for dry seasons (Fig. 5a) with a mean value equals 80 ( $m^3/s$ ). It is due to extreme variation between flood events and dried season with maximum value of 15000 ( $m^3/s$ ). However, the main application of this model is prediction of flood and that the neural network could predict it with a high accuracy (Fig 5b). Figure 5b indicates the peak flows (floods) are well simulated. There is a potential further study to investigate the application of different learning algorithm as well as the sensitivity of mapping scale to avoid the over-prediction in dry seasons.

## CONCLUSION

An artificial neural network model with a feed-forward Back-propagation learning algorithm is adopted to predict the daily flow discharge in the Fitzroy River up to four days ahead by a learning and recalling process. Two different network structures were compared. The network with two hidden layers had better outcomes. A cross validation method was used to prevent the network from over-fitting. Results show the neural network provides a high accuracy prediction of flow discharge for the next three days and has a reliable prediction for the fourth day. One of the advantages of the presented model compared to the ordinary numerical models is that it is not dependent on the initial and boundary conditions. However, the reliability of results is dependent on how much data is available for training the network.

The proposed neural network model can predict flood events very accurately. However, it over predicts low flow discharge

with average value of 80 (m<sup>3</sup>/s). Further investigation is suggested to use other network structures, transfer functions and mapping scale to expand this model from flood prediction to drought prediction.

### ACKNOWLEDGEMENTS

The authors are grateful to Prof. Bofu Yu for his helpful suggestions. The authors are also thankful for support from Griffith University Research Grant (GURG, 2006). The Department of Natural Resources and Water is gratefully acknowledged for providing data.

### LITERATURE CITED

- ASCE (Task Committee on Application of Artificial Neural Networks in Hydrology), 2000b. Artificial neural networks in hydrology. II: Hydrologic applications. *Journal of Hydrologic Engineering* 5(2), 124–137.
- BOM (Bureau Of Meteorology), 2005. Flood warning system for the Fitzroy River. [online] Available from: <<http://www.bom.gov.au/hydro/flood/qld/brochures/fitzroy/fitzroy.shtml>> (Accessed 15/08/2006).
- CAMPOLO, M., ANDREUSSI, P. and SOLDATI, A., 1997. River flood forecasting with a neural network model. *Water Resources Research*, 35(4), 1191–1197.
- FAHLMAN, S.E. and LEBIERE, C., 1990. The cascade-correlation learning architecture. In: TOURETZKY, D.S. (Ed.), *Advances in Neural Information Processing Systems 2*. Morgan Kaufmann San Mateo, CA.
- FORD, P.W. 2006. Nutrient dynamics and sediment budgets in the Fitzroy estuary during a flood event. *Coastal CRC Technical Report No. 72*. CRC for Coastal Zone, Estuary and Waterway Management; Indooroopilly, 20p.
- FRENCH, M.N., KRAJEWSKI, W.F., and CUYKENDALL, R.R., 1992. Rainfall forecasting in space and time using a neural network. *Journal of Hydrology*, 137, 1–31.
- HAYKIN, S., 1994. *Neural networks: a comprehensive foundation*. Mac-Millan, New York, 842p.
- KELLY, J. N. and WONG W. T., 1996. Sediment transport in the Fitzroy River during flood events. *Proceedings Stream Management '96: 1st Australian Stream Management Conference, 19-23 February 1996, Merryjig, Victoria*. (Eds: I RUTHERFORD, and M. WALKER) pp. 19-21. CRC for Catchment Hydrology, Clayton, Victoria.
- MARGVELASHVILI, N., ROBSON B., SAKOV P., WEBSTER I.T., PARSLow J., HERZFELD M. and ANDREWARTHA J., 2003. Numerical modelling of hydrodynamics, sediment transport and biogeochemistry in the Fitzroy estuary. *Coastal CRC Technical Report No. 10*. CRC for Coastal Zone, Estuary and Waterway Management; Indooroopilly, 70p.
- MARGVELASHVILI, N., HERZFELD M., and WEBSTER I.T., 2005. Modelling of fine sediment transport in Fitzroy Estuary and Keppel Bay. *CRC CZEWM Technical Report 9*. Cooperative Research Centre for Coastal Zone, Estuary, and Waterway Management. Indooroopilly, Australia, 40p.
- NRW (Department of Natural Resources and Water), 2006. Watershed database. [online] Available from: <<http://www.nrw.qld.gov.au/watershed/index.html>> (Accessed 12/06/2006).
- RUMELHART, D.E., HINTON, G.E., and WILLIAMS, R.J., 1986. Learning representations by Back-propagating errors. *Nature* 323, 533–536.
- SAHOO, G.B., RAY, C., and DE CARLO, E.H., 2006. Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii. *Journal of Hydrology*, 327(3-4), 525-538.
- WARNER, R.F., 1986. Hydrology. In: JEAN, D.N., (ed.), *Australia – A Geography Volume One: The Natural Environment*, Sydney University Press. pp. 49-79
- WASSERMAN, P. D., 1989. *Neural computing: theory and practice*. Van Nostrand Reinhold, New York, 230p.
- WEBSTER, I.T., FORD P.W., ROBSON B., MARGVELASHVILI N., and PARSLow J., 2003. Conceptual model of the hydrodynamics, fine sediment dynamics, biogeochemistry and primary production in the Fitzroy estuary. *Coastal CRC Technical Report No. 8*. CRC for Coastal Zone, Estuary and Waterway Management; Indooroopilly, 43 p.