Inflow forecasting using Artificial Neural Networks for reservoir operation

CHUTHAMAT CHIAMSATHIT, ADEBAYO J. ADELOYE & SOUNDHARAJAN BANKARU-SWAMY

Institute for Infrastructure and Environment, Heriot-Watt University, Edinburgh, EH14 4AS, UK
a.j.adeloye@hw.ac.uk

Abstract In this study, multi-layer perceptron (MLP) artificial neural networks have been applied to forecast one-month-ahead inflow for the Ubonratana reservoir, Thailand. To assess how well the forecast inflows have performed in the operation of the reservoir, simulations were carried out guided by the systems rule curves. As basis of comparison, four inflow situations were considered: (1) inflow known and assumed to be the historic (Type A); (2) inflow known and assumed to be the forecast (Type F); (3) inflow known and assumed to be the historic mean for month (Type M); and (4) inflow is unknown with release decision only conditioned on the starting reservoir storage (Type N). Reservoir performance was summarised in terms of reliability, resilience, vulnerability and sustainability. It was found that Type F inflow situation produced the best performance while Type N was the worst performing. This clearly demonstrates the importance of good inflow information for effective reservoir operation.

Key words Reservoir operation, Inflow forecasting, Rule curves, Ubonratana-Thailand, Sustainability

INTRODUCTION

The planning of reservoirs for various purposes including flood and drought control relies on the historic inflow data at the reservoir site. Due to natural variability and other factors (e.g. climate and land-use changes), however, the inflow situation when the reservoir is being operated will be different. It is therefore important that reservoirs are properly operated so that they continue to perform satisfactorily during changing hydro-climatology.

Reservoir operation concerns taking decisions on water release from a reservoir based on the amount of water available vis-à-vis the demand placed on the system. The available water is the sum of starting period storage and the inflow expected during the period. Consequently, effective reservoir operation relies on reliable forecast of the inflow into the reservoir. Traditional forecasting methods using hydrologic, hydraulic and time-series models are notoriously uncertain (Zhang, 1998), which is why focus has recently shifted to the use of data-driven techniques. In particular, artificial neural networks (ANN) have been widely used to forecast reservoir inflows (see e.g. Edossa and Babel, 2012; Mohamadi et al., 2005) due to their effectiveness and flexibility and have been proven to be superior to other approaches such as regression-based and time series models.

The aim of this study is to apply multi-layer perceptron (MLP)-ANN for the one-month-ahead inflow forecasting for the Ubonratana reservoir, Thailand. To investigate the effect of the forecasts on reservoir operation performance, four situations were considered for the one-month-ahead inflow: (1) inflow is known and assumed to be the historic (Type A); (2) inflow is known and assumed to be the ANN forecast (Type F); (3) inflow is known and assumed to be the historic average for the given month (Type M); and (4) inflow is not known and the release decision is conditioned only on the starting reservoir storage (Type N). Simulations of the Ubonratana reservoir were then carried out with these alternative inflow scenarios and the resulting reservoir performance was summarised in terms of reliability, resilience, vulnerability and sustainability.

In the next section, further details about the methodology will be given. This is then followed by the presentation of the case study. Next the results are presented and discussed and finally, the main conclusions are given.

METHODOLOGY

Artificial neural networks modelling

The theory and mathematical basis of ANN have been described excellently by Shamseldin (1997). Essentially, the structure of ANN comprises an input layer, an output layer and one or more hidden layers as illustrated in Fig.1. The schematic in Fig.1 has a single hidden layer which is generally sufficient to approximate any complex, non-linear function (Mulia et al., 2015). The layers contain nodes or neurons which are connected by weights. Determining optimal values for these weights and other parameters of the network is the purpose of the ANN training exercise.

For a given problem, the number of nodes in the output layer is fixed by the problem, e.g. in the current work, it is the 1-month ahead inflow forecast. The input nodes must be determined by the factors known to affect the output variable and this has been achieved through an examination of the cross-correlation matrix (see Adeloye and De-Munari, 2006). The number of neurons in the hidden layer is much more difficult to arrive at and is normally determined as part of the training by trial and error as described by Adeloye and De-Munari (2006).

Training is often improved through the use of early-stop-rule (ESR) that helps to avoid over-fitting. In ESR, the available data are divided into three parts: (i) a training set, used to determine the network weights and biases, (ii) a validation set, used to estimate the network performance and decide when the training should be stopped, and (iii) a test set, used to verify the effectiveness of the stopping criterion and to estimate the expected performance in the future.

