23rd International Congress on Modelling and Simulation, Canberra, ACT, Australia, 1 to 6 December 2019
mssanz.org.au/modsim2019

Estimating Lake Mulwala Diversions for calibration of a semi-distributed hydrologic model of the Murray River

T. L. Ng[a], D. E. Robertson[a] and J. C. Bennett[a, b]

[a] *CSIRO Land & Water, Bayview Ave, Clayton, Victoria*
[b] *Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania*

Email: seline.ng@csiro.au

Abstract: In hydrologic model calibration, storage diversions are often neglected due to the lack of sufficient data. While it may make little difference to include diversions where the points of interest are far downstream of the diversions, the same is not true where the distances between the former and latter are small. In this paper, we develop a series of equations using least squares optimisation to estimate the flow diversions from Lake Mulwala on the Murray River to Mulwala Canal in the north and Yarrawonga Main Channel in the south. We then use the equations to estimate the diversions for times when historical observations are unavailable. This is to provide data for calibrating a hydrologic model of the section of the Murray River between Hume Dam and Lake Mulwala. The hydrologic model is developed using the Short-term Water Information and Forecasting Tools (SWIFT2) and is intended for generating forecasts of the hourly inflow to Lake Mulwala when provided rainfall forecasts. The results show the equations to be most accurate when predicting the diversions in summer when the flows are highest, and when predicting the diversion to Mulwala Canal, the larger of the two diversions. Further, results of the calibration and validation performance of the hydrologic model show including the Lake Mulwala diversions (as opposed to neglecting them as is often the case where real data are unavailable) substantially improves the model fit, thus increasing the model credibility.

Keywords: *Storage diversions, hydrologic modelling, model calibration, SWIFT2*

1. INTRODUCTION

When calibrating hydrologic models of catchments with reservoirs and other storage facilities, flow diversions such as for irrigation are often ignored, typically due to the lack of detailed flow data over the desired time interval. While it may make little difference to include diversions where the points of interest are far downstream of the diversions, the same is not true where the distances between the former and latter are small. When this is the case, it may be advantageous to include diversions in the model calibration process for more accurate results. To do so, however, requires first estimating their values.

In the literature, a common approach to estimating flow diversions, particularly for irrigation, is to apply a crop evapotranspiration model. Wu and Chen (2013) applied an adapted Penman model (Doorenbos and Pruitt, 1977) to estimate crop evapotranspiration as a function of solar radiation, temperature, humidity, wind speed, a weighting factor, and a crop coefficient that is specific to the crop type and stage of growth. From there, the authors estimated the irrigation demands of a river basin in China to define the objective of a reservoir operation optimisation model. Similarly, Kirby et al. (2013) computed crop evapotranspiration using the Food and Agriculture Organisation (FAO) Penman-Monteith equation (Allen et al., 1998) and used it, together with available reservoir storage, to estimate monthly irrigation diversions in the Murray-Darling Basin in Australia. Another common approach to estimating flow diversions is to use statistical models, which, once parameterised, are relatively easier to use than crop evapotranspiration models. Martin (1990) used a series of linear regression equations to predict monthly irrigation diversions to three rice irrigation systems in Texas, United States given rainfall, crop acreage, crop planting time, and lake evaporation. The equations are useful for future planning and assessment purposes. Pulido-Calvo (2007) also provide examples of neural network models for forecasting the daily irrigation demands of an irrigation district in Spain in real time.

In this paper, we develop statistical models to predict the flow diversions from Lake Mulwala in the Murray-Darling Basin in Australia. We have selected this approach considering the availability of historical flow data necessary for parameterising the equations. Also, the approach is more direct and less data intensive than using a crop evapotranspiration model.

Lake Mulwala is an artificial lake that was formed with the damming of the Murray River by Yarrawonga Weir. The lake is a major source of irrigation water to surrounding farms. We develop two sets of equations; the first to predict the daily diversion to Mulwala Canal (MC) in the north, and the second to predict the daily diversion to Yarrawonga Main Channel (YMC) in the south. In both sets of equations, the equations are differentiated by day of the year, and are dependent on a single input representing the sum of flows at upstream gauges.

