CHAPTER 4

MODEL TESTING ON STREAMFLOWS:
I. ANNUAL MEAN FLOWS

4.1 RESULTS AND DISCUSSION

4.1.0 Preamble

The methodology as described in the previous chapter was first tested on three streamflow gauges, i.e. D1H003, D1H006 and D1H009 as depicted in Figure A1 of Appendix A and section 3.4 of Chapter 3). The regime of the time series considered was annual mean flows (1960-1989). These data are useful for water resources projects, water balance studies etc.

The theory of the EM algorithms and the entropy computations are much easier for normal distributions than for other distributions. The first assumption of the serial independence of the annual mean flow data at each gauge was tested by computing the value of the first order serial correlation of the flow series at each gauge. The computed values of the first order serial correlation were 0.10, 0.14, and 0.15 for the gauges D1H003, D1H006, and D1H009 respectively. At 95% confidence interval, the computed critical values for the annual mean flow data series were -0.211 and 0.182 respectively (Refer to Section 3.3.1, methodology). Thus, the confidence interval is (-0.211, 0.182). The computed values of the first order serial correlations, i.e. 0.20, 0.14 and 0.25 for the gauges fall into the above-mentioned interval. It was concluded that the hypothesis that the maxima data series at gauges D1H003, D1H006 and D1H009 are independent could not be rejected statistically. (In other words the null hypothesis was accepted). The rule of thumb, i.e. Elshorbagy et. al (2000b), was just mentioned to confirm the statistical decision made thereof. Thus, the annual mean flow data series at different gauges were considered to be independent (or non-autocorrelated).

According to equations 3.2, Chapter 3, the second assumption, i.e. normality, was tested for the three gauges D1H003, D1H006 and D1H009. The correlation coefficients of the data points in the Q-Q plots for these gauges (refer to Section 3.3.2, Chapter 3) were computed and found to be 0.76, 0.88, 0.80 respectively. The normality test was then carried out. The hypothesis that the data series at these gauges are normal, at a significance level of 5 %, was rejected as the correlation coefficients in the Q-Q plots were below the critical value of 0.965. To force the data series to follow the normality assumption before application of entropy concept and EM techniques, the Box Cox transformation power families (i.e. inverse, logarithmic, untransformed, cubic, square root and square equations) were applied to data. Then the values of maximum loglikelihood function (for the Box Cox transformation power families) for each gauge data series were computed. It was found that the square root transformation applied to the annual mean flow data series could achieve better (than the others) the normality assumption for these data. In other words, the square root transformation gave the highest value of the maximum loglikelihood function for the annual mean flows at gauges D1H003, D1H006 and D1H009. (The hypothesis that the data series at these gauges are normal, at significance level of 5 %, was not rejected as the correlation coefficients in the Q-Q plots were above the critical value of 0.965.). Hence, the square root transformation was applied to the raw data to follow the normality assumption at these gauges.

4.1.1 Selection of base/subject station

The entropy computations for the respective gauges were performed on the data after transformation as shown in Table 4.1 (a-c). The cross-correlation values of lag zero for the gauging site pairs, i.e. D1H003- D1H006, D1H003- D1H009 and D1H006- D1H009 were found to be 0.9023, 0.9554 and 0.8812 respectively. Tale 4.1a shows that the values of uncertainty in annual mean flows at gauges D1H003 and D1H009 are relatively higher than the uncertainty at gauge D1H006. Thus, gauges D1H003 and D1H009 could contain more information than D1H006.

Table 4.1a Marginal entropy of annual mean flows for different gauges

	D1H003	D1H006	D1H009
Marginal entropy (napiers)	3.074	1.916	2.983

Table 4.1b Informational matrix (e.g. T) of annual mean flows for different station pairs.

	D1H003	D1H006	D1H009
D1H003	1	0.8413	1.2196
D1H006	0.8413	1	0.7492
D1H009	1.2196	0.7492	1

Table 4.1c Informational matrix (e.g. DIT) of annual mean flows for different station pairs.

	D1H003	D1H006	D1H009
D1H003	1	0.2737	0.39679
D1H006	0.4391	1	0.39104
D1H009	0.408823	0.2511	1

By assuming a threshold value of 0.3 (e.g. Threshold1) for the directional information transfer index (above which, the first station within a given station-pair is potentially considered as the one transferring physically information to the second station), the following station pairs were selected: D1H003 (base)-D1H009 (subject), D1H006 (base)-D1H003 (subject); D1H006 (base)-D1H009 (subject) and D1H009 (base)-D1H003 (subject). From these selected gauge pairs, it could be said that the potential predictor (base) gauges for the predicted (subject) gauge D1H003 are gauges D1H006 and D1H009. Similarly, D1H003 and D1H006 are potential predictor (base) gauges for D1H009.

4. 1.2 Training and assessment of streamflow data infilling techniques

Since the streamflow data exhibited no gaps, different gap sizes (duration) of missing annual mean flows were created at the target (subject) station, starting from the year 1965. For the different station-pair selected in the previous section, the different values of gap duration of 6.7 %, 13.4 %, 20 % and 30 % were considered and then infilled by the selected potential base station(s).

A three-layered ANN was used as in Zealand et al. (1999). The learning rate was set to the range between 0.25 and 0.35 for quite reasonable results, unless stated otherwise. The number of nodes in the hidden layer was set to the range between 2 and 4 for reasonable results. The ANN techniques were then trained on the concurrent parts of the observed values (e.g. annual mean flows) and the parameters obtained (weights, etc.) were subsequently used to estimate the missing values. Training comprises of the annual mean flow series pertaining to input and output to the network and obtaining the inter-connection weights for the backpropagation network. Initially, the transfer function (e.g. sigmoid, hyperbolic tangent) is defined and the network is assigned values to the interconnected weights. A sigmoid function was first tried unless stated otherwise. The input data and corresponding output data were scaled within the interval (0.1, 0.9).

On the other hand, the EM algorithms were applied, thus the observed data could remain intact throughout and only the missing values were repeatedly estimated until convergence.

The different techniques were first assessed among themselves and in turn compared to the other category of techniques. Based on the results from the application of the different techniques, the proposed and existing techniques were assessed predominantly through entropy concept. The graphical and statistical criteria were also done to verify the results from entropy concept. The graphical representation is done here as in Bennis et al. (1997), where the observed parts remain intact and only the missing values are estimated. This can enhance essentially one of the steps of the EM techniques.

4.1.3 Infilling missing annual mean flows at D1H003 with base gauge D1H009

4.1.3.1 Using 6.7 % missing annual mean flows at gauge D1H003

4.1.3.1.1 Selection of ANN and EM techniques for flow infilling at D1H003

The results of performance for the different streamflow data infilling techniques (e.g. ANN and EM) at D1H003 are summarized in Table 4.2 and Figures 4.1 (a-l). D1H009 is the base station for interpolating (filling in) the missing annual mean flows at the subject gauge D1H003. Table 4.2 summarizes entropy calculations and statistics at D1H003 using ANN and EM techniques. Gauge D1H006 is not retained as base station since it could not be considered as a promising estimator of the missing annual mean flow series at D1H003, i.e. the computed value of DIT for the data series at D1H003, with base gauge D1H006, was less than the threshold (of 30 %).

The techniques that are presented in the above-mentioned table are thought to have performed well, i.e. satisfying the entropy criterion. The following has been added to elucidate the criteria of goodness of fit: the values of T or DIT as depicted in Table 4.2 should be seen as satisfying the entropy criterion. This criterion stipulates that "a given in-filling data technique is selected only if the computed value for DIT is at least 0.3, i.e. 30 %. Otherwise the technique is rejected (Refer to equation 3.20, Chapter 3). In other words, DIT which is a generalization of T represents the amount of uncertainty removed physically from the data series at the subject gauge, via a given in-filling data technique. The values for DIT (T) range from 0.387 (2.457 napiers) to 0.516 (3.28 napiers) and fall under the entropy criterion. In other words, a technique is selected when it has physically a high capability of simulating information contained in the annual mean flows at the subject gauge. The standard BP which is the most popular technique performed as well as its variants. The same is noticed for the standard EM technique and its variants. A value of 0.47 (i.e. 47 %) for the dimensionless indicator of fit of goodness, i.e. DIT (which is the generalization of T), satisfies the entropy criterion for model selection. In fact, the value of this indicator is greater than the threshold value set for the entropy criterion. In other words, 47 % of uncertainty can be removed physically from the data series at the subject gauge by applying a given in-filling data technique. Thus a technique for which the value of DIT is greater of equal to 30 % (entropy criterion for model selection), is thought to be good. Recall that the directional information transfer index (DIT) is a generalization of the transinformation T. DIT which is a dimensionless indicator of goodness of fit, expresses physically the information transfer.

The statistical criteria (Table 4.2) and the graphical plot (Figures 4.1 (a-l)) for each respective technique were just made to crosscheck the results found from entropy calculations. The values of the statistical indicators (i.e. RMSEp, RMEp and EV) range respectively from 30.20 to 65.99 Mm^3, from 0.071 to 0.142 and 0.0057 to 0.057. With regard to the magnitude of mean annual data series at gauge D1H003, the lower limit of 65.99 for RMSE could be acceptable in the in-filling data process. To support this, the statistical parameters (i.e. RMEp and EV) corresponding to the lower limit, which are dimensionless, were computed. The values of these parameters were found to be 0.142 and 0.057 respectively. Thus, these values are relatively small for the different in-filling data techniques and satisfy the entropy criterion (i.e. at least 30 % of uncertainty could be removed from the data series at the subject gauge, by applying the different techniques). Therefore, the lower limit for the dimensioned parameter RMSE was thought to have led finally to good results for entropy.

The following explains that the value of 0.142 for the RME is better than the conventional methods of filling data series at gauge D1H003. In fact, this value corresponds to the last best infilling data technique selected (i.e. QBP ANN), which showed its superiority when compared with conventional methods (i.e. linear regression) for which the value for RME was found to be 0.32. This shows that ANNs can map better (than conventional methods) the non-linearity in the annual mean flows at gauge D1H003. On the other hand, the graphical plot enabled a visual observation between the observed and the estimated hydrographs for each selected technique. The two hydrographs were found to be close for each respective technique (Figures 4.1 (a-l)).

Thus the selected techniques are thought to be good estimators of the missing annual mean flows at D1H003, for gap duration of 6.7 %.

Table 4.2 Entropy calculations and statistics at the target gauge D1H003 (6.7 % missing annual mean flows) using the base gauge D1H009

	Description	T (napiers)	DIT	RMSEp (Mm^3)	RMEp (n = 2)	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	3.101	0.4880	36.45	0.0873	0.0071	0.9870	1.000
	MBP (sigmoid hidden layer), u = 0.05	3.169	0.4990	36.199	0.083	-0.0088	0.9862	0.999
	McL1BP (sigmoid hidden layer)	3.128	0.44992	40.265	0.847	-0.027	0.988	0.995
	GenerBP (sigmoid hidden layer), lr = 0.25, s=5	3.28	0.5160	30.200	0.071	0.00573	0.9963	1.002
	BP (hyperbolic tangent hidden layer)	2.982	0.4690	39.34	0.0935	0.0232	0.9864	1.050
	QBP (hyperbolic tangent hidden layer), acc = 0.025, lr = 0.25, Weight Cond	2.457	0.3870	65.99	0.142	0.057	0.9865	1.054
	VLR (sigmoid hidden layer)	2.962	0.4660	38.22	0.0810	0.0594	0.990	0.9863
EM Techniques	Standard EM	3.044	0.4792	35.054	0.0749	0.0539	1.00	1.00
	MEM1 (u = 0.95)	3.044	0.4792	35.054	0.0749	0.0539	1.00	1.00
	MEM2, u = 0.89, lr = 0.001	3.044	0.4792	35.054	0.0749	0.0539	1.00	1.00
	MEM3	3.046	0.4794	35.024	0.0749	0.0542	1.00	1.00
	ECM1-2	3.044	0.4792	35.054	0.0749	0.0539	1.00	1.00
	ECME1-2 (k = 2), 3	3.044	0.4792	35.054	0.0749	0.0539	1.00	1.00

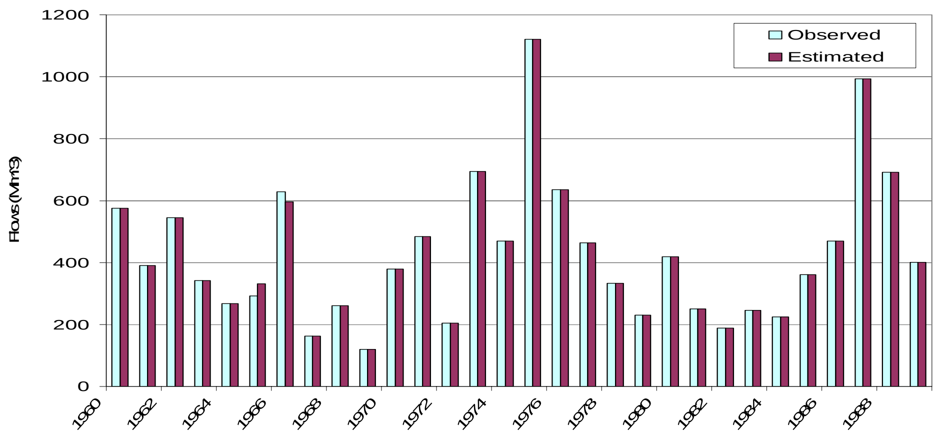


Figure 4.1a Annual mean flows at D1H003 (6.7 % missing from 1965) using standard BP with base gauge D1H009

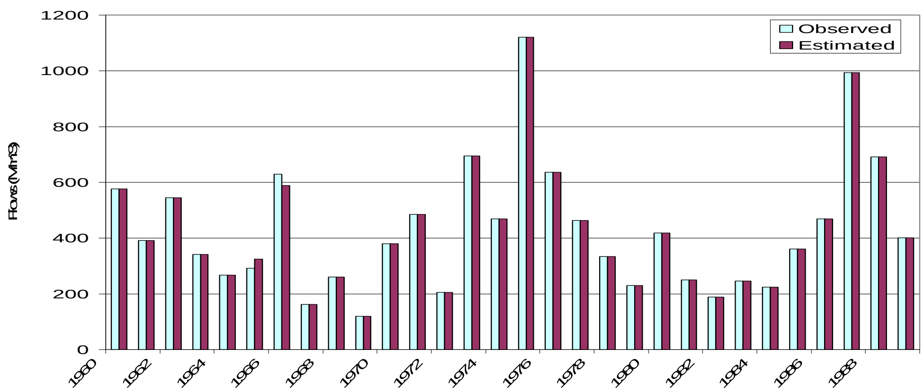


Figure 4.1b Annual mean flows at D1H003 (6.7 % missing from 1965) using MBP (u = 0.05) with base gauge D1H009

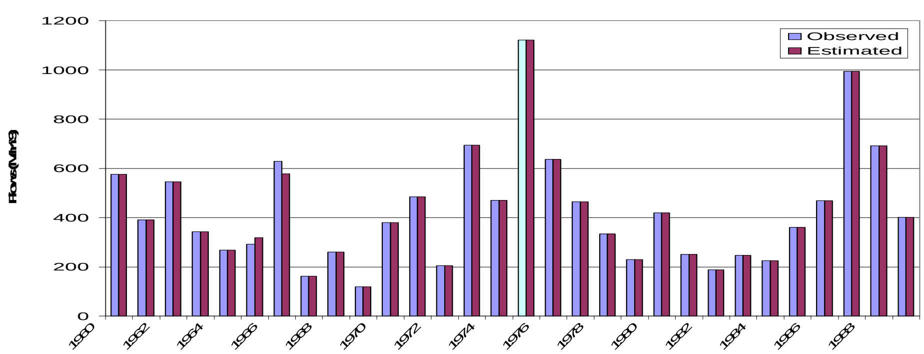


Figure 4.1c Annual mean flows at D1H003 (6.7 % missing from 1965) using McL1BP with base gauge D1H009

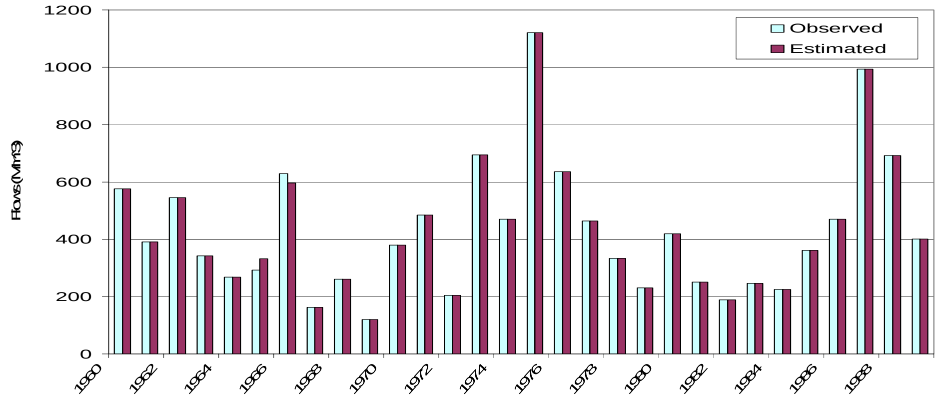


Figure 4.1d Annual mean flows at D1H003 (6.7 % missing from 1965) using GenerBP with base gauge D1H009