The tested ANN architectures (in trying to arrive at the best value for the number of hidden neurons) were compared using the correlation coefficient (R) criterion, i.e.:

$$R = \frac{\sum y_{sim} y_{obs} - \frac{\sum y_{sim} \sum y_{obs}}{N}}{\sqrt{(\sum y_{sim} - \frac{(\sum y_{sim})^2}{N})(\sum y_{obs}^2 - \frac{(\sum y_{obs})^2}{N})}} \quad (1)$$

where y_{sim} and y_{obs} are respectively the simulated and observed values of the output variable and N is the number of exemplars used.

Reservoir performance simulation

Reservoir behaviour simulation employed the mass balance equation (McMahon and Adeloye, 2005):

$$S_{t+1} = S_t + Q_t - D'_t - E_t \quad (2)$$

subject to the operational policy for the reservoir, where S_t and S_{t+1} are respectively storage at the beginning and end of time t; Q_t is the inflow to the reservoir during t; E_t is the net evaporation (evaporation minus direct rainfall) in period t; $D_t^{'}$ is the total water release towards meeting the target demand of D_t during t.

As noted previously, the water available for allocation during t, WA_t, is:

$$WA_t = S_t + Q_t \quad (3)$$

and assumes that the inflow is known at the start of the month when making the release decision. In practice, however, this is not the case and assumptions about the size of the anticipated inflow must be made. If the actual inflow turns out to be exactly the same as the assumed inflow, then the end of period storage will be exactly as given by Eq. (2). If, however, there is a discrepancy, the actual end of period storage will be different from Eq. (2).

Let the actual end-of-period storage be $S_{end,t}$, the relationships between this and S_{t+1} for each of the assumed inflow knowledge assumptions become:

(1) Type A: $WA_t = S_t + Q_t$ and $S_{end,t} = S_{t+1}$

(2) Type F: $WA_t = S_t + Q'_t$ and $S_{end,t} = S_{t+1} + Q_t - Q'_t$

(3) Type M: $WA_t = S_t + \overline{Q}_t$ and $S_{end,t} = S_{t+1} + Q_t - \overline{Q}_t$

(4) Type N: $WA_t = S_t$ and $S_{end,t} = S_{t+1} + Q_t$

where Q_t is the observed (correct) inflow during time t, Q'_t is the corresponding forecast inflow, $\overline{Q}_t$ is the historic mean flow for the month of time t, and $S_{end,t}$ is the adjusted end-of-period storage .

With the available water determined, release then takes place guided by the rule curves as follows:

Case 1: For $WA_t \geq URC_m$ this is the excess operation case, i.e., $D'_t \geq D_t$

$$D'_t = S_t + Q_t - E_t - URC_m \quad (4)$$

$$Y_t = D'_t - D_t \quad (5)$$

Case 2: For $LRC_m < WA_t < URC_m$ this is the normal operation case, i.e., $D'_t \leq D_t$

$$Y_t = 0 \quad (6)$$

If $WA_t - D_t \geq LRC_m, D'_t = D_t$ (7)

If $WA_t - D_t < LRC_m, D'_t = WA_t - LRC_m$ (8)

Case 3: For $WA_t \leq LRC_m$ this is the deficit operation case, i.e., $D'_t = 0$ (No water released)

where URC_m is the upper rule curve during month *m(=1, 2.,3,…,12)* of the year*;* LRC_m is the lower rule curve during month *m;* Y_t is the excess water released during period *t*. In general, t = 12(y-1) + m for years y = 1, 2, 3…., n, where n is the number of years in the data record.

Once the simulation is complete, performance indices are then evaluated as follows (McMahon and Adeloye, 2005):

(i) *Time-based Reliability* (R_t) :

$R_t = N_s / N$, where N_s is the total number of intervals out of N that the demand was met.

(ii) Volume-based Reliability (R_v):

$$R_v = \sum_{t=1}^{N} D'_t \Big/ \sum_{t=1}^{N} D_t \, , \; \forall D'_t \leq D_t$$

(iii) Resilience:

$\varphi = 1 \Big/ \left(\frac{f_d}{f_s}\right) = \frac{f_s}{f_d}; \quad 0 < \varphi \leq 1$, where φ is resilience, f_s is number of continuous sequences of failure periods and f_d is the total duration of the failures, i.e. $f_d = N - N_s$.