We use the models developed to predict the diversions from Lake Mulwala for times when historical diversion data are unavailable. This is to supply data for calibrating and validating a hydrologic model of the 237 km section of the Murray River from Hume Dam at the upstream end to Lake Mulwala at the model outlet. The model includes the Ovens and Kiewa Rivers, both tributaries of the section of the Murray modelled. The model is developed using version 2 of the Short-term Water Information and Forecasting Tools (SWIFT2) (Perraud et al., 2015), and is intended for generating forecasts of the hourly inflow to Lake Mulwala when provided rainfall forecasts. See Figure 1 below for a schematic of the system.

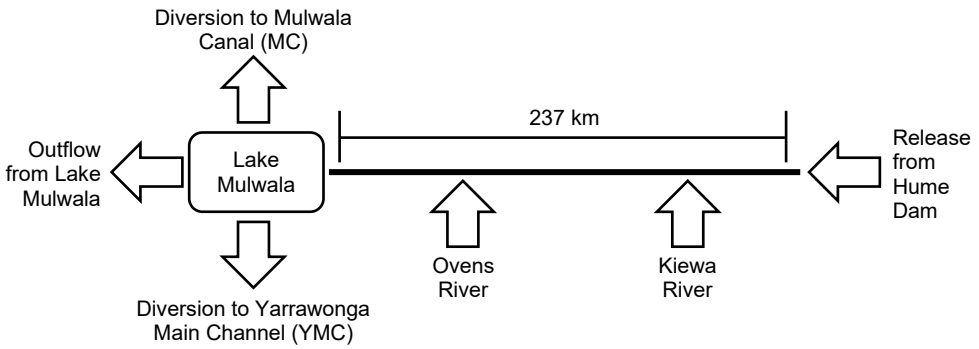


Figure 1. Schematic of the system

2. METHODS AND DATA

2.1. Statistical Models of Lake Mulwala Diversions

To model the two diversions from Lake Mulwala to MC and YMC, we consider the following equations:

$$y_1 = c_1(a_1/b_1)(x/b_1)^{a_1-1}e^{-(x/b_1)^{a_1}} \quad (1)$$

$$y_2 = c_2 - c_2 e^{-(x/b_2)^{a_2}} \quad (2)$$

$$y_3 = \frac{c_3}{x b_3 \sqrt{2\pi}} \exp\left(-\frac{(\ln x - a_3)^2}{2b_3^2}\right) \quad (3)$$

$$y_4 = \frac{c_4}{2} + \frac{c_4}{2} \operatorname{erf}\left[\frac{\ln x - a_4}{\sqrt{2} b_4}\right] \quad (4)$$

$$y_5 = c_5 \sqrt{\frac{b_5}{2\pi x^3}} \exp\left(-\frac{b_5(x-a_5)^2}{2a_5^2 x}\right) \quad (5)$$

$$y_6 = a_6 + b_6 x \quad (6)$$

Equation (1) is adapted from the Weibull probability distribution, equation (3) the log-normal probability distribution, and equation (5) the inverse Gaussian probability distribution. They are selected for consideration here for their positive skew. Equation (2), adapted from the Weibull cumulative probability distribution, and equation (4), adapted from the log-normal cumulative probability distribution, are selected for their non-decreasing property. Lastly, equation (6), which is linear, is selected for its simplicity and for completeness sake. y_i is the dependent output from equation (i) representing either the MC or YMC diversion on a particular day. x is the sole independent input to the equations representing the sum of the flows on the same day at an upstream at the Murray River at Corowa and at a second upstream gauge at the Ovens River at Peechelba. x is the same across all six equations.

a_i, b_i and c_i are parameters. To estimate them, we fit equations (1) to (6) to historical daily values of y_i and x by least squares optimisation. In the case of y_i, the historical data are from J. Hughes (personal communication, 2018) of CSIRO Land and Water, and the Department of Environment, Land, Water and Planning (DELWP) of Victoria. In the case of x, the data are from the Victorian DELWP and the Department of Industry (DoI) of New South Wales (NSW). All data from the two state agencies are from Water Data Online (www.bom.gov.au/waterdata/). Figure 2 shows the time periods corresponding to the historical y_i and x data used in this study. For comparison, the figure also shows the calibration and validation period of the SWIFT2 hydrologic model of the Murray River developed (Section 2.2). There are altogether 2652 x-y_i data pairs for fitting equations (1) to (6) in the case of the MC diversion, and 3048 pairs in the case of the YMC diversion.