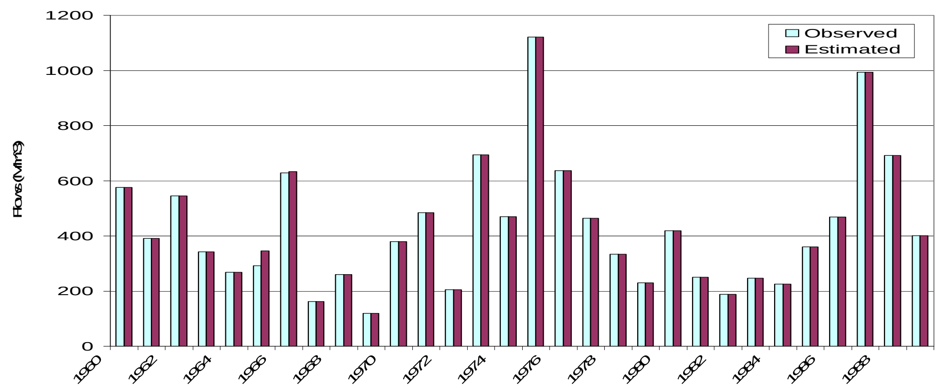


Figure 4.1e Annual mean flows at D1H003 (6.7% missing from 1965) using VLRBP with gauge base D1H009

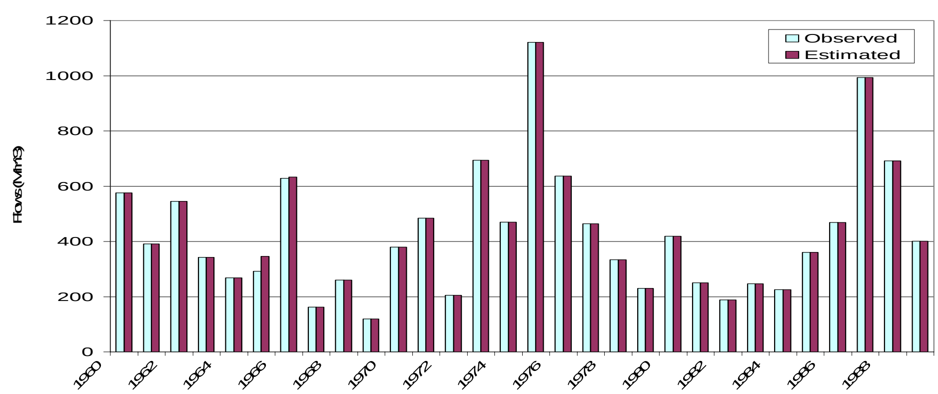


Figure 4.1f Annual mean flows at D1H003 (6.7% missing from 1965) using BP (hyperbolic tangent hidden layer) with gauge D1H009

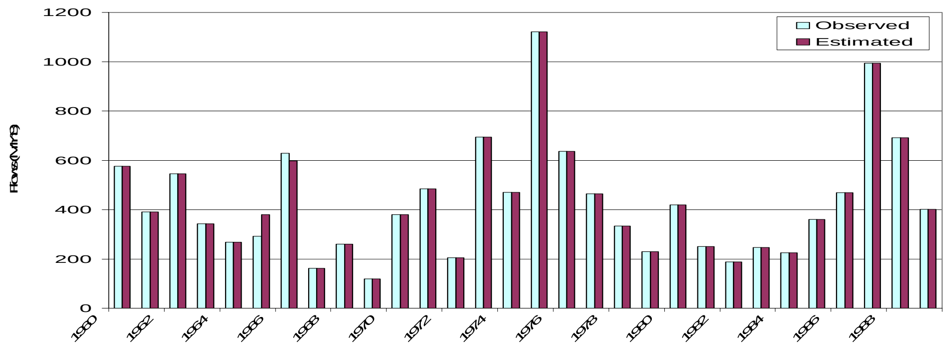


Figure 4.1g Annual mean flows at D1H003 (6.7% missing from 1965) using QBP with base gauge D1H009

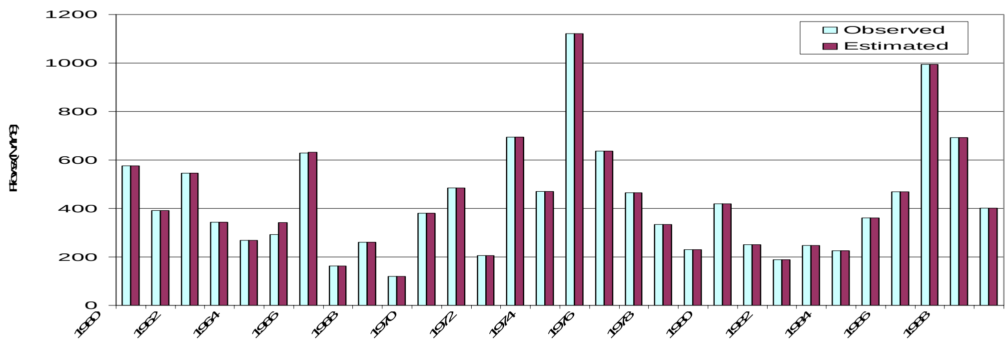


Figure 4.1h Annual mean flows at D1H003 (6.7% missing from 1965) using Standard EM with base gauge D1H009

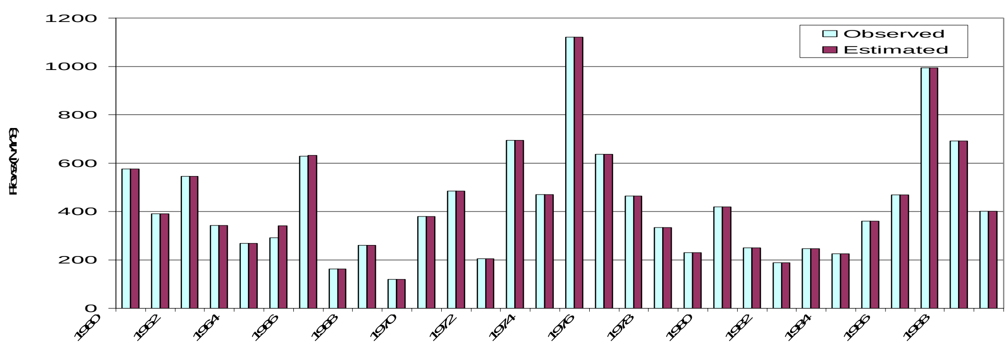


Figure 4.1i Annual mean flows at D1H003 (6.7% missing from 1965) using MEM1-2 with base gauge D1H009

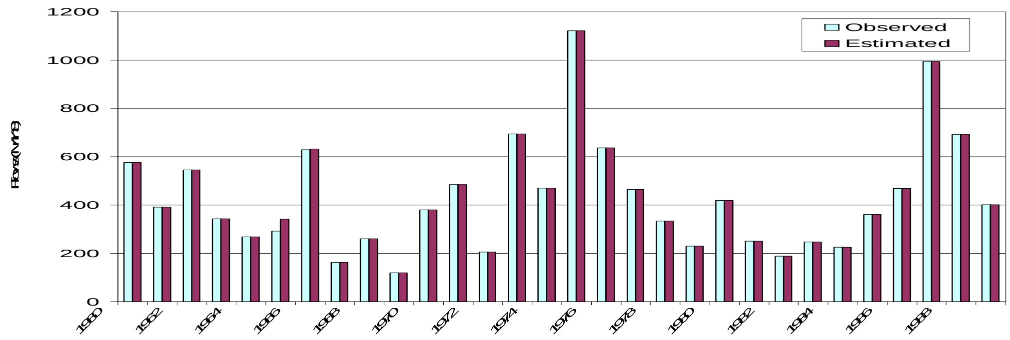


Figure 4.1j Annual mean flows at D1H003 (6.7% missing from 1965) using MEM3 with base gauge DD1H009

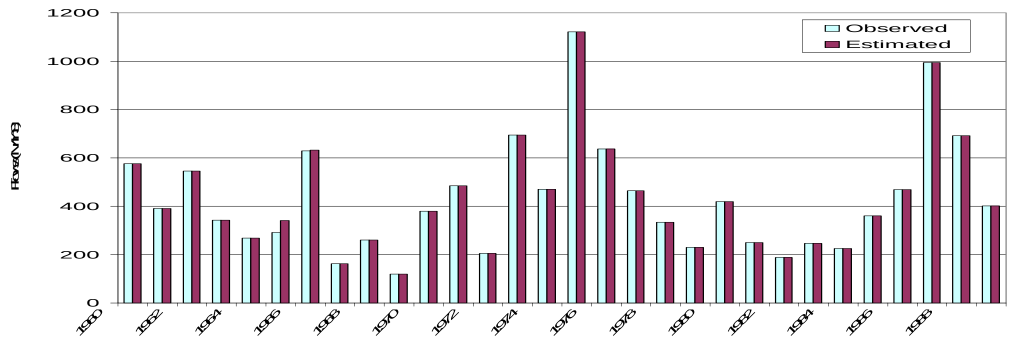


Figure 4.1k Annual mean flows at D1H003 (6.7% missing from 1965) using ECM1-2 with base gauge D1H009

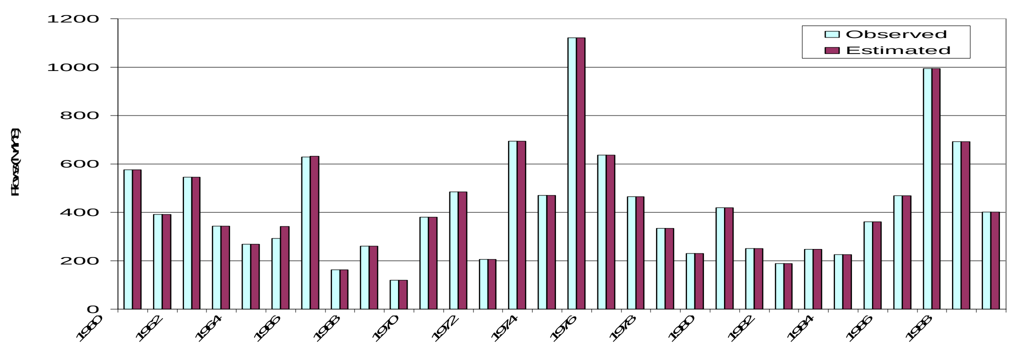


Figure 4.1l Annual mean flows at D1H003 (6.7% missing from 1965) using ECME1-2-3 using base gauge D1H009

4.1.3.1.2 Comparison of performance of ANN techniques at gauge D1H003 using base gauge D1H009

Based on the entropic values for the selected techniques, model performance assessment was then made for the estimation of missing annual mean flow values at gauge D1H003. Considering D1H009 as base gauge and for gap duration of 6.7 % at D1H003, the performance (in a decreasing order) of the different ANN techniques in terms of DIT (T) is as follows:

-GenerBP, lr = 0.25, s = 5

-MBP

-McL1BP

-Standard BP (sigmoid hidden layer)

-Standard BP (hyperbolic tangent hidden layer)

-VLRBP

-QBP (hyperbolic tangent hidden layer), acc = 0.025, lr = 0.25 and Weight cond.

From Table 4.2, the values of DIT (T) range from 0.387 (2.457 napiers) to 0.516 (3.28 napiers) and Figure 4.2 shows the values for DIT. Thus, the first best model (i.e. GenerBP) among the selected ANN techniques has the highest entropic value of 0.516 (3.28 napiers) for DIT (T). In other words, 51.6 % of uncertainty can be removed from the annual mean flow data at gauge D1H003 via the GenerBP technique. This can be also seen as the proportion of information physically transferred by the knowledge of the estimated series (through this technique) into the process to make the original series at D1H003 better defined. The GenerBP technique is thought to be relatively more capable (than other ANNs) of estimating the missing values at D1H003. It is believed that the GenerBP could make it through its generalization parameter (e.g. s), which was selected above 1.

The McL1BP is slightly better than the standard BP since it is believed that the Mac Laurin power series of order 1 is a better approximation for the sigmoid function within the scaled input data interval (0.1, 0.9) for the gap duration of 6.7 % at D1H003. Despite that, it could be noticed that 48.8 % of information could be physically inferred about gauge D1H003 via the standard BP. This technique performed better than the rest of ANNs. Thus, the standard BP is thought, in this study, to be a good estimator of missing annual mean flows at D1H003. This technique leads to some solutions in most cases as stated by Minns and Hall (1996).

The last best technique among these selected ANN techniques (i.e. QBP) has the lowest entropic value of 0.387 (2.457napiers) for DIT(T). It can be said that 38.7 % of uncertainty can be removed from the annual mean flows at gauge D1H003 via the QBP technique.

On the other hand, the values for RMSEp range from 30.20 to 65.99 Mm^3 of annual mean flows. These two values are for GenerBP and QBP respectively. The values for RMEp range from 0.071 to 0.142. The values for EV range from 0.057 to 0.0057. The GenerBP technique has the lowest values of statistical indicators while the QBP has the highest values of the same statistical indicators. These results just confirm the conclusions drawn from the entropy calculations.

Figure 4.2 summarizes the comparison of the performance of the different ANN techniques in terms of entropic values. Figure 4.3 did not show any substantial difference between any two consecutive techniques. However, the difference can be noticed between for example the first best technique and the last best one. This confirms the results from entropy calculations.

In general, these techniques appear to be good estimators of the missing annual flow data (6.7 %) at gauge D1H003. For 6.7 % missing values, the statistical parameters for D1H003 such as mean and variance did not vary as shown (in Table 4.2) by the ratio of observed statistic to the estimated statistic. In other words, the above-selected ANNs preserved the mean and the variance of the estimated series compared to the observed one.

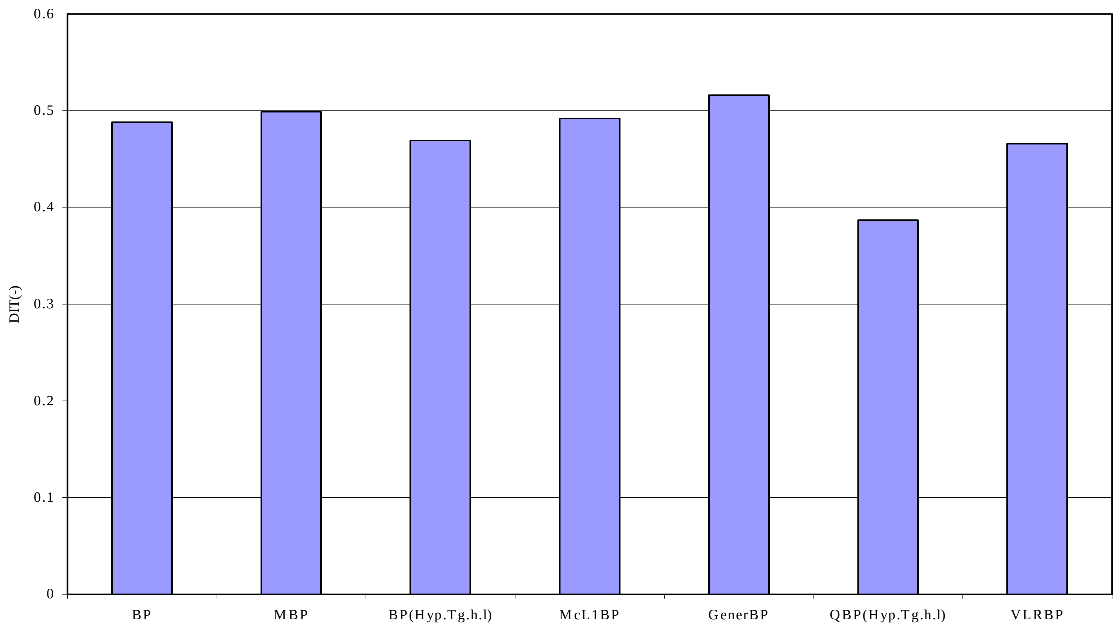


Figure 4.2 Comparison of ANNs in terms of DIT (6.7 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

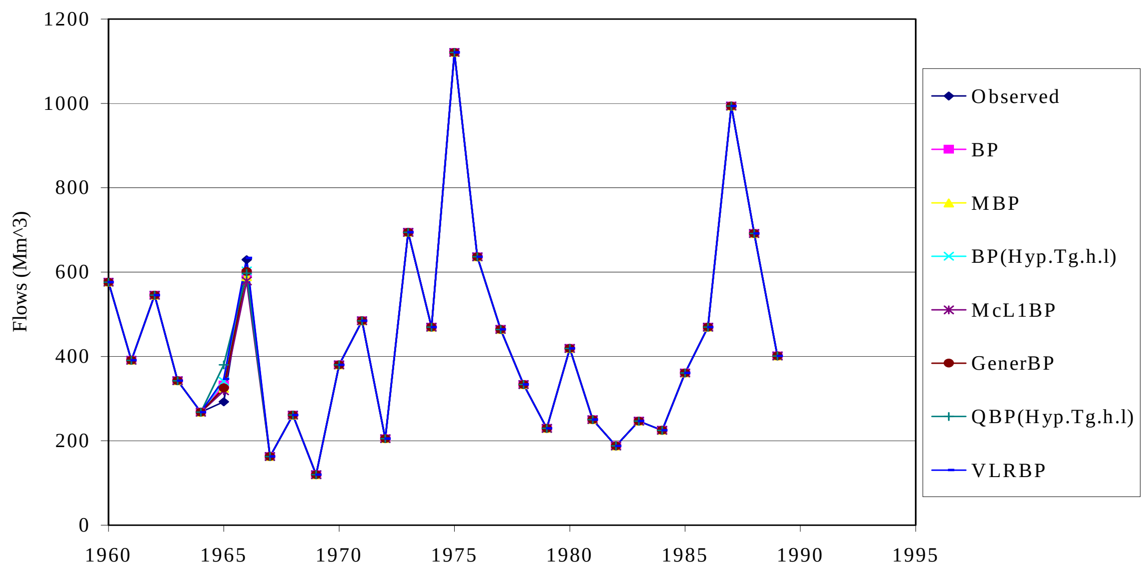


Figure 4.3 Comparison of ANN techniques in terms of hydrographs (6.7 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

4.1.3.1.3 Comparison of performance of EM techniques at gauge D1H003 using base gauge D1H009

From the results shown in Table 4.2 and Figure 4.4, it can be seen that the standard EM technique and its variants have identical entropic values although very slight (negligible) improvements were noticed only for MEM3 compared to the standard EM technique. In other words, the proportion of the uncertainty in the data series at gauge D1H003, which has been physically removed by the application of EM and MEM3 techniques, is almost the same. This shows that the standard EM leads to the same solution as its variants in some cases; especially for cases where the complete data maximum likelihood estimation of parameters is not complicated as chosen in this thesis. However, for complicated cases Meng and Rubin (1993) and Rubin and Xu (1994) recommended using other variants.