(iv) Vulnerability:

$\eta = \frac{\sum_{k=1}^{f_s} (\frac{\max(sh_k)}{D_k})}{f_s}$, where max ($sh_k$) is the maximum water shortage in failure sequence k and D_k corresponding demand.

(v) Sustainability index (Sandoval-Solis *et al.*, 2011):

$$\lambda = \left(R_t \varphi (1-\eta)\right)^{1/3}$$

Where there are multiple users or sectors, each of the above indices will be evaluated for each sector and these can later be combined to determine a weighted group (or global) index. This was done for the sustainability index λ using:

$$\lambda_G = \sum_{j=1}^{M} w_j \lambda_j \quad (9)$$

where w_j is a weight, given by (Sandoval-Soils et al., 2011):

$$w_j = \frac{DS^j}{\sum_{j=1}^{M} DS^j} \quad (10)$$

and λ_G is the group sustainability; λ_j is the sustainability for users category *j*; w_j is the weighting for user *j*; *M* is the total number of users sectors and DS_j is the average annual water demand for users sector *j*.

STUDY AREA AND DATA

The Ubonratana reservoir is the largest, single multi-purpose reservoir in the upper Chi River Basin in north-eastern Thailand. The dam provides water for consumptive uses (domestic, industrial, irrigation), Pong River in-stream flow augmentation as well as flood control (EGAT, 2002). However, the water deliveries first pass through turbines for power generation (installed capacity = 25.2 MW) before being allocated to the other uses. The release is prioritised in the order of public (i.e. domestic and industrial), instream flow augmentation and irrigation. The maximum storage capacity of the reservoir is 2,431 Mm^3 at elevation of 182 m above mean sea level (mamsl). Minimum water level of the reservoir is 175 mamsl or 581.67 Mm^3 which has been prescribed for the purpose of hydropower generation.

Data collected for the study included daily reservoir inflows, evaporation, area-height-storage relationship, weekly and monthly water requirements and operating rule curves for the reservoir. The observed monthly inflow from April 1970 to March 2012 and rainfall from April 1981 to March 2012 were provided by the Electricity Generating Authority of Thailand (EGAT) and the Royal Irrigation Department (RID). The analysis, however, used the overlapping period of April 1982 to March 2012 (i.e. 360 months) for which the rainfall and runoff data were complete. Data on historical water releases to the various sectors were also provided by the RID. The gross water requirements for the analysis period were 28,952 Mm^3, i.e. average monthly of: 5.9 Mm^3 for public (municipal and industrial) demands; 113 Mm^3 for downstream requirements; and 363.7 Mm^3 for irrigation. The original rule curves were also provided by the EGAT; the improved versions of these (see Fig. 2) developed by Chiamsathit et al. (2014) were used in the current study.

RESULTS AND DISCUSSION

ANN inflow forecasts

Based on extensive testing involving the examination of the auto-correlation function (acf- Fig (3a)), partial-autocorrelation function (pcf- Fig. 3(b)) and cross-correlation function (ccf- Fig 3(c)), six input variables (i.e. current month historic mean inflow, lagged inflows (t-1, t-2, t-3), and lagged rainfall (t-1, t-2)) were used for the ANN modelling. The acf (Fig.3 (a)) shows infinite attenuation with only the first three lags of inflow being significant. Additionally, the ccf in Fig.3(c) indicates that the first two lags of the rainfall are significant. With these, the functional form of the forecast model becomes:

$$Q_t = f(Q_{t-1}, Q_{t-2}, Q_{t-3}, R_{t-1}, R_{t-2}, \overline{Q}_t) \quad (11)$$

where Q_t is the one-month ahead inflow forecast; Q_{t-1}, Q_{t-2} and Q_{t-3} are lagged inflows of one-month, two-month and three-month, respectively; R_{t-1} and R_{t-2} are lagged rainfall of one-month and two-month, respectively; and $\overline{Q}_t$ is historic mean inflow for the current month.

The ESR was used for the ANN training and for this the 360 months of data were split into three (90:5:5) for training, validation and testing, respectively. The number of hidden neurons was varied between 1 and 35 and based on the R criterion the best architecture had 33 neurons in the hidden layer. Indeed, the final model performed very well with the R exceeding 0.9 in each of the training, validation and testing. Fig.4 (a), (b) and (c) compare the predicted and observed inflow during training, validation and testing, respectively and further confirm the good performance of the forecasting model. The time series of the forecast inflows (April 1982 to March 2012, i.e. 360 months) are also compared in Fig.5 and this together with the estimated Nash–Sutcliffe efficiency (NSE) of

0.75 is further evidence of the efficacy of the forecasting model. Additionally, the fact that the NSE was higher than zero is an indication that the model has been a better predictor than the mean value of the observed time series.