We derive two sets of a_i, b_i and c_i, one for computing the MC diversion and the other the YMC diversion. Each set of a_i, b_i and c_i comprises 37 subsets of values to differentiate the parameters by day of the year. The values in the first subset apply when computing results for the first 10 days of the year, the second subset the second 10 days of the year, and so on. The values of a_i, b_i and c_i in the 37th subset apply when computing results for the last 5 days of the year, or last 6 days in a leap year.

To derive the a_i, b_i and c_i values, we utilize the fminsearch function in MATLAB (an unconstrained nonlinear optimisation routine based on the Nelder-Mead simplex algorithm). To ensure a good solution quality, we run fminsearch in a multistart manner. We then select the best performing equation for each of the 37 ordinal subsets of each of the two sets of the parameters, according to whichever equation yields the smallest root mean square error (RMSE).

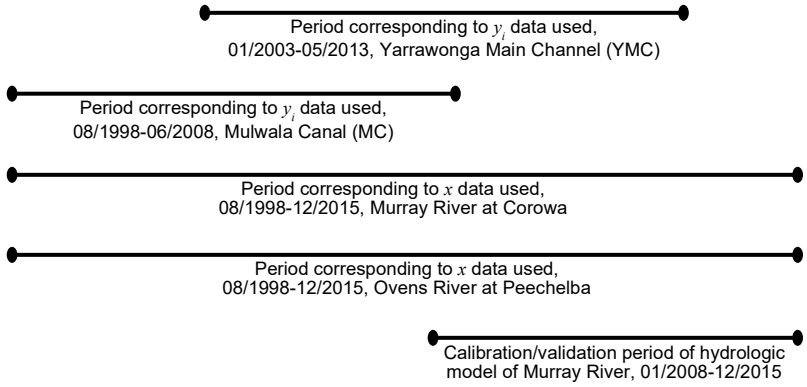


Figure 2. Time periods corresponding to the historical x and y_i data used in this study, and calibration and validation period of the hydrologic model of the Murray River developed; x and y_i are as defined in equations (1) to (6)

2.2. Hydrologic Model of Murray River

As mentioned above, the statistical models from Section 2.1 are to provide estimates of the MC and YMC diversions from Lake Mulwala for calibrating a hydrologic model of the section of the Murray River between Hume Dam and the lake. In this section, we describe the hydrologic model and the steps taken to calibrate and validate it.

We develop the model using SWIFT2 (Perraud et al., 2015). It combines the 4-parameter GR4J rainfall-runoff model (Perrin et al., 2003), the lag-and-route flow routing model and the Error Reduction and Representation in Stages (ERRIS) model (Li et al., 2017). ERRIS is an error model comprising 4 stages and is designed to reduce errors in streamflow forecasts.

The model runs on an hourly time step. It comprises 87 nodes, 86 links and 86 subareas. About 30% of the nodes, links and subareas are to model the stretch along the Murray River of interest, 60% to model the Ovens River (a tributary to the Murray), and 10% to model the Kiewa River (also a tributary to the Murray).

Outputs from the model are the hourly streamflow at each node of the model. Inputs to the model are potential evapotranspiration (PET) and rainfall for each subcatchment. We interpolate monthly PET to subareas by simple weighted averaging of gridded data from the Australian Water Availability Project (AWAP) (Raupach et al., 2012; Raupach et al., 2009), then disaggregate the monthly data to hourly time steps according to theoretical solar insolation patterns based on latitude and time of the year. We interpolate precipitation to subareas from gauge observations using inverse distance squared interpolation; hourly historical rain gauge data are obtained from the Bureau of Meteorology's Australian Integrated Forecast System (AIFS) database. A third input to the model is the flow release from Hume Dam to the Murray, which we fix according to historical values as reported by the NSW DoI on Water Data Online.

We calibrate the hydrologic model of the Murray to hourly streamflow observations at 21 sites, including the inlet to Lake Mulwala, and outlets of the Ovens and Kiewa at Peechelba and Bandiana respectively. The historical observations are from the NSW DoI and the Victorian DELWP, and are taken from Water Data Online. In the case of the inlet to Lake Mulwala, which is the most downstream of the 21 sites, real observations are unavailable. Thus, to estimate the historical flow at that point, we use the equations developed in Section 2.1 to estimate the diversions from Lake Mulwala, then add those estimates to real observations from a gauge just downstream of the lake.