The purpose in this study was to determine the accuracy of the streamflow data infilling techniques, pre-dominantly through entropy criterion, not their computational efficiency. However, it was just noticed some differences in terms of iterations between the standard method and its different variants. It can be concluded that, from entropy calculations, both the standard EM and its variants perform likely equally in estimating the missing annual mean flow values (e.g. 6.7 %) at gauge D1H003. The momentum or the learning rate in the different versions of EM, as introduced in this study, would seem to be optional in the sense that they don't have any substantial impact on the accuracy of the estimated values.

The statistical indicators were almost identical for all the EM techniques. These results enhance the conclusions drawn from entropy calculations. The EM techniques preserved the mean and variance as shown in Table 4.2.

Figure 4.5 summarizes the comparison of the performance in terms of hydrographs for the different EM techniques. From a visual observation, all the estimated hydrographs were almost identical. Again, this is a confirmation of the results obtained from the entropy calculations.

In an overall assessment, the EM techniques lead almost to the same solution and are thought to be promising estimators of the missing annual mean flow values at D1H003 for gap duration of 6.7 % with base gauge D1H009.

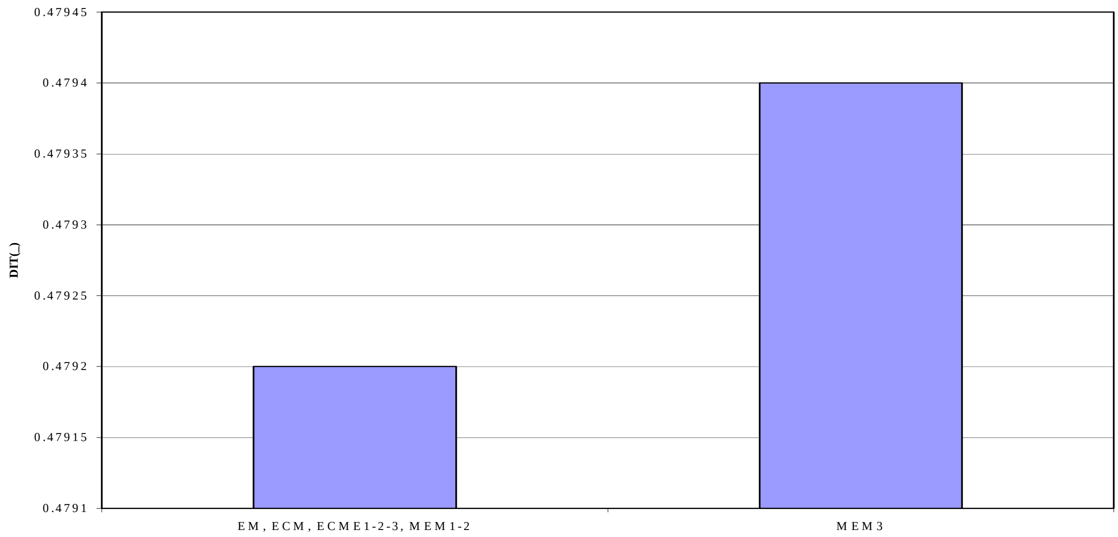


Figure 4.4 Comparison of EM techniques in terms of DIT (6.7 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

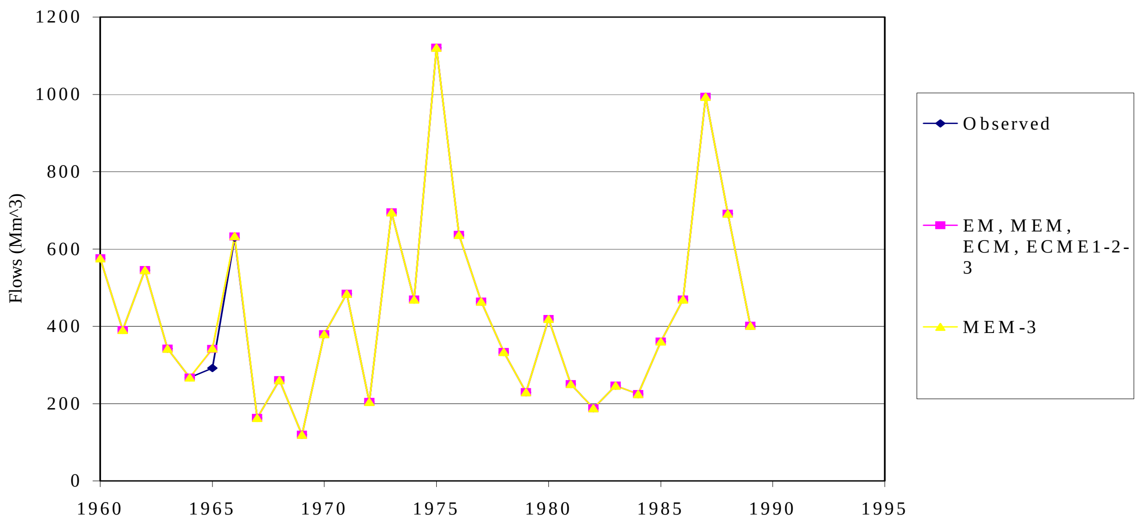


Figure 4.5 Comparison of EM techniques in terms of hydrographs (6.7 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

4.1.3.1.4 Comparison of performance of ANN and EM techniques at gauge D1H003 using base gauge D1H009

The best first four ANN techniques and the EM techniques are presented in Figure 4.6. The EM techniques come in the fourth and fifth position after the standard BP technique. Nonetheless, the EM techniques showed good estimation capabilities for the missing annual mean flows at gauge D1H003. This fact may be due to the presence of linear relationship in the annual mean flows between gauges D1H003 and D1H009 since the EM techniques used here were developed within the linear regression context. A similar observation was also made by Khalil et al. (2001), when using linear regression methods.

The differences in entropic calculations between the best four ANNs, viz GenerBP, MBP, McL1BP and standard BP and the EM techniques were small. In other words, the above-mentioned ANNs are thought to have slightly higher capabilities (than EM techniques) of mapping the non-linear characteristics in the missing values through the use of a sigmoid/hyperbolic hidden layer. On the other hand, the EM techniques were shown to perform slightly better than the rest of ANNs. The graphical comparison (hydrographs) between the EM and the ANN techniques for the gauge D1H003 is exhibited in Figure 4.7. The visual observation revealed that the hydrographs were very close.

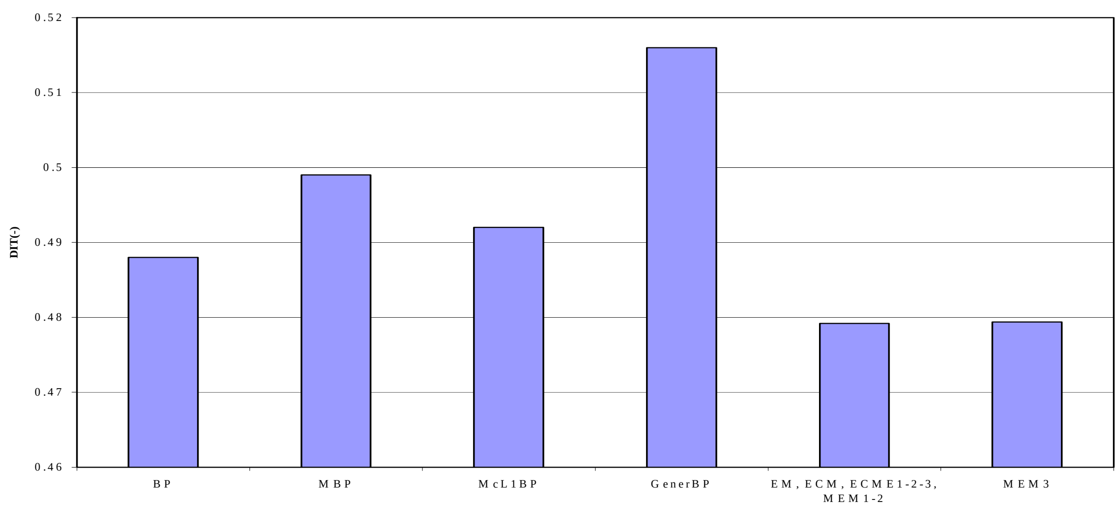


Figure 4.6 Comparison of ANN and EM techniques in terms of DIT (6.7 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

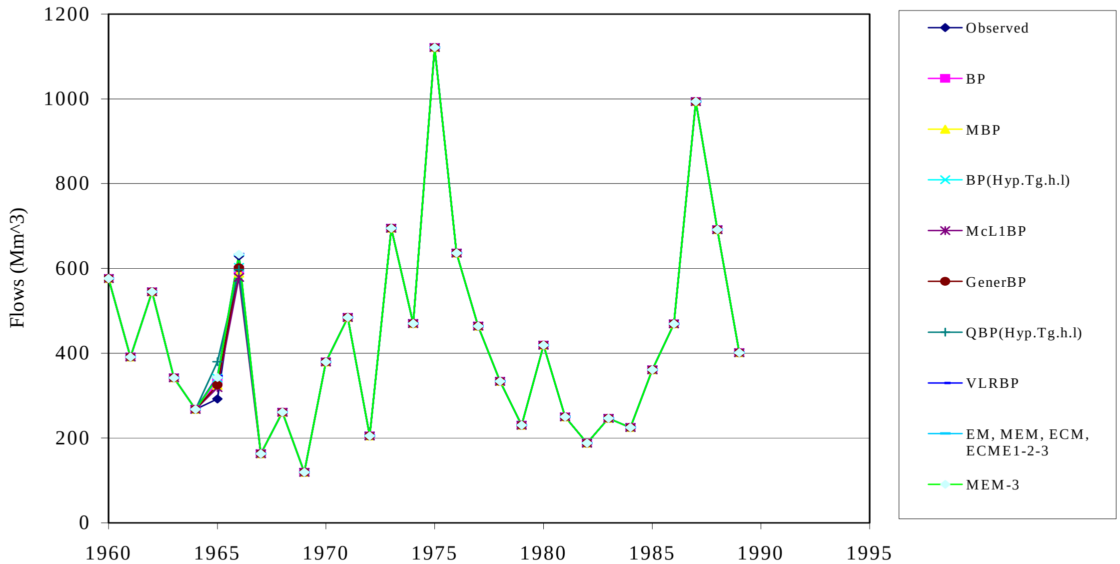


Figure 4.7 Comparison of ANN and EM techniques in terms of hydrographs (6.7 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

4.1.3.2 Using 13.4 % missing annual mean flows at gauge D1H003

4.1.3.2.1 Selection of ANN and EM techniques for flow infilling at D1H003

The results of performance for the different streamflow data infilling techniques (e.g. ANN and EM) at the predicted station D1H003 are summarized in Table 4.3 and Figures 4.8 (a-l). The techniques presented in Table 4.3 are thought to have performed well. The values for DIT (T) range from 0.363 (2.309 napiers) to 0.390 (2.478 napiers). These two limits comply with the entropy criterion. In terms of percentage, these limits are 36.3 % and 39.0 % respectively. It could be said that the standard BP performed as well as its variants. The same could be said for EM techniques. The statistical indicators viz. RMSEp, RMEp and EV range respectively from 37.97 to 49.00 Mm^3, from 0.135 to 0.1544 and from 0.007 to 0.0256 (Table 4.3). The values for RMEp and EV, which are dimensionless, are relatively small (e.g. approaching zero). The two hydrographs (i.e. observed and estimated) were found to be close for each respective technique (Figure 4.8 (a-l)). Both statistical indicators and graphical plots confirmed the results from entropy calculations.

Table 4.3 Entropy calculations and statistics at the subject gauge D1H003 (13.4 % missing annual mean flows) using base gauge D1H009

	Description	T (napiers)	DIT	RMSEp (Mm^3)	RMEp n = 4	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	2.460	0.387	39.942	0.1446	0.029	0.9902	1.002
	BP (sigmoid hidden layer), u = 0.05	2.465	0.388	39.935	0.144	-0.023	1.002	0.998
	McL1BP (sigmoid hidden layer), momentum	2.425	0.382	42.7218	0.142	-0.003	1.011	0.996
	GenerBP (sigmoid hidden layer), s = 2	2.477	0.3899	39.32	0.1426	0.0136	1.009	0.995
	BP(hyperbolic tangent hidden layer)	2.452	0.3860	40.040	0.1438	0.0138	1.020	1.000
	QBP(hyperbolic tangent hidden layer), acc = 0.025, lr = 0.25, Weight Cond	2.478	0.3907	37.97	0.135	0.0256	1.018	0.999
	VLR (sigmoid hidden layer)	2.309	0.3630	49.00	0.1544	0.0079	0.997	0.993
EM Techniques	Standard EM	2.473	0.3890	38.138	0.1358	0.0254	1.016	0.998
	MEM1 (u = 0.95)	2.473	0.3890	38.138	0.1358	0.0254	1.016	0.998
	MEM2, u = 0.95, lr = 0.001	2.419	0.3808	41.438	0.143	0.05	1.016	0.998
	MEM3	2.473	0.3890	38.138	0.1358	0.0254	1.016	0.998
	ECM1-2	2.473	0.3890	38.138	0.1358	0.0254	1.016	0.998
	ECME1-2-3(u = 0.90)	2.473	0.3890	38.138	0.1358	0.0254	1.016	0.998

Figure 4.8a Annual mean flows at D1H003 (13.4 % missing from 1965) using standard BP with base gauge D1H009

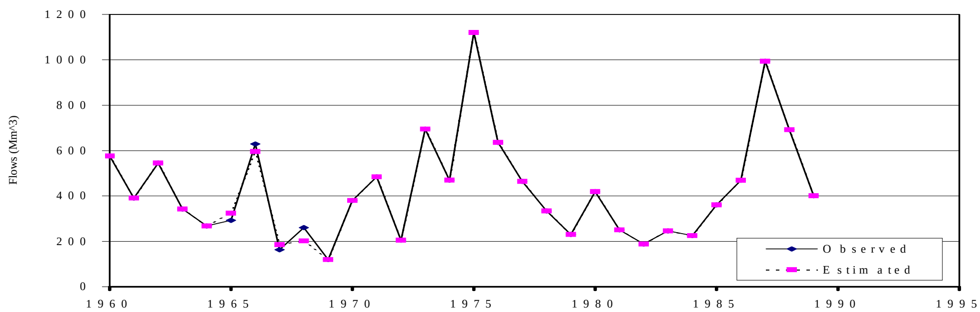


Figure 4.8b Annual mean flows at D1H003 (13.4 % missing from 1965) using MBP with base gauge D1H009