Reservoir performance evaluation

The results of the performance evaluation are summarised in Table 1. For convenience, the operating policy with Type A, Type F, Type M and Type N are denoted by P-A, P-F, P-M and P-N, respectively.

As seen in Table 1, in terms of the total amount of water released, P-A, P-F and P-M were significantly better than P-N, which is not surprising given that P-N did not have any additional water from inflows. In terms of reliability (R_t and R_v), the P-F was marginally better than using P-A and significantly better than P-N; P-F was, however, inferior to P-M. A possible reason for this is that in some of the months, the historic monthly mean and forecast inflows were higher than the actual inflows, implying that more water will be released in those months with P-M and P-F than with the other two inflow situations. However, the net effect of such large releases (based on the upwardly-biased inflow forecasts) is the increased number of excursions of the end-of-period storage ($S_{end,t}$) into the region below the LRC as shown in Table 1 for both the P-F and P-M. .

The other performance indices reported in Table 1 all reveal the superiority of P-F relative to the other inflow situations. For example, the group sustainability index for P-F was the highest of all four; indeed, the same better performance of P-F was recorded across all three (public, instream and irrigation) demand sectors supplied by the reservoir. As expected, the conservative nature of P-N resulted in the least number of excursions below the LRC. This is likely to benefit the hydro-power generation potential of the reservoir albeit, as revealed by this study, at the expense of its performance in meeting the consumptive demands.

CONCLUSION

This study has developed MLP-ANN model to forecast one-month-ahead inflow for the Ubonratana reservoir in north-eastern Thailand. Extensive testing of the model showed that it was able to provide inflow forecasts with reasonable accuracy. The performance of the ANN forecasts was tested against those of three other inflow scenarios and the reservoir simulation results showed that the ANN forecasts produced superior reservoir performance. The worst performing inflow situation was when there was complete lack of knowledge about the inflow and release decision was based on the starting storage alone. All this represents an objective demonstration of good inflow forecast knowledge for effective reservoir operation.

REFERENCES

Adeloye, A.J. and De Munari, A. (2006) Artificial neural network based generalized storage-yield-reliability models using Levenberg-Marquardt algorithm, Journal of Hydrology. 362(1-4), p.215-230.

Chiamsathit, C., Adeloye, A. J. and Soundharajan, B. (2014) Assessing competing policies at Ubonratana reservoir, Thailand. Proceedings ICE (Water Management). 167(WM10). p.551 –560.

Edossa, D. C. and Mukand S., B. (2012) Forecasting Hydrological Droughts Using Artificial Neural Network Modeling Technique. translated by Pretoria. South Africa: University of Pretoria.

EGAT (2002) Improved Rule Curve. Procedure of the Ubonratana reservoir operation: Electricity Generating Authority of Thailand (EGAT) in the Ubonratana dam.

McMahon, T. A. and Adeloye, A. J. (2005) Water resources yield. Water Resources Publications, LLC. Colorado, USA.

Mohammadi, K., Eslami, H. R. and Dardashti, S. D. (2005) Comparison of Regression, ARIMA and ANN Models for Reservoir Inflow Forecasting using Snowmelt Equivalent (a Case study of Karaj). Journal of Agricultural Science and Technology. 7. p.17-30.

Mulia, E. I., Asano, T. and Tkalich, P. (2015) Retrieval of missing values in water temperature series using a data-driven model. Earth Science Informatics. p.1-12.

Sandoval-Soils, S, Mckinney, DC and Loucks, DP (2011) Sustainability index for water resources planning and management, Water Resources Planning and Management, ASCE, 137(5), 381-389

Shamseldin, A.Y. (1997) Application of Neural Network Technique to Rainfall-Runoff Modelling. Hydrol, J.. 199. p.272-294.

Zhang, G., Eddy Patuwo, B. and Y. Hu, M. (1998) Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting. 14(1). p.35-62.