To calibrate the hydrological model, we take a sequential approach. We first calibrate the links and subareas above the sites that are most upstream, then those between the most upstream sites and the sites next closest downstream, and so on. To evaluate the model, we conduct a buffered leave-1-year-out cross-validation, with a buffer of 1 year, over the 8-year period from 2008 to 2015 (Figure 2). That is, for example, to generate cross-validated simulations for 2011, we use all years' data except 2011-2012 to calibrate the model.

Each link in the model is defined by 2 lag-and-route parameters, and each subcatchment by 4 GR4J parameters. In the model calibration process, shuffled complex evolution (SCE) (Duan et al., 1994) is used to jointly determine their values. SCE is also used to determine the 11 parameters of the ERRIS model for each of the 21 calibration sites. The lag-and-route and GR4J parameters are determined by maximising for the Nash-Sutcliffe Efficiency (NSE) between the simulated and observed data, while the ERRIS parameters by maximising a likelihood function (Li et al., 2017).

3. RESULTS AND DISCUSSION

Figure 3 compares observed and predicted values of the MC and YMC diversions. (The predicted values are as simulated by the statistical models from Section 2.1.) Overall, the predicted values of the MC diversion fit better to the observed data than the predictions of the YMC diversion (Figure 3) according to two goodness-of-fit measures, the coefficient of determination, R^2 and the NSE. The relative underperformance of the YMC predictions, while less than ideal, is not a major concern as the magnitudes of the YMC diversion are much smaller than the MC diversion. Hence, considering the ultimate purpose of estimating the two diversions (i.e., for calibrating the SWIFT2 hydrologic model of the Murray from Section 2.2), it is not as critical the accuracy of the YMC predictions as for the MC predictions.

Figure 4 shows the individual performances, in terms of R^2, of the statistical models from Section 2.1. The figure also shows the average MC and YMC diversions for each 10-day period in the year. As seen in Figure 3, the YMC predictions are less accurate than the MC predictions. In general, the equations are most accurate when predicting the diversions in summer (when they are highest), but less so at other times. Note though that the diversions are mostly zero in winter when conditions are generally wetter, and therefore, no models are developed for that period.

When fitting the six equations in Section 2.1 to observed data, equation (1) fits best most often, followed by equation (3), then equation (4). Equations (2) and (5) are least utilized. The calibrated equation parameters are available by email request to the first author of this paper.

Figure 5 gives the calibration and validation performance of the hydrologic model of the Murray River (Section 2.2) when neglecting the diversions from Lake Mulwala in the model calibration and validation process, and when including them. The results are for the most downstream node in the model corresponding to the inlet to Lake Mulwala. They are obtained by calibrating the model against observed hourly streamflow data for the periods 2008-2010 and 2013-2015, then validating it against the same for the year 2011. This is done first to yield one set of model parameters and with it, generate the results in the top row of Figure 5. The process next repeats to yield a second separate set of parameters and to generate the results in the bottom row. Due to space limitation, we omit showing the results for other calibration and validation years. This is justified as they all show the same trends as in Figure 5. Neither do we show the results for other calibration sites as flows at those other sites, being upstream of Lake Mulwala, are unaffected by the MC and YMC diversions.

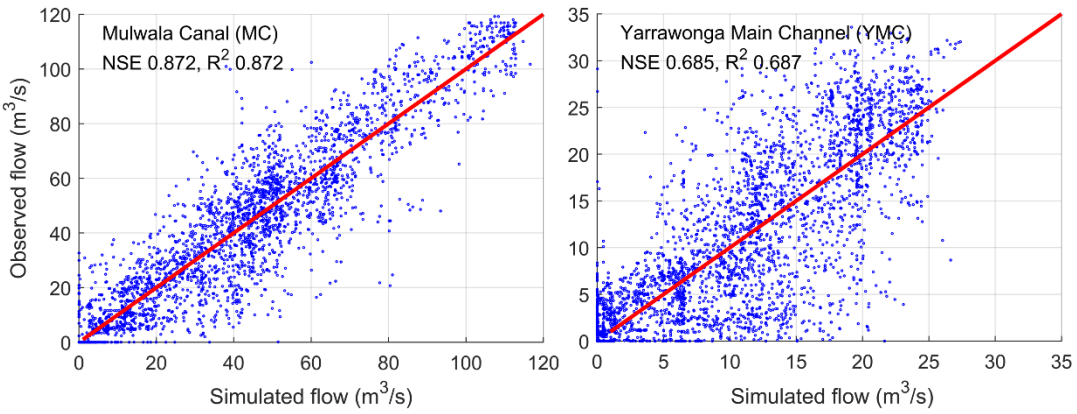


Figure 3. Overall fit of the statistical models developed to predict the MC (left panel) and YMC (right panel) diversions; insets give the model goodness-of-fit in terms of R^2 and NSE