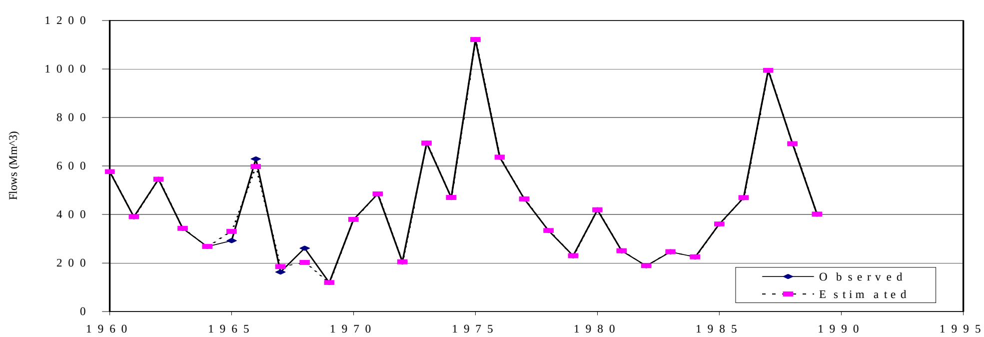


Figure 4.8c Annual mean flows at D1H003 (13.4% missing from 1965) using GenerBP (s = 2) with base gauge D1H009

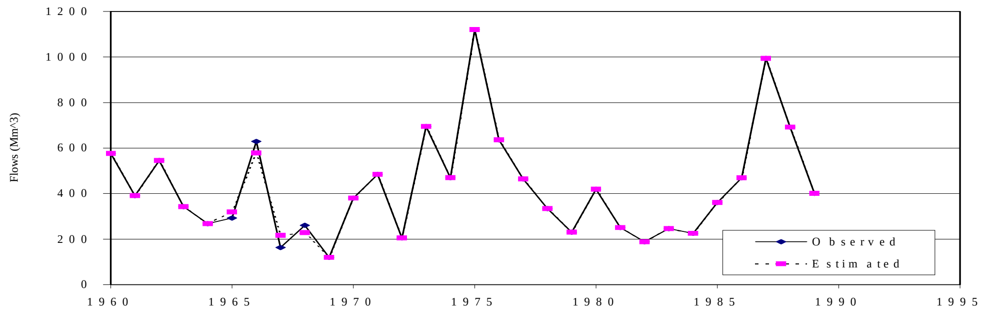


Figure 4.8d Annual mean flows at D1H003 (13.4% missing from 1965) using McL1BP with base gauge D1H009

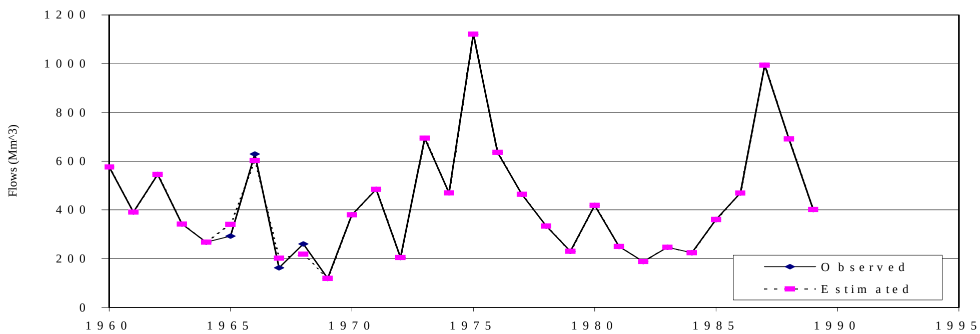


Figure 4.8e Annual mean flows at D1H003 (13.4% missing from 1965) using BP (hyperbolic tangent hidden layer) with base gauge D1H009

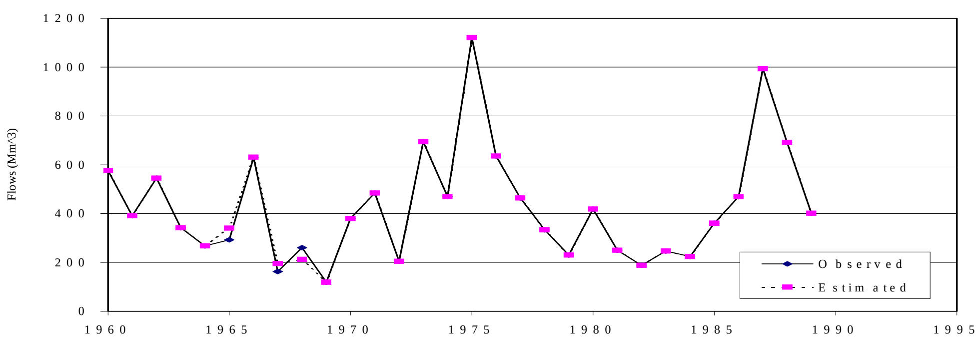


Figure 4.8f Annual mean flows at D1H003 (13.4% missing from 1965) using QBP with base gauge D1H009

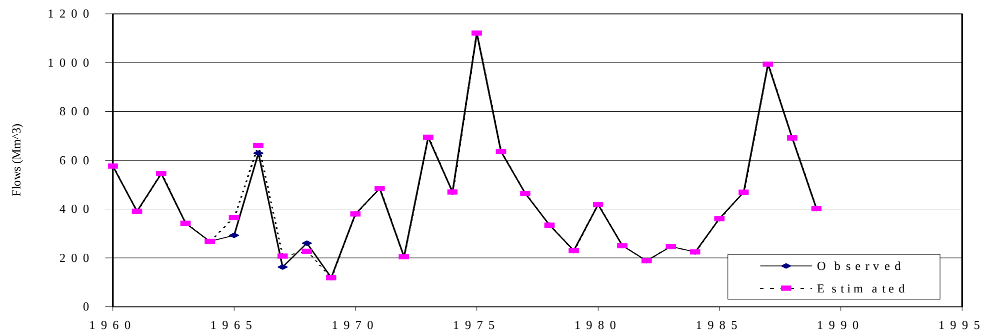


Figure 4.8g Annual mean flows at D1H003 (13.4 % missing from 1965) using VLRBP with base gauge D1H009

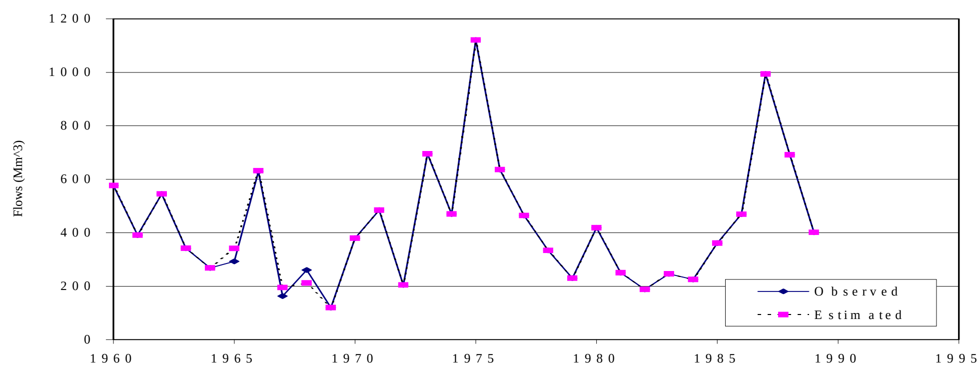


Figure 4.8h Annual mean flows at D1H003 (13.4 % missing from 1965) using standard EM with gauge base D1H009

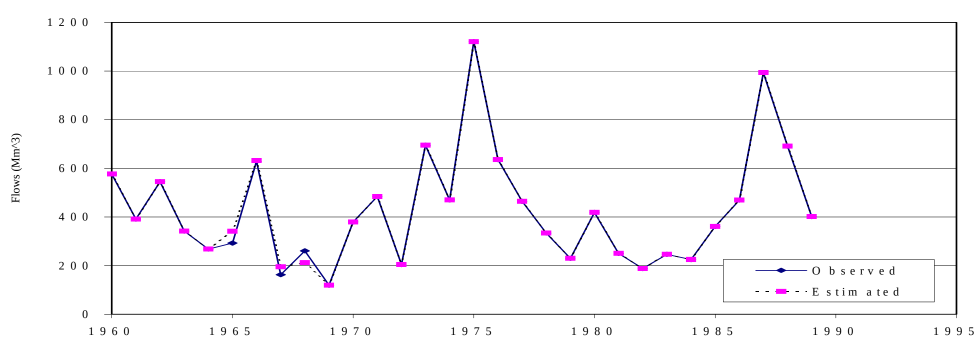


Figure 4.8i Annual mean flows at D1H003 (13.4 % missing from 1965) using MEM1-3 with base gauge D1H009

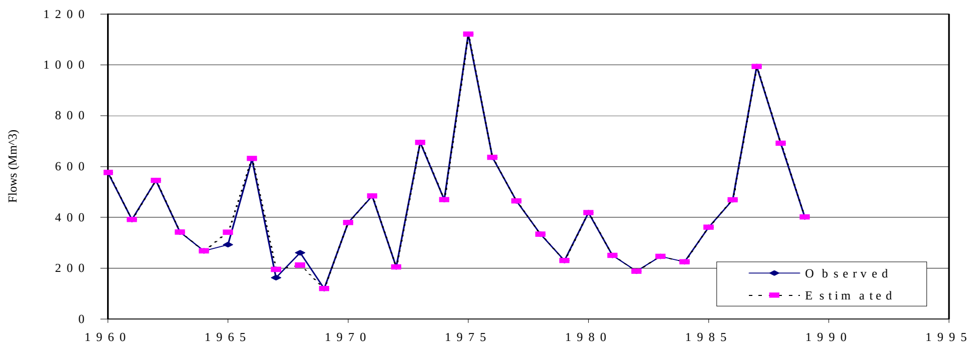


Figure 4.8j Annual mean flows at D1H003 (13.4 % missing from 1965) using MEM2 with base gauge D1H009

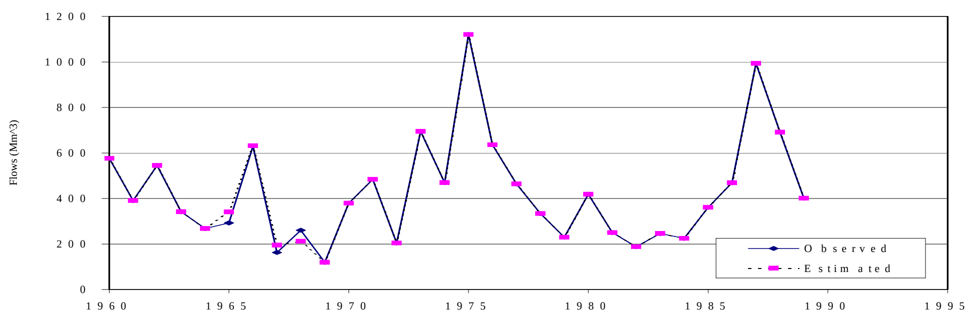


Figure 4.8k Annual mean flows at D1H003 (13.4 % missing from 1965) using ECM1-2 with base gauge D1H009

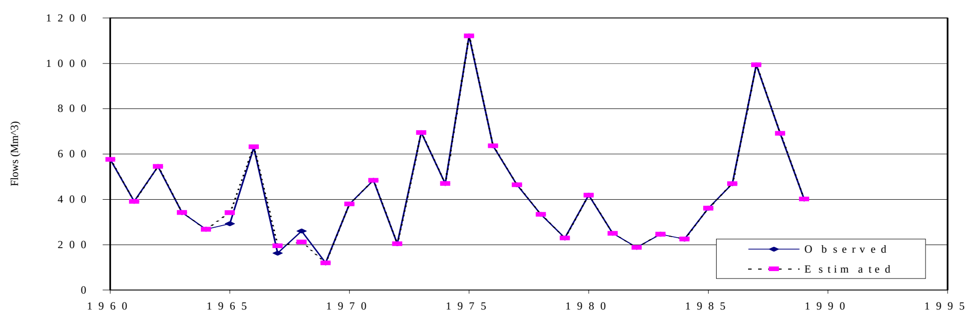


Figure 4.8l Annual mean flows at D1H003 (13.4 % missing from 1965) using ECME1-2-3 with base gauge D1H009

4.1.3.2.2 Comparison of performance of ANN techniques at gauge D1H003 using base gauge D1H009

From the entropy calculations (Table 4.3), technique performance assessment was made for the estimation of missing annual mean flows at gauge D1H003. Considering D1H009 as base gauge and for gap duration of 6.7 % at D1H003, the performance (in a decreasing order) of the different ANN techniques in terms of DIT (T) is as follows:

-QBP (hyperbolic tangent hidden layer), acc = 0.025, lr = 0.25 and Weight cond.

-GenerBP, s = 2

-MBP, u = 0.05

-Standard BP

-Standard BP (hyperbolic tangent hidden layer)

-McL1BP

-VLRBP

From Table 4.3, the values of DIT (T) range from 0.363 (2.309 napiers) to 0.390 (2.477 napiers) and Figure 4.9 shows the values for DIT. Thus, the first best technique (i.e. QBP) among the selected ANN techniques have the highest entropic value of 0.390 (2.478 napiers) for DIT (T). In other words, 39.0 % of uncertainty can be removed from the annual mean flows at gauge D1H003 via the QBP technique. It can be noticed that 38.7 % of information can be physically inferred about gauge D1H003 via the standard BP. This technique is thought, in this study, to be also a good estimator of missing annual mean flow data at the subject gauge D1H003. The last best technique among ANNs (i.e. VLR) has the lowest entropic value of 0.363 (2.457 napiers) for DIT (T). Hence, it can be said that 36.3 % of uncertainty can be removed from the annual mean flow data at gauge D1H003 via VLRBP.

On the other hand, the values for RMSEp range from 37.32 to 49.00 Mm^3. The values for RMEp range from 0.135 to 0.1544. The values of EV range from -0.003 to 0.0256. The QBP technique has the lowest values of RMSEp and RMEp, except for EV while the VLR has the highest value for RMSEp and RMEp, except the value for EV. Nonetheless the results obtained from RMSEp and RMEp calculations confirm the conclusions drawn from the entropy calculations.

Figure 4.10 shows the hydrographs for the different techniques. These plots did not show any substantial difference between two consecutive techniques. However, a small difference can be noticed, for example, between the first best technique (GenerBP or QBP) and the last one (VLR). This confirms the results from Figures 4.9.

In general, the techniques presented above appear to be good estimators of the missing annual mean flow values (13.4 %) at gauge D1H003. For this gap duration, the statistical parameters for D1H003 such as mean and variance did not vary considerably as shown (in Table 4.3) by the ratio of observed statistic to the estimated statistic.

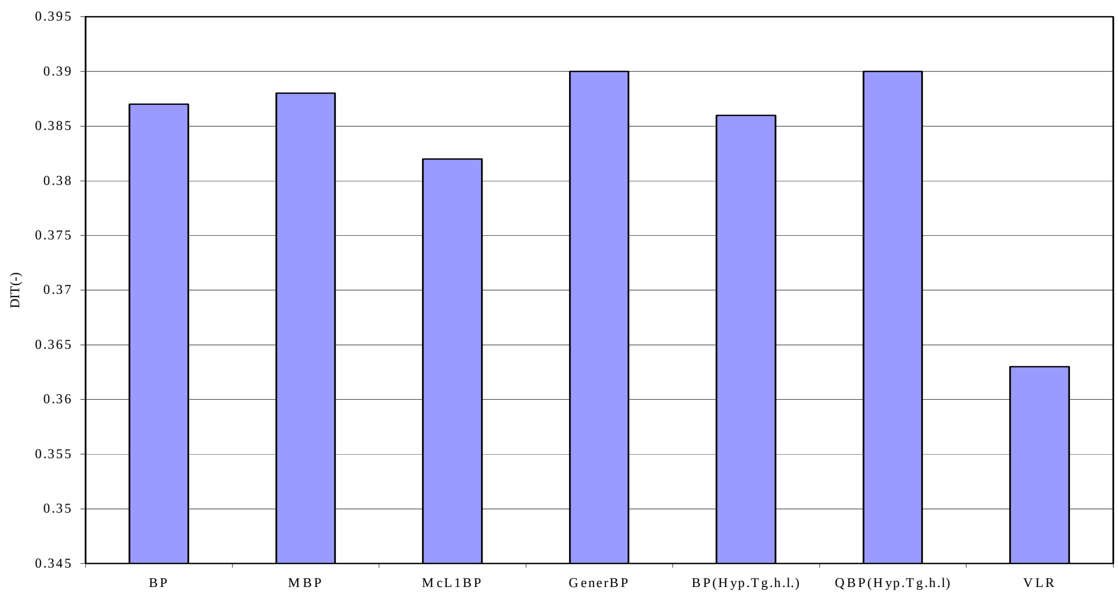


Figure 4.9 Comparison of ANNs in terms of DIT (13.4 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

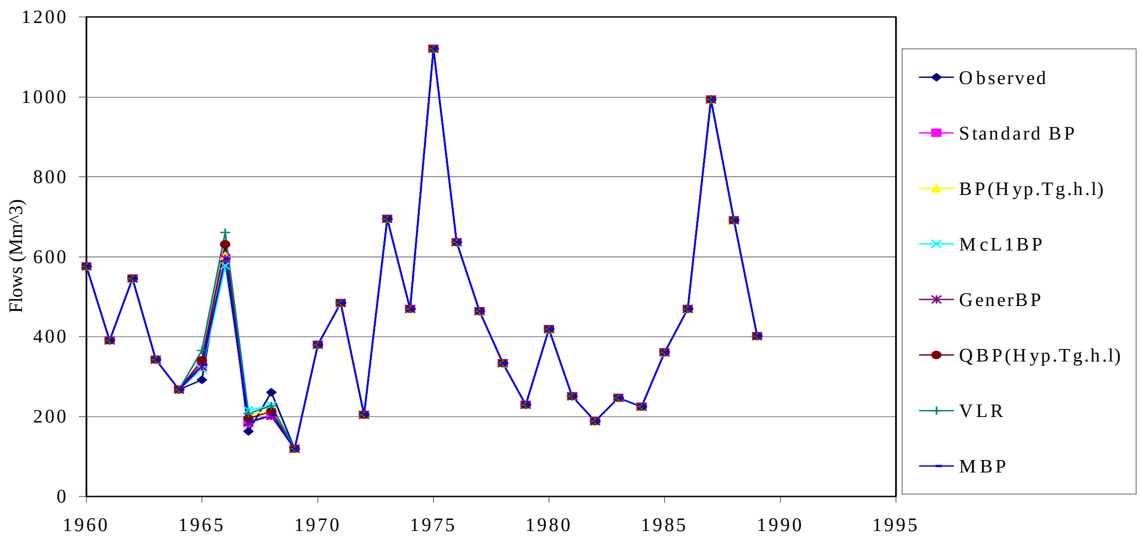


Figure 4.10 Comparison of ANNs in terms of hydrographs (13.4 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

4.1.3.2.3 Comparison of performance of EM techniques at gauge D1H003 using base gauge D1H009

From the results shown in Figure 4.11, the standard EM technique and its variants have almost identical entropic values. This shows that the standard EM leads to the same solution as its variants in some cases. The reason of this has been given so far.