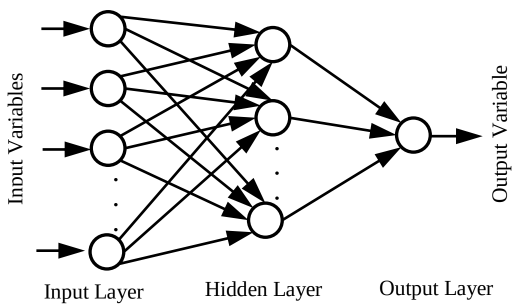


Figure 1 Schematic of artificial neural network

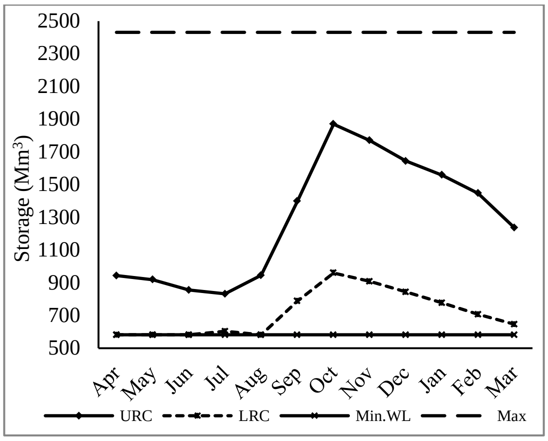
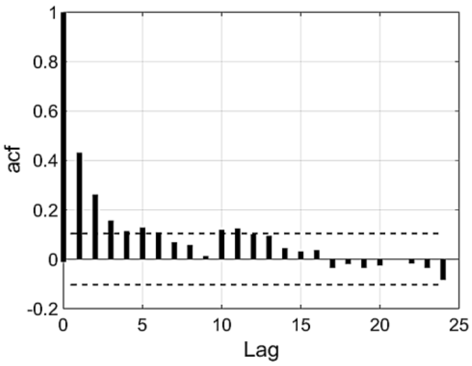
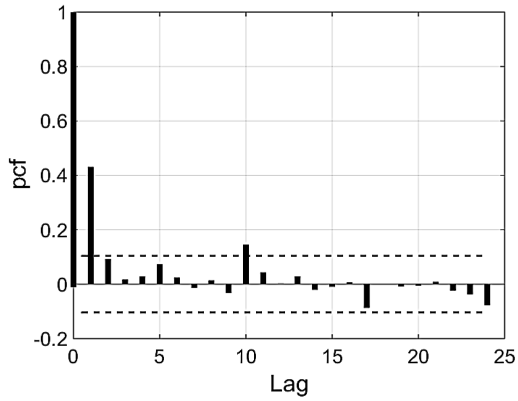


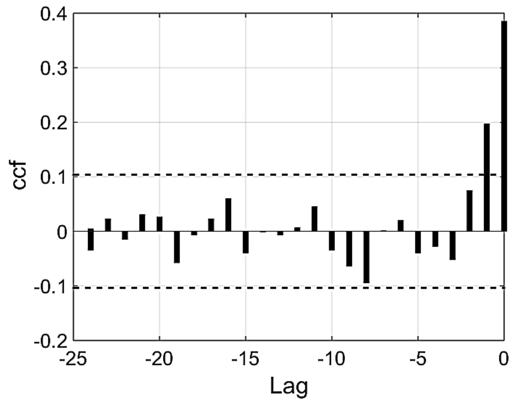
Figure 2 Rule curves for Ubonratana reservoir



(a)

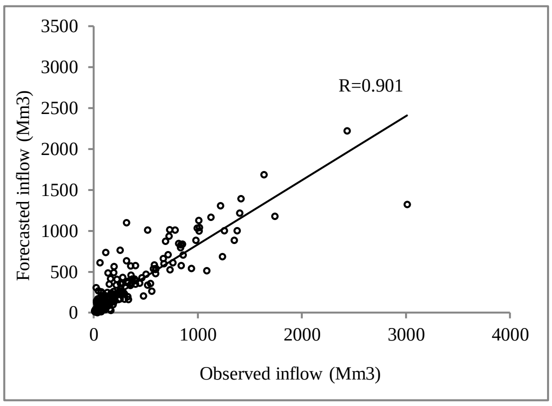


(b)