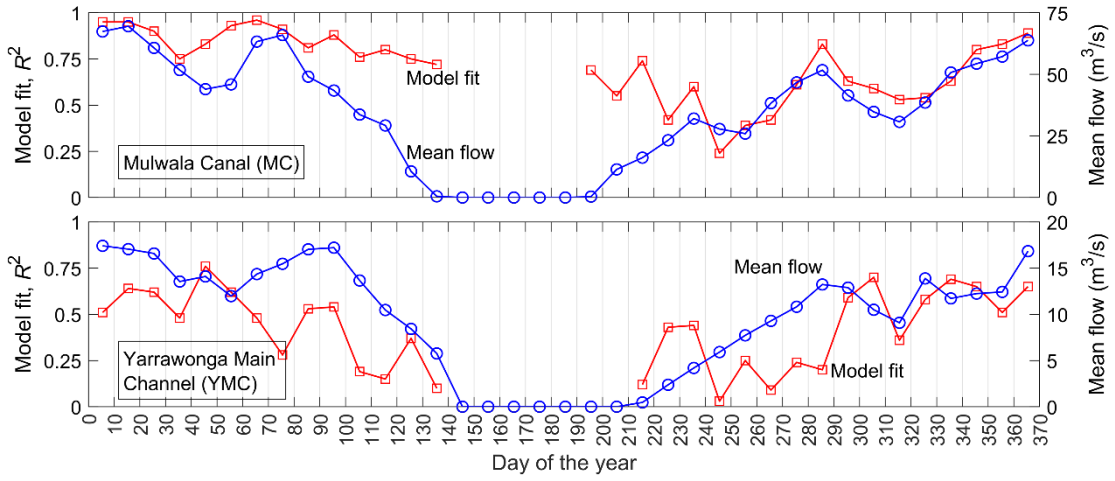


Figure 4. Individual performances, in terms of R^2, of the statistical models developed (red lines, left axis), and the average observed flows to MC and YMC (blue lines, right axis) for each 10-day period in the year; note that the flows in winter are mostly zero, and thus, no models are developed for that period

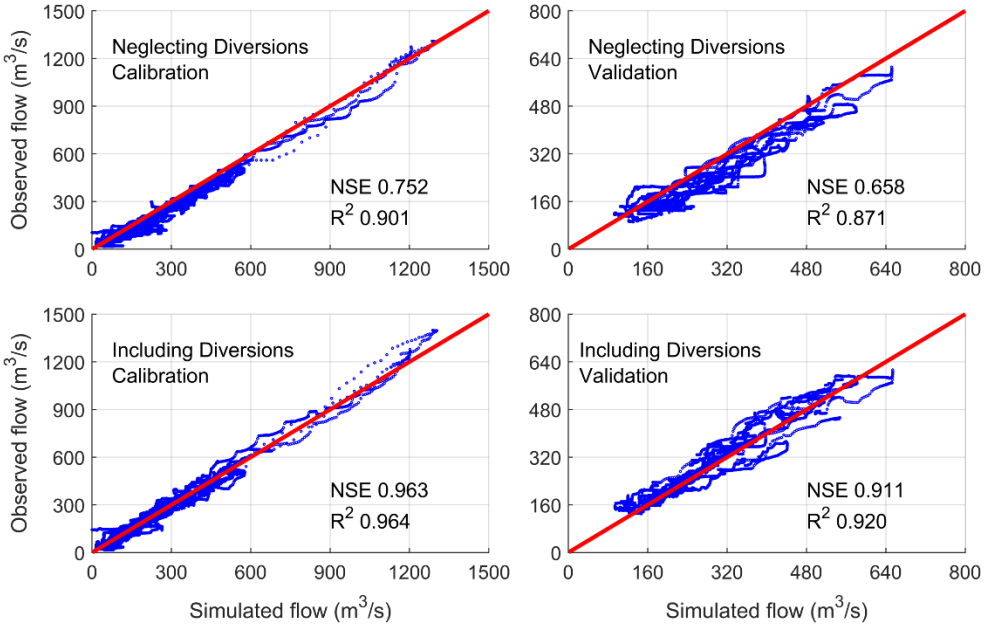


Figure 5. Calibration (left panels) and validation (right panels) performance of the hydrologic model of the Murray River from Section 2.2; top panels show the performance when the MC and YMC diversions are neglected in the model calibration and validation; bottom panels show the performance when the MC and YMC diversions are included in the model calibration and validation; insets show the model fit in terms of R^2 and NSE