From entropy calculations, it can be concluded that both the standard EM and its variants perform likely equally in estimating the missing annual mean flows (e.g. 13.4 %) at gauge D1H003.

The statistical indicators were almost identical for all the EM techniques. The EM techniques preserved the mean and variance as shown in Table 5.3.

Figure 4.12 summarizes the comparison of the performance in terms of hydrographs for the different EM techniques. From a visual observation, all the estimated hydrographs were almost identical. Again, this is a confirmation of the results obtained from the entropy calculations.

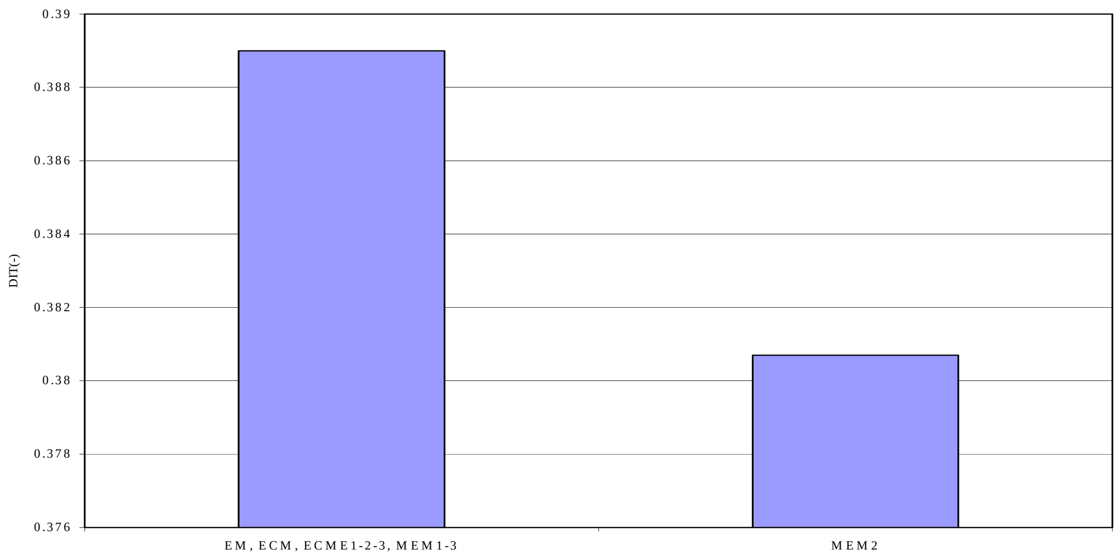


Figure 4.11 Comparison of EM in terms of DIT (13.4 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

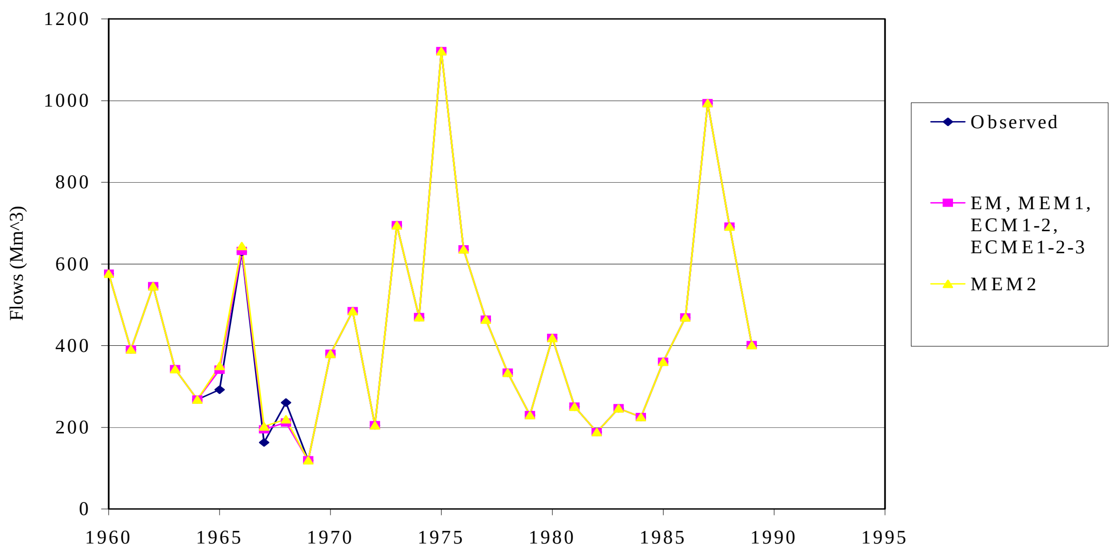


Figure 4.12 Comparison of EM in terms of hydrographs (13.4 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

4.1.3.2.4 Comparison of performance of ANN and EM techniques at gauge D1H003 using base gauge D1H009

Results for the comparison between ANNs and EM techniques are depicted in Figure 4.13 and Figure 4.14. The EM techniques come in the third position after QBP and GenerBP techniques in terms of performance (through entropy criterion). Nonetheless, the EM techniques show good estimation capabilities of the missing annual mean flows at D1H009. This may be due to the presence of linear relationship in the data series between gauges D1H003 and D1H009, since the EM techniques used here were developed within the linear regression context.

The first two ANN techniques, viz GenerBP and QBP were found to be marginally better than the EM techniques. However, the difference in entropic values between these ANNs and the EM techniques were very small. In other words, the above-mentioned ANNs are thought to have slightly higher capabilities of mapping the non-linear characteristics in the missing annual mean flows through the use of a sigmoid/hyperbolic hidden layer. On the other hand, the EM techniques were shown to perform slightly better than the rest of ANNs.

The graphical comparison (hydrographs) between the EM techniques and the first two ANN techniques for the gauge D1H003 is exhibited in Figure 4.14. The visual observation reveals that the hydrographs were very close.

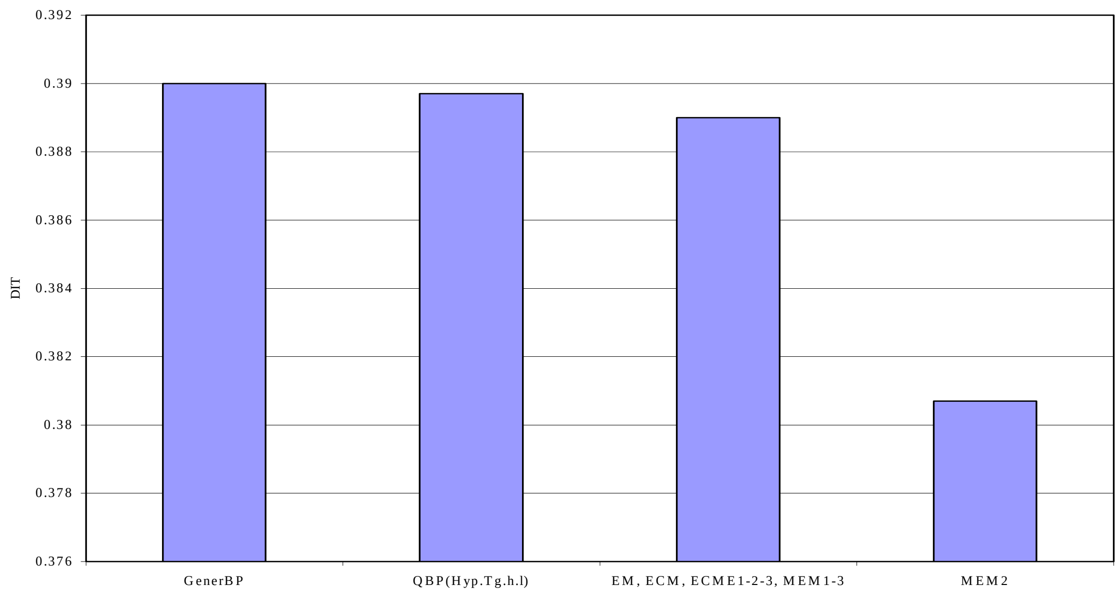


Figure 4.13 Comparison of ANN and EM in terms of DIT (13.4 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

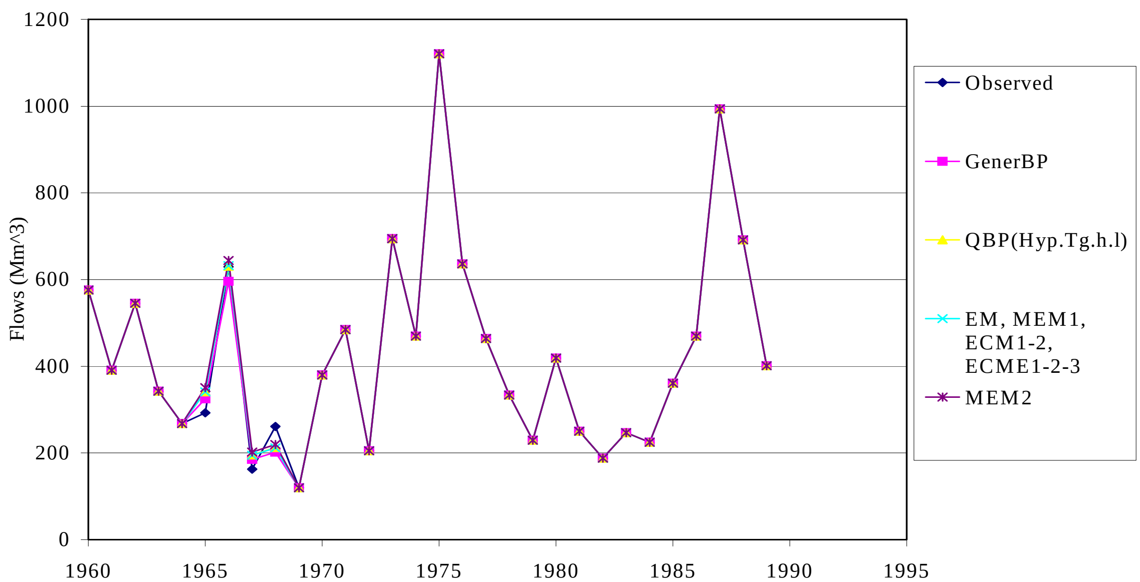


Figure 4.14 Comparison between ANNs and EM techniques in terms of hydrographs (13.4 % missing annual mean flows from 1965 at D1H003) with base gauge D1H009

4.1.3.3 Using 20 % of missing annual mean flows at gauge D1H003

4.1.3.3.1 Selection of ANN and EM techniques for flow infilling at D1H003

The results of performance for the different streamflow data infilling techniques (i.e. ANN and EM) at the subject gauge D1H003 are summarized in Table 4.4 and Figures 4.15 (a-d). Table 4.4 summarizes the entropy calculations and the statistics at D1H003.

The techniques presented in Table 4.4 are thought to have performed well. The values for DIT (T) range from 0.3006 (1.91 napiers) to 0.3479 (1.974 napiers). These two limits fall under the entropy criterion. It could be said that the standard BP technique performed as well as its variants did. The same could be also said for the standard EM and its variants.

The values of statistical indicators, e.g. RMSEp, RMEp and EV, range respectively from 41.62 to 67.87 Mm^3, from 0.200 to 0.250 and from 0.042 to 0.058. The values for RMEp, EV, which are dimensionless, are relatively small (Table 4.4). This could eventually confirm the results obtained from entropy criterion. The two hydrographs (i.e. observed and estimated) were found to be close for each respective technique (Figures 4.15 (a-d)). Thus, the above techniques are thought to be good estimators of the missing annual mean flows at D1H003 for gap duration of 6.7 %. This also confirms the results from entropy calculations.

Table 4.4 Entropy calculations and statistics at the target gauge D1H003 (20 % missing annual mean flows) using the base gauge D1H009

	Description	T (napiers)	DIT	RMSEp (Mm^3)	RMEp (n = 6)	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	1.927	0.3033	51.87	0.250	0.056	1.021	1.005
	McL1BP (sigmoid hidden layer)	1.860	0.2927	55.40	0.262	0.058	1.0215	1.009
	GenerBP (sigmoid hidden layer), s = 5	1.914	0.3006	52.81	0.252	0.048	1.019	1.001
	QBP (sigmoid hidden layer), acc = 0.025, lr = 0.85, Weight Cond	1.974	0.3107	48.91	0.200	0.042	1.000	0.998
EM Techniques	Standard EM	2.21	0.3479	41.62	0.0232	0.992	1.052	0.999
	MEM1 (u = 0.5), 2, 3 (u = 0.89, lr = 0.001)	2.21	0.3479	41.62	0.0232	0.039	1.052	0.999
	ECM1-2	2.21	0.3479	41.62	0.0232	0.039	1.052	0.999
	ECME1,2 (k = 2), 3	2.21	0.3479	41.62	0.0232	0.039	1.052	0.999

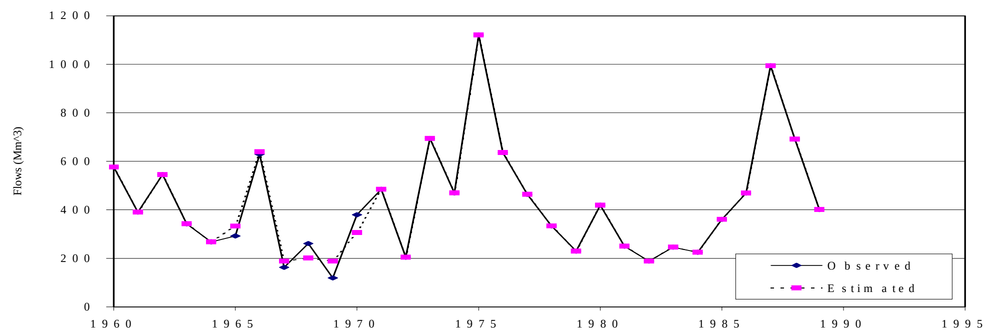


Figure 4.15a Annual mean flows at D1H003 (20 % missing from 1965) using standard BP with base gauge D1H009

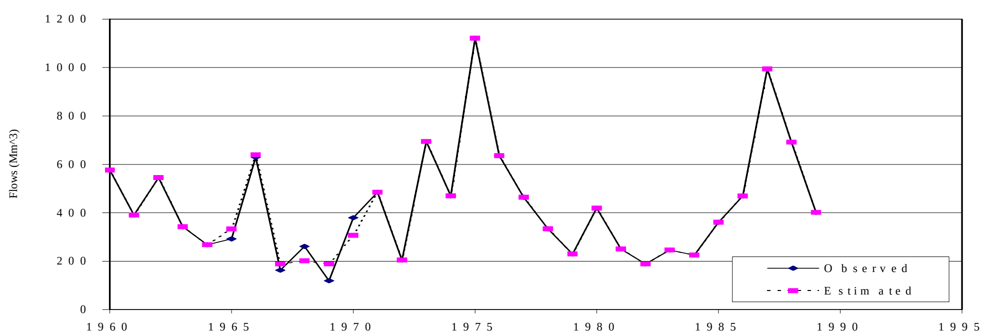


Figure 4.15b Annual mean flows at D1H003 (13.4 % missing from 1965) using QBP with base gauge D1H009

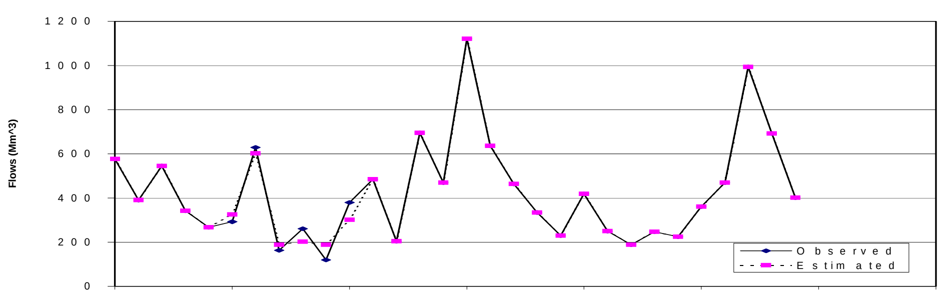


Figure 4.15c Annual mean flows at D1H003 (20 % missing from 1965) using GenerBP with base gauge

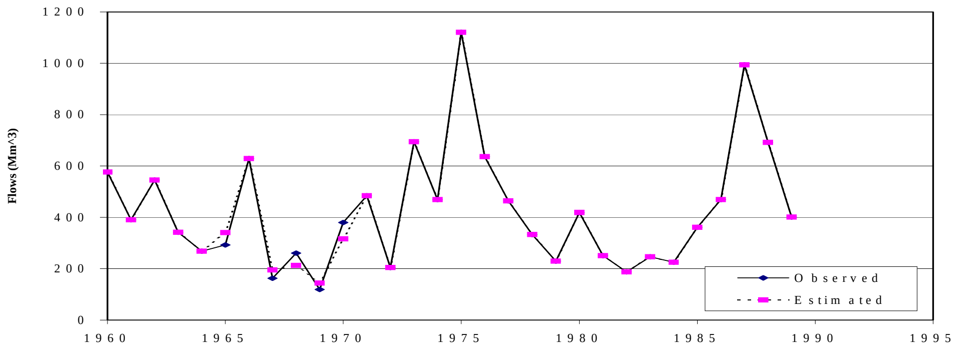


Figure 4.15d Annual mean flows at D1H003 (20 % missing from 1965) using EM, MEM1-2-3, ECM1-2, ECME1-2-3 with base gauge D1H009

4.1.3.3.2 Comparison of performance among ANN techniques at gauge D1H003 using base gauge D1H009

With D1H009 as base station and for a gap duration of 20 % at D1H003, the performance in a decreasing order of the different ANN techniques in terms of DIT (T) values was then possible and is as follows (Table 4.4): QPBP, BP, Generalized BP.