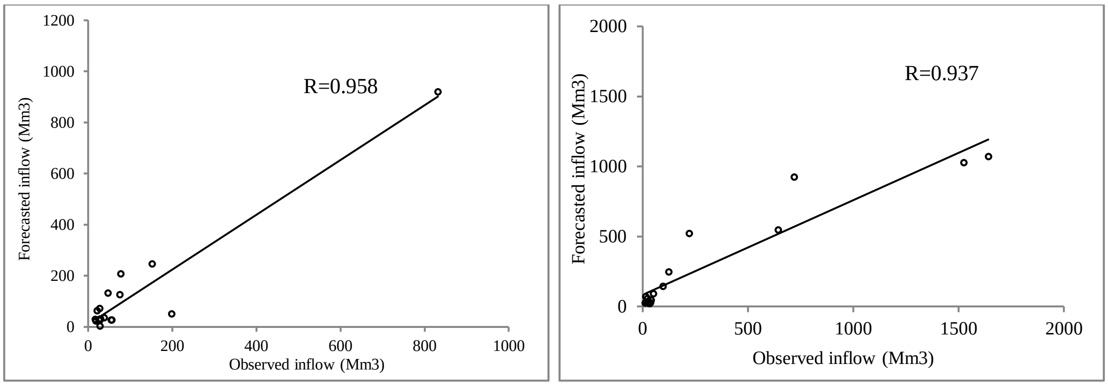


(c)

Figure 3 Inflow (a) auto-correlation, (b) partial autocorrelation functions, and (c) inflow-rainfall cross-correlation function for Ubonratana system.



(a)



(b) (c)

Figure 4: Comparing the 1-month ahead observed and forecast inflow during (a) training, (b) validation, and (c) testing

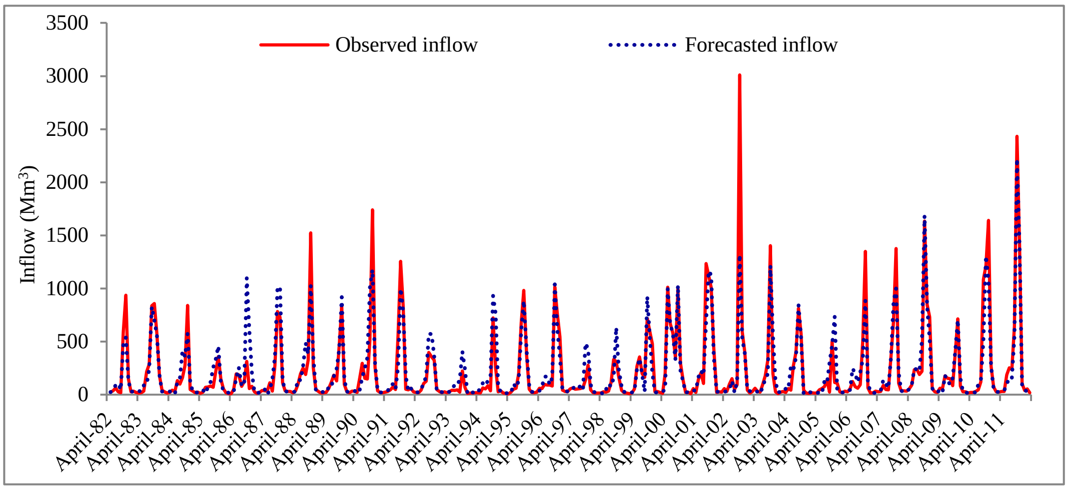


Figure 5: Time series of 1-month ahead of observed and forecast inflows for the complete data record

Table 1 Summary of evaluated reservoir performance indices for Ubonratana reservoir

Policy	Water user	Total water shortage (Mm^3)	Excursions of $S_{end,t}$ below LRC	f_d	f_s	Reliability (%) R_t	Reliability (%) R_v	φ	η	λ_{user}	λ_G
P-A	Domestic	0.0	8.0	0	0	100.00	100.00	-	0.000	1.000	0.557
	Downstream	0.5		1	1	99.72	99.99	1	0.026	0.990	
	Irrigation	309.4		15	3	95.83	98.58	0.200	0.626	0.415	
P-F	Domestic	0.0	14.0	0	0	100.00	100.00	-	0.000	1.000	0.655
	Downstream	0.0		0	0	100.00	100.00	-	0.000	1.000	
	Irrigation	244.5		10	4	97.22	98.88	0.400	0.591	0.542	
P-M	Domestic	0.0	16.0	0	0	100.00	100.00	-	0.000	1.000	0.464
	Downstream	0.0		0	0	100.00	100.00	-	0.000	1.000	
	Irrigation	166.8		6	1	98.33	99.24	0.167	0.853	0.289	
P-N	Domestic	3.2	4.0	6	5	98.33	99.09	0.833	1.000	0.000	0.543
	Downstream	132.7		10	9	97.22	98.04	0.900	0.770	0.586	
	Irrigation	1062.6		28	15	92.22	95.13	0.536	0.684	0.539	