Figure 5 shows that including the diversions from Lake Mulwala greatly improves the model fit, as can be seen from the considerably larger R^2 and NSE values when including the diversions. This increases the model credibility, allowing users more confidence when using the model. Note though, that including the diversions does not necessarily lead to significant differences in the simulated inflow to Lake Mulwala. This is because

the simulated flow entering the lake is mostly governed by upstream flows that in turn are a function of parameters that have been determined independently from the Lake Mulwala diversions.

4. SUMMARY AND CONCLUSIONS

In this paper, we develop a series of equations to predict the flow diversions from Lake Mulwala to MC and YMC. The equations are developed by data fitting using least squares optimisation. R^2 and NSE values show the predictions to be most accurate in summer when the flows are highest. Further, predictions of the MC diversion (the larger of the two diversions) are more accurate than predictions of the YMC diversion. We then use the equations developed to estimate the diversions for times when historical observations are unavailable for informing the calibration and validation of a hydrologic model of the Murray River between Hume Dam and Lake Mulwala. Results of the calibration and validation performance of the hydrologic model show that including the Lake Mulwala diversions (as opposed to neglecting them as is often the case where real data are unavailable) substantially improves the model fit, thus increasing the model credibility.

ACKNOWLEDGMENT

We thank the Bureau of Meteorology for providing hourly rainfall observations used in this study.

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D., Smith, M. (1998). Crop evapotranspiration: guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56. Food and Agriculture Organization of the United Nations, Rome.

Doorenbos, J., Pruitt, W.O. (1977). Guidelines for predicting crop water requirements. Water Resources and Development Service, Food and Agriculture Organization of the United Nations, Rome.

Duan, Q., Sorooshian, S., Gupta, V.K. (1994). Optimal use of the SCE-UA global optimization method for calibrating watershed models. Journal of Hydrology, 158(3-4), 265-284.

Kirby, J.M., Mainuddin, M., Ahmad, M.D., Gao, L. (2013). Simplified monthly hydrology and irrigation water use model to explore sustainable water management options in the Murray-Darling Basin. Water Resources Management, 27, 4083-4097.

Li, M., Wang, Q.J., Robertson, D.E., Bennett, J.C. (2017). Improved error modelling for streamflow forecasting at hourly time steps by splitting hydrographs into rising and falling limbs. Journal of Hydrology, 555, 586-599.

Martin, Q.W. (1990). Estimating irrigation diversions for major Texas rice-growing area. Journal of Irrigation and Drainage Engineering, 116(2), 243-260.

Perraud, J.-M., Bridgart, R., Bennett, J.C., Robertson, D. (2015). SWIFT2: high performance software for short-medium term ensemble streamflow forecasting research and operations. Proceedings of the 21st International Congress on Modelling and Simulation, 29 Nov-4 Dec, Gold Coast, Australia.

Perrin, C., Michel, C., Andreassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. Journal of Hydrology, 279, 275-289.

Pulido-Calvo, I., Montesinos, P., Roldan, J., Ruiz-Navarro, F. (2007). Linear regressions and neural approaches to water demand forecasting in irrigation districts with telemetry systems. Biosystems Engineering, 97(2), 283-293.

Raupach, M.R., Briggs, P.R., Haverd, V., King, E.A., Paget, M., Trudinger, C.M. (2009). Australian Water Availability Project (AWAP): CSIRO Marine and Atmospheric Research Component: Final report for Phase 3. CAWCR Technical Report No. 013 p. 67.

Raupach, M.R., Briggs, P.R., Haverd, V., King, E.A., Paget, M., Trudinger, C.M. (2012). Australian Water Availability Project (AWAP). CSIRO Marine and Atmospheric Research, Canberra, Australia. <http://www.csiro.au/awap>. Accessed 10 December 2018.

Wu, Y., Chen, J. (2013). Estimating irrigation water demand using an improved method and optimizing reservoir operation for water supply and hydropower generation: a case study of the Xinfengjiang reservoir in southern China. Agricultural Water Management 116, 110-121.