From Table 4.4, the values of DIT (T) range from 0.3007 (1.91 napiers) to 0.3107 (1.974 napiers) and Figure 4.16 shows the values for DIT. Thus the first best model (i.e. QBP) among the selected ANN techniques has the highest entropic value of 0.3107 (1.974 napiers) for DIT (T). In other words 31.07 % of uncertainty can be removed from the observed data at gauge D1H003 via the QBP technique. The QBP technique is thought to be relatively more capable of estimating the missing values at D1H003 than other ANNs.

On the other hand the values for RMSEp range from 48.91 to 52.819 Mm^3 respectively for the ANNs. The values of RMEp range from 0.200 to 0.262. The values of EV range from 0.0392 to 0.0583. The QBP technique has the lowest values of statistical indicators while the GenerBP has the highest value for the same statistical parameters. The results obtained from the statistical indicators confirm the conclusions drawn from the entropy calculations. Figure 4.17 summarizes the comparison of the performance of the different ANN techniques in terms of hydrographs. Figure 4.17 shows that the results are very close. In general these techniques appear to be good estimators of the missing values (20

%) at gauging station D1H003. For 20 % missing values, the statistical parameters for D1H003 such as mean and variance did not vary considerably as shown (in Table 4.4) by the ratio of observed statistic to the estimated statistic.

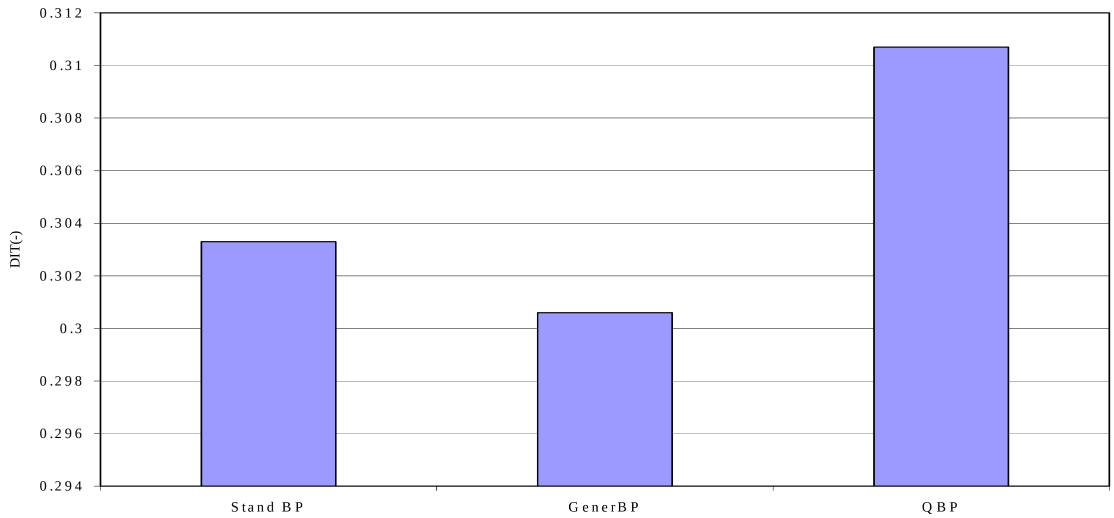


Figure 4.16 Comparison of ANNs techniques in terms of DIT (20 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

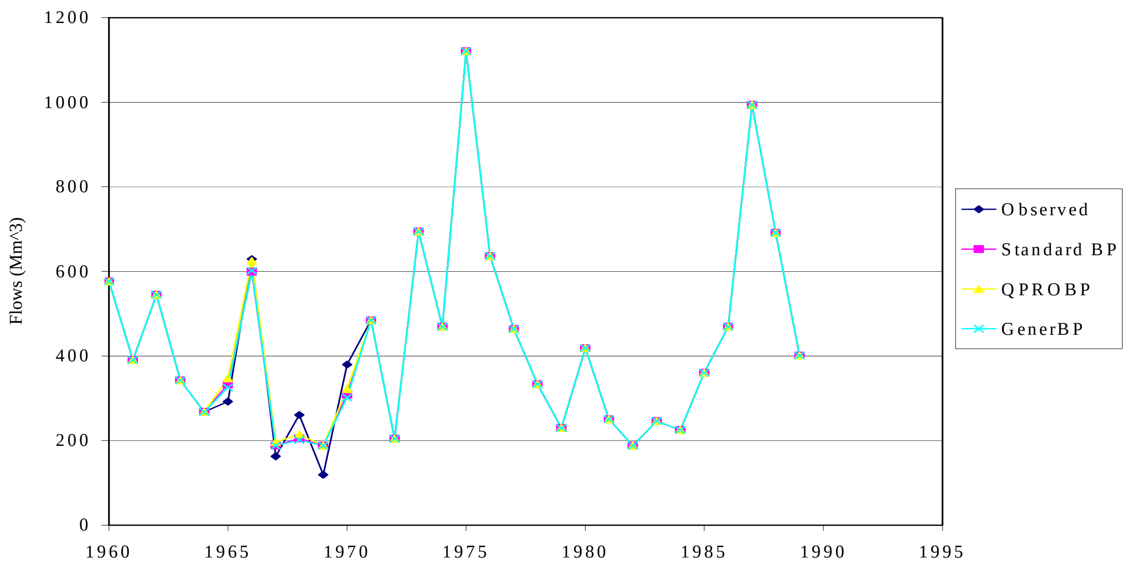


Figure 4.17 Comparison of ANN techniques in terms of hydrographs (20 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

4.1.3.3.3 Comparison of performance of EM techniques at gauge D1H003 using base gauge D1H009

The selected techniques are thought to be promising estimators of the missing values at D1H003 for gap duration of 20 % as long as they satisfy the entropy criterion.

From the results presented in Table 4.4, it can be seen that the standard EM technique and its variants have identical entropic values. Hence, the values for DIT (T) were 0.3479 (2.21 napiers) everywhere. This shows that the standard EM leads to the same solution as its variants in some cases. The raison behind that could be most probably what has been said so far. No graphical plot was made here, as the results were identical everywhere.

It can be concluded from entropy calculations that both the standard EM and its variants perform likely equally in estimating the missing annual mean flow values (i.e. 20 %) at gauge D1H003.

The statistical indicators (e.g. RMSEp, RMEp and EV) were also identical for all the EM techniques. These results enhance the conclusions drawn from entropy calculations.

4.1.3.3.4 Comparison of performance of ANN and EM techniques at gauge D1H003 using base gauge D1H009

Comparison between these techniques is exhibited in Figure 4.18 and Figure 4.19. The EM techniques and its variants showed better estimation capabilities (than ANNs) for the missing annual mean flow data at gauge D1H003. This may be due to the presence of linear relationship between gauges D1H003 and D1H009. Recall that the EM techniques used here were developed within the linear regression context.

The graphical comparison (hydrographs) between the EM and ANN techniques for gauge D1H003 is exhibited in Figure 4.19. The visual observation reveals that the hydrographs were very close.

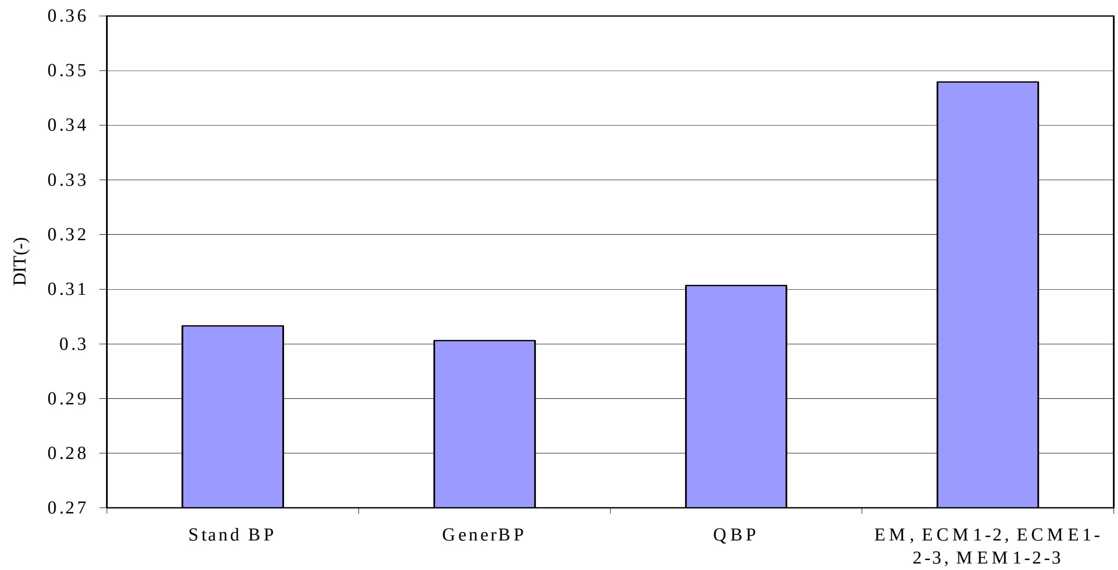


Figure 4.18 Comparison between ANNs and EM techniques in terms of DIT (20 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

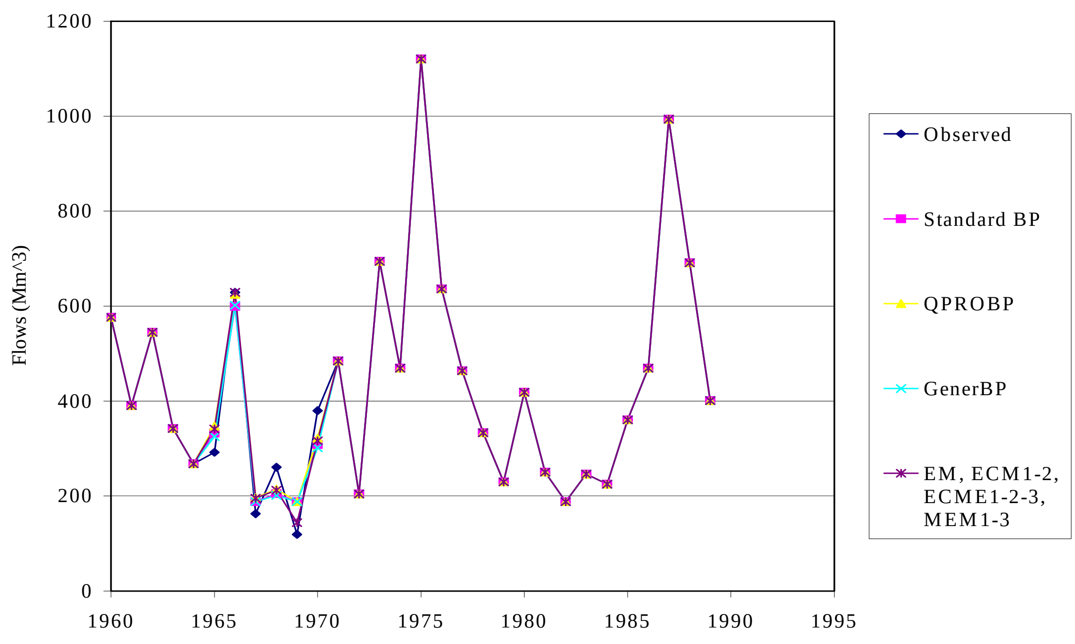


Figure 4.19 Comparison between ANN and EM techniques in terms of hydrographs (20 % missing annual mean flows from 1965 at D1H003) using base gauge D1H009

4.1.4 Infilling annual mean flows at gauge D1H009 with base gauge D1H003

4.1.4.1 Using 6.7 % missing annual mean flows at gauge D1H009

4.1.4.1.1 Selection of ANN and EM techniques for flow infilling at gauge D1H009

The results of performance for the different streamflow data infilling techniques (e.g. ANN and EM) at gauge D1H009 are summarized in Table 4.5 and Figures 4.19 (a-l). D1H003 is the base gauge for interpolating (infilling) missing annual mean flows at D1H009. Table 4.5 summarizes entropy calculations and statistics at D1H009 using ANN and EM techniques. Like in the previous section, D1H006 is not retained as base station since the computed DIT for the data series at D1H009 with base gauge D1H006 is less than the threshold (e.g. 30 %).

The values for DIT (T) range from 0.48132 (3.058 napiers) to 0.8358 (5.310 napiers). The two limits fall under the entropy criterion. In other words, the techniques presented in Table 4.5 are thought to have high capabilities of simulating information contained in the annual mean flow data at the subject gauge. It could be said that the standard BP performed as well as its variants. The same can be noticed for the EM and its variants.

With regard to the statistical indicators, the values for RMSEp, RMEp and EV range respectively from 4.29 to 30.42 Mm^3 of flows, from 0.0071 to 0.1005 and from 0.0063 to -0.079 (Table 4.5). The values for RMEp and EV, which are dimensionless, are relatively small. This can eventually confirm the results obtained from entropy criterion.

The two hydrographs (i.e. observed and estimated) were found to be close for each respective technique (Figures 4.19 (a-l)). Thus, the selected techniques could be thought to be good estimators of the missing annual mean flows at D1H009. This also confirms the results from entropy calculations.

Table 4.5 Entropy calculations and statistics at the target gauge D1H009 (6.7 % missing annual mean flows) using the base gauge D1H003

	Description	T	DIT	RMSEp	RMEp (n = 2)	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	3.77	0.5934	20.43	0.031	0.031	1.011	1.0025
	BP (sigmoid hidden layer), u = 0.05	3.633	0.5718	23.62	0.039	0.038	1.003	1.000
	McL1BP (sigmoid hidden layer)	5.310	0.8358	4.29	0.0071	0.0063	0.998	1.000
	GenerBP (sigmoid hidden layer), s=5	4.137	0.6512	10.56	0.0318	-0.006	0.999	1.020
	BP (hyperbolic tangent hidden layer),	3.300	0.5194	24.54	0.0762	-0.0078	1.018	0.998
	QBP (sigmoid hidden layer), acc = 0.015, lr = 0.85, Weight Cond	3.718	0.5852	18.96	0.0496	-0.048	1.020	1.000
	VLR (sigmoid hidden layer)	3.058	0.4813	30.42	0.1005	-0.079	1.000	0.998
EM Techniques	Standard EM	3.115	0.4903	29.30	0.0954	-0.077	1.006	0.996
	MEM3	3.116	0.4905	29.30	0.0954	-0.077	1.006	0.996
	MEM1-2, u = 0.89, lr = 0.001	3.115	0.4903	29.30	0.0954	-0.077	1.006	0.996
	ECM1-2	3.115	0.4903	29.30	0.0954	-0.077	1.006	0.996
	ECME1, 2 (k = 2), 3	3.115	0.4903	29.30	0.0954	-0.077	1.006	0.996

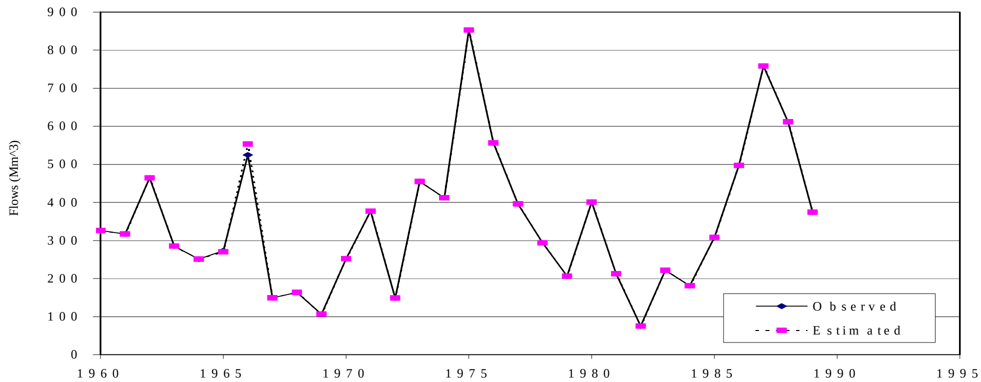


Figure 4.20a Annual mean flows at D1H009 (6.7 % missing from 1965) using standard BP with base gauge D1H003

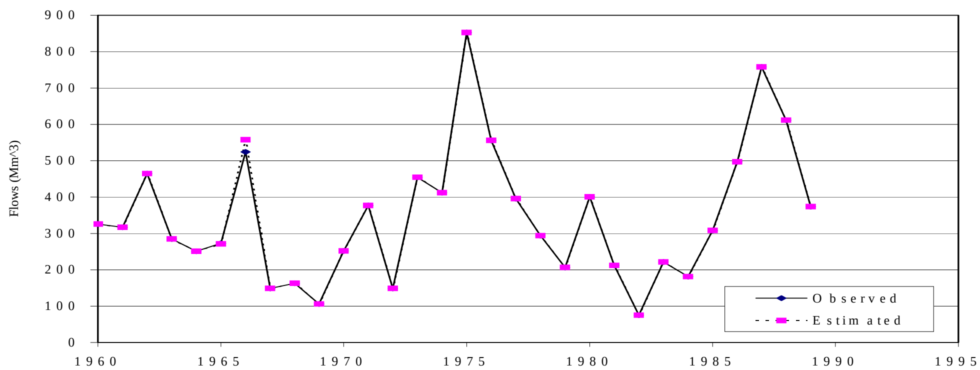


Figure 4.20b Annual mean flows at D1H009 (6.7 % missing from 1965) using MBP with base gauge D1H003

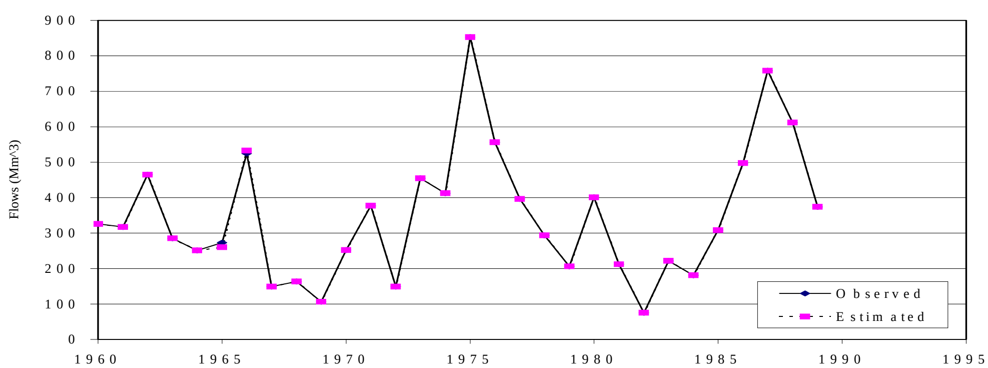


Figure 4.20c Annual mean flows at D1H009 (6.7 % missing from 1965)using McL1BP with base gauge D1H003

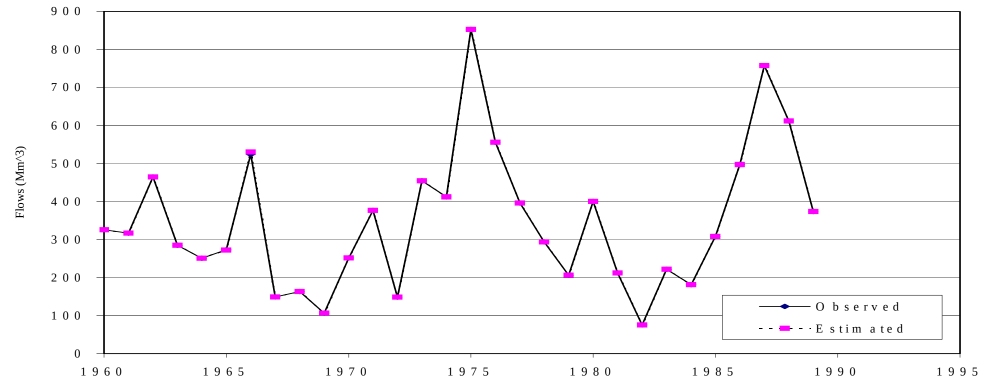


Figure 4.20d Annual mean flows at D1H009 (6.7%) missing from 1965) using GenerBP with base gauge D1H003

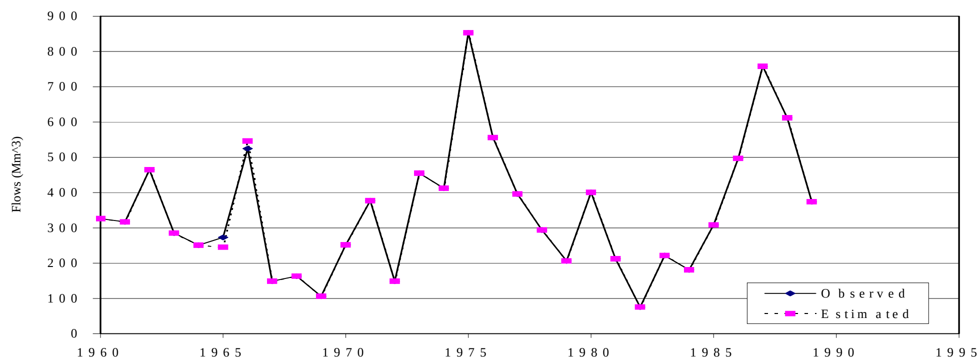


Figure 4.20e Annual mean flows at D1H009 (6. 7 % missing from 1965) using BP (hyperbolic tangent hidden layer) with base gauge D1H003

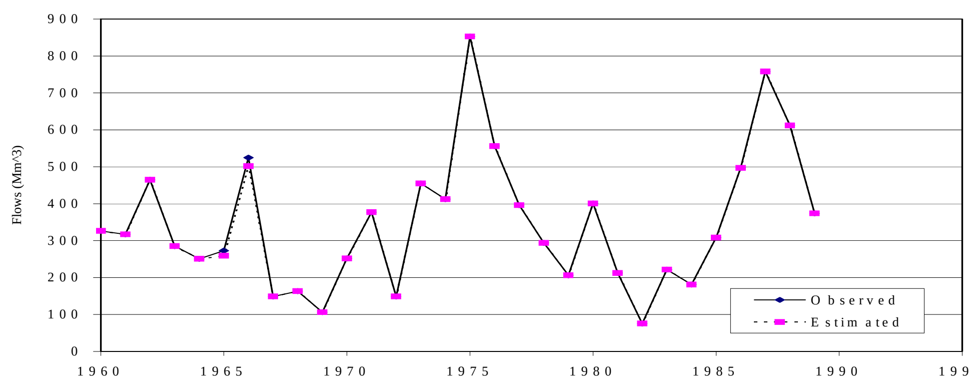


Figure 4.20f Annual average flows at D1H009 (6.7 % missing from 1965) using QBP with base gauge D1H003

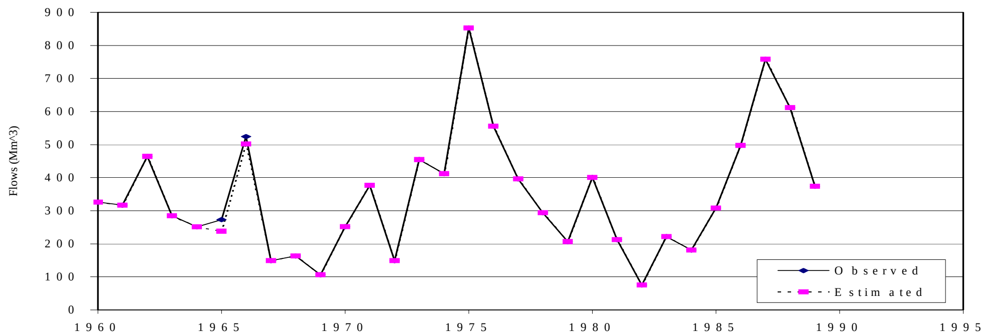


Figure 4.20g Annual mean flows at D1H009 (6.7 % missing from 1965) using VLRBP with base gauge D1H003

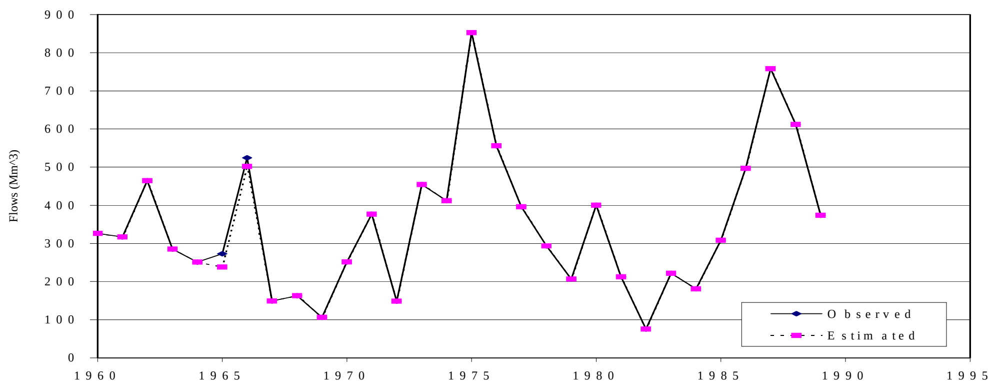


Figure 4.20h Annual mean flows at D1H009 (6.7 % missing from 1965) using standard EM with base gauge D1H003

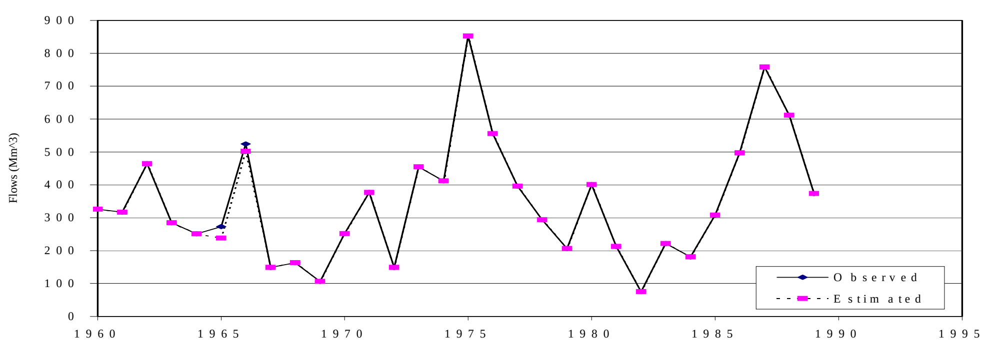


Figure 4.20i. Annual mean flows at D1H009 (6.7 % missing from 1965) using MEM1-2 with base gauge D1H003

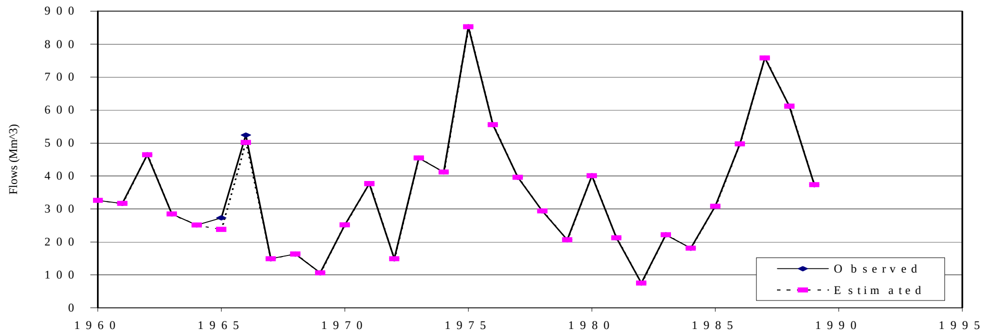


Figure 4.20j Annual mean flows at D1H009 (6.7 % missing from 1965) using MEM3 with base gauge D1H003

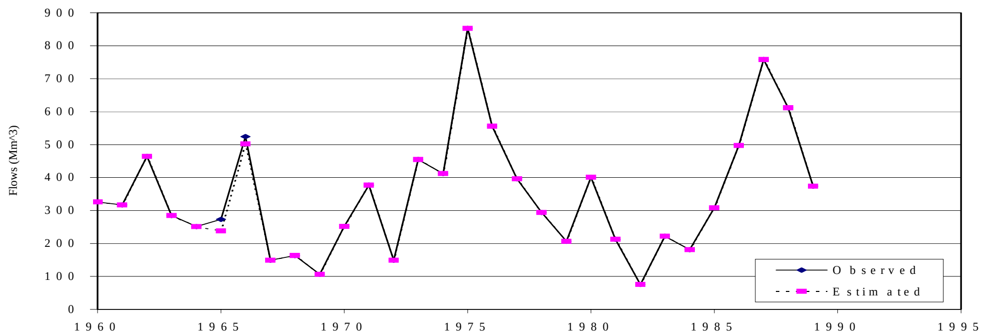


Figure 4.20k Annual mean flows at D1H009 (6.7 % missing from 1965) using ECM1-2 with base gauge D1H003

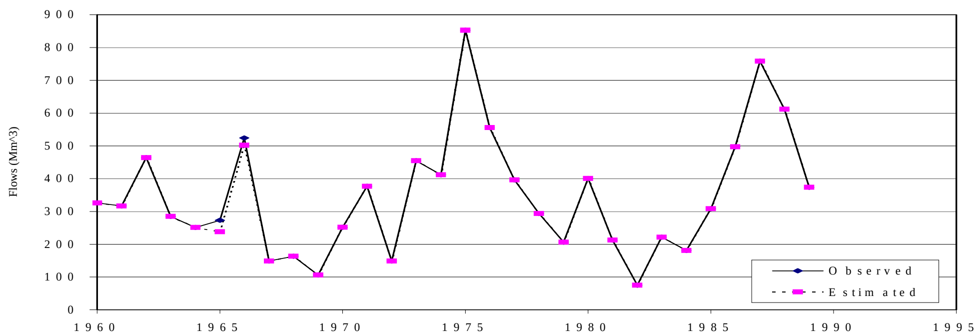


Figure 4.20l Annual mean flows at D1H009 (6.7 % missing from 1965) using ECME1-2 with base gauge D1H003

4.1.4.1.2 Comparison of performance of ANN techniques at gauge D1H009 using base gauge D1H003

Considering D1H003 as base gauge and for gap duration of 6.7 % at D1H009, the performance (in a decreasing order) of the different ANN techniques in terms of DIT (T) values is as follows:

-McL1BP

-GenerBP

-Standard BP

-QBP

-MBP

-BP (hyperbolic tangent hidden layer)

-VLRBP

From Table 4.5 and Figure 4.21, the values for DIT (T) range from 0.48132 (3.058 napiers) to 0.8358 (5.310 napiers). Thus, the first best technique (i.e. McL1BP) among the selected ANN techniques has the highest entropic value of 0.8358 (5.310 napiers) for DIT (T). In other words, 83.58 % of uncertainty can be removed from the annual mean flows at gauge D1H009 via the McL1BP technique. The McL1BP technique is thought to be relatively more capable (than the rest) of estimating the missing annual mean flows at D1H009. It is believed that the Mac Laurin order 1 power series could be a better approximation for the sigmoid function within the scaled input data interval (0.1, 0.9) for this specific gap size at D1H009. The GenerBP is also believed to make it through the generalization parameter. The standard BP was shown to perform better than QBP, MBP BP (hyperbolic tangent hidden layer) and VLRBP.

The last best technique among these selected ANN techniques VLRBP has the lowest entropic value of 0.48132 (3.058 napiers) for DIT(T). It can be said that 48.13 % of uncertainty can be removed from the annual mean flows at gauge D1H009 via the VLRBP technique.

Figure 4.21 shows also that the values for DIT are close for two successive techniques, except for McL1BP and GenerBP.

On the other hand, the values for RMSEp range from 4.29 to 30.42 Mm^3 of flows. The values of RMEp range from 0.0071 to 0.1005. The values of EV range from 0.0063 to 0.0079. The McL1BP technique has the lowest values of statistical indicators while the VLRBP has the highest values of the same statistical indicators. These results just confirm the conclusions drawn from the entropy calculations.

Figure 4.22 shows the graphical comparison of the hydrographs for the different techniques. A small difference could be noticed between McL1BP and the rest of the ANN techniques. Again, this confirms the results from entropy calculations.

In general, these techniques appear to be good estimators of the missing annual mean flows (6.7 %) at gauge D1H009. The selected ANNs preserved the mean and the variance for gauge D1H009 (see Table 4.5).

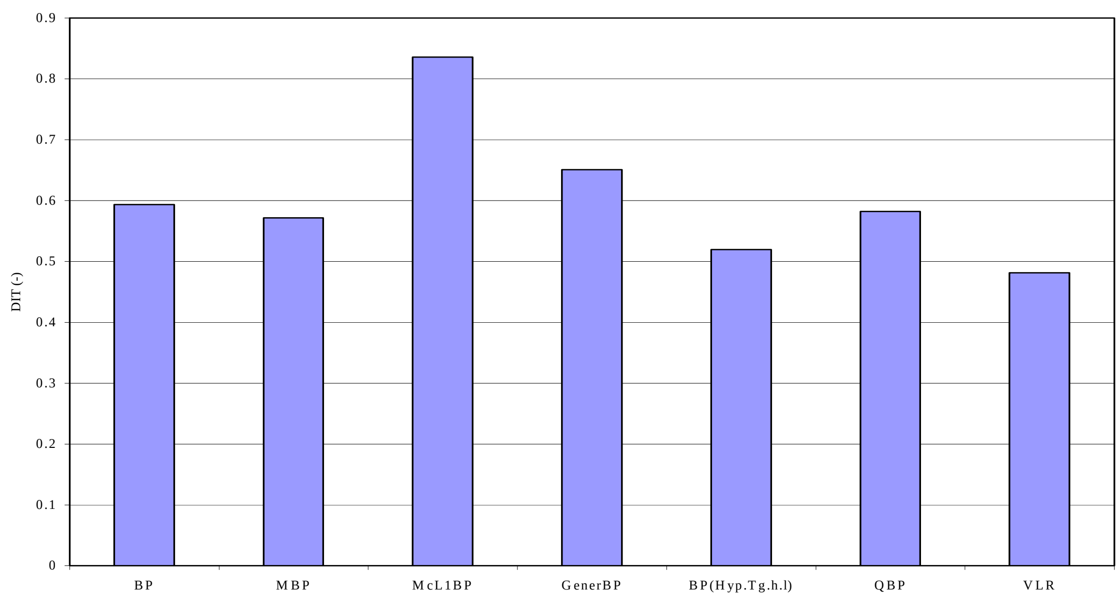


Figure 4.21 Comparison of ANNs in terms of DIT (6.7 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

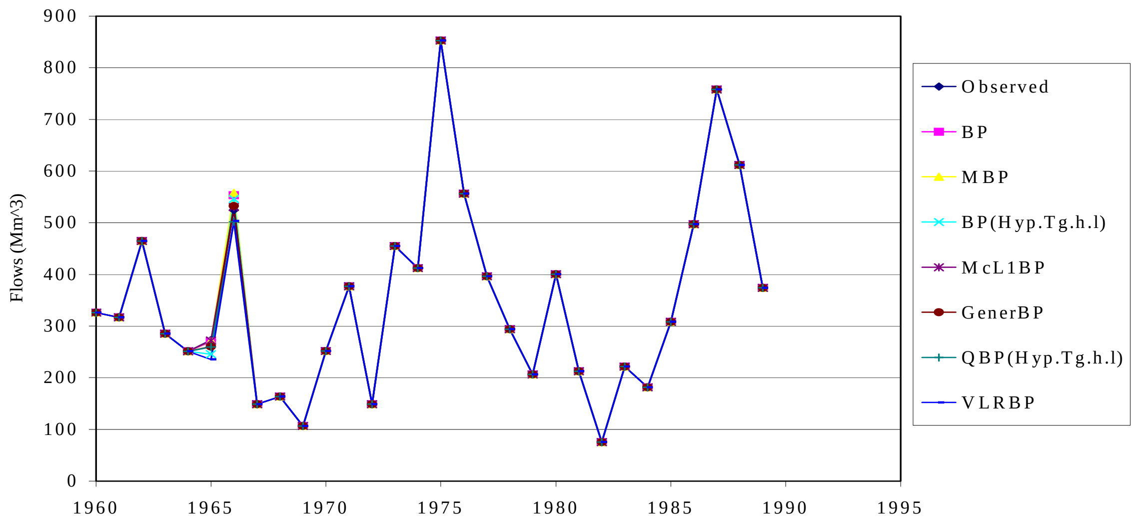


Figure 4.22 Comparison of ANNs in terms of hydrographs (6.7 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

4.1.4.1.3 Comparison of performance of EM techniques at gauge D1H009 using base gauge D1H003

From the results shown in Table 4.5 and Figure 4.23, it can be seen that the standard EM technique and its variants have almost identical entropic values although very slight (negligible) improvements were noticed only for MEM3, compared to the standard EM technique. This shows that the standard EM leads to the same solution as its variants as noticed previously.

It can be concluded from entropy calculations that both the standard EM and its variants perform likely equally in estimating the missing annual mean flows (e.g. 6.7 %) at gauge D1H009. The use of the momentum or the learning rate in the different versions, as done in this study, would seem to be optional since they don't have any substantial impact on the accuracy of the estimated annual mean flows at D1H009.

The statistical indicators were also identical for all the EM techniques. These results enhance the conclusions drawn from entropy calculations. These techniques also preserved the mean and the variance at D1H009 as shown in Table 4.5. The graphical plot (Figure 4.24) showed that the hydrographs for the different EM techniques are identical.

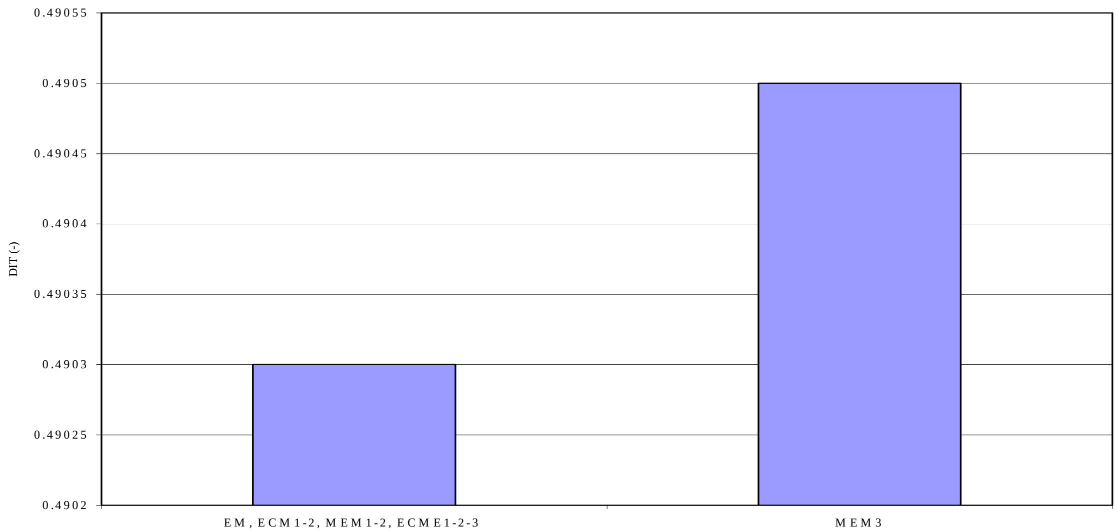


Figure 4.23 Comparison of EM techniques in terms of DIT (6.7 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

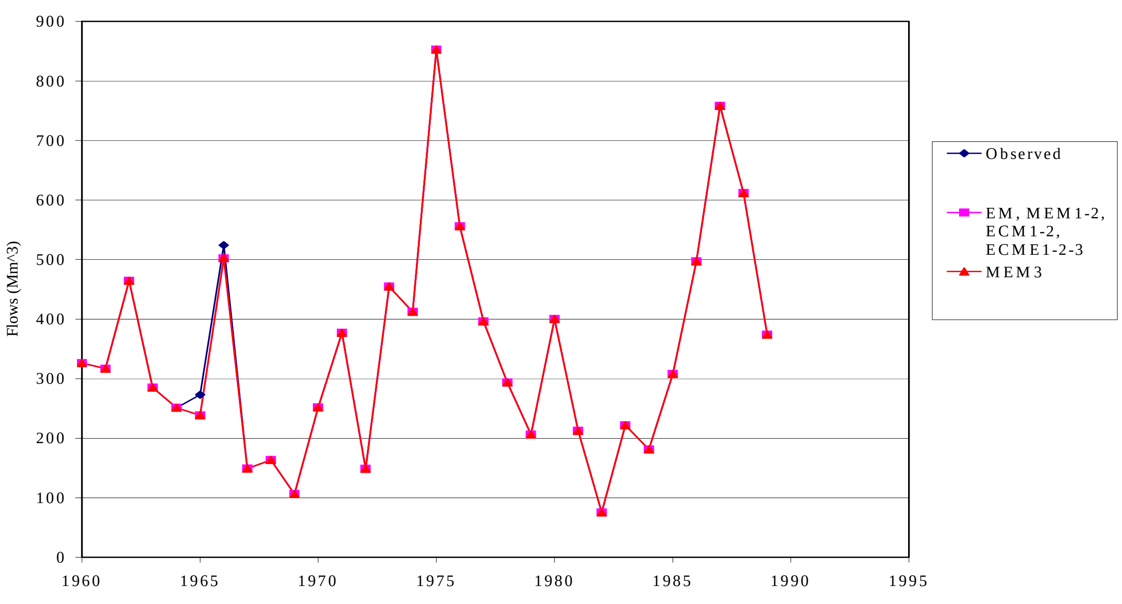


Figure 4.24 Comparison of EM techniques in terms of DIT (6.7 % missing annual mean flows from 1965 at D1H009) using base gauge DH1003

4.1.4.1.4 Comparison of performance of ANN and EM techniques at gauge D1H009 using base gauge D1H003

Comparison between ANNs and EM techniques is exhibited in Figure 4.25 and Figure 4.26. The EM techniques come in the seventh position after the MBP technique (Table 4.5). Despite their position, the EM techniques show good estimation capabilities for the missing annual mean flows at gauge D1H009. This may be due to the presence of linear relationship between the gauges D1H003 and D1H009, since the EM techniques used here were developed within the linear regression context. A similar observation was also made by Khalil et al. (2001), when using linear regression methods.

The best first five ANNs are thought to have higher capabilities (than EM techniques) of mapping the non-linear characteristics in the missing annual flows at D1H009, through the use of a sigmoid/hyperbolic hidden layer. On the other hand, the EM techniques were shown to perform slightly better than the VLR. The graphical comparison (hydrographs) between the EM and ANN techniques for the subject gauge D1H009 is exhibited in Figure 4.26. The visual observation revealed that the hydrographs were close between ANN and EM techniques.

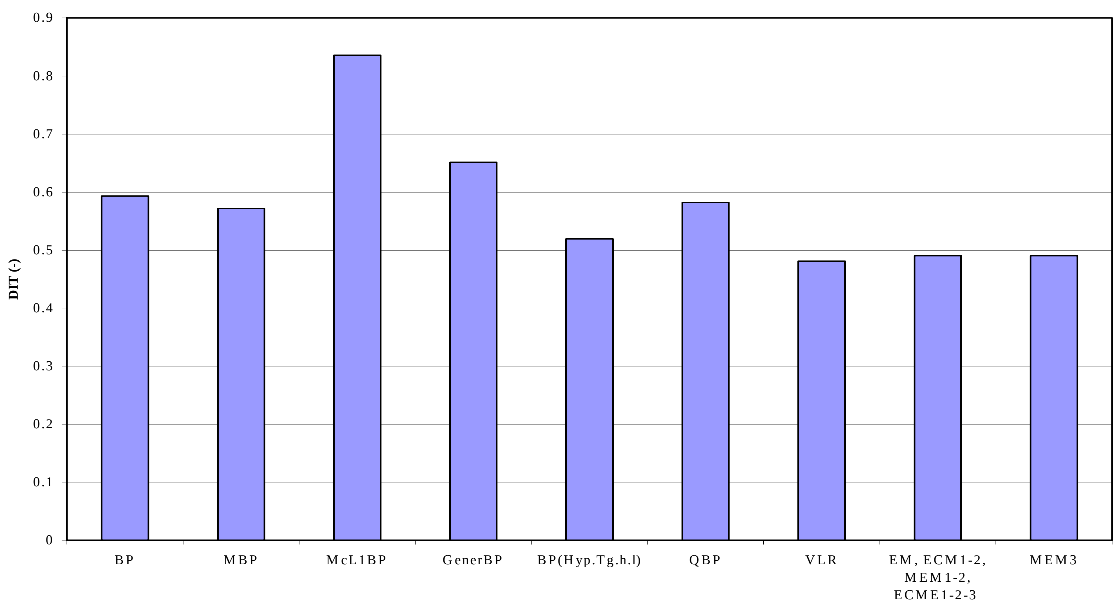


Figure 4.25 Comparison between ANN and EM in terms of DIT (6.7 % missing annual mean flows from 1965 at D1H009 from 1965) using base gauge D1H003

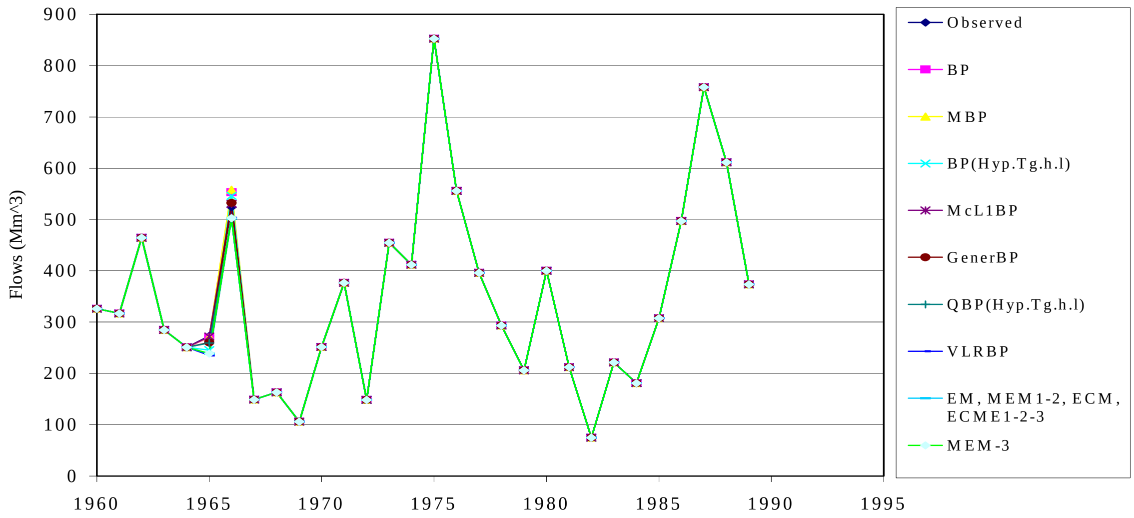


Figure 4.26 Comparison between ANN and EM in terms of hydrographs (6.7 % missing annual mean flows from 1965 at D1H009 from 1965) using base gauge D1H003

4.1.4.2 Using 13.4 % of missing annual mean flow values at gauge D1H009

4.1.4.2.1 Selection of ANN and EM techniques for flow infilling at D1H009

The results of performance for the different streamflow data interpolation (infilling) techniques (e.g. ANN and EM) at gauge D1H009 are summarized in Table 4.6 and Figures 4.27 (a-l). The values for DIT (T) range from 0.3280 (2.084 napiers) to 0.3861 (2.453 napiers). These two limits fall under the entropy criterion. It could be said that the standard BP performed as well as its variants. The same could be said for the standard EM technique and its variants. The values for RMSEp, RMEp and EV range respectively from 32.28 Mm^3 to 48.12 Mm^3, from 0.12 to 0.14 and from –0.013 to 0.105, as shown in Table 4.6. The values for RMEp and EV, which are dimensionless, are relatively small. The two hydrographs (i.e. observed and estimated) were found to be close for each respective technique (Figure 4.27 (a-l)). The results from both statistical indicators and graphical plots confirm the results from entropy calculations. Thus, the selected techniques are thought to be good estimators of the missing annual mean flows at gauge D1H009.

Table 4.6 Entropy calculations and statistics at the target gauge D1H009 (13.4% missing annual mean flows) using the base gauge D1H003

	Description	T (napiers)	DIT	RMSEp (Mm^3)	RMEp (n = 4)	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	2.283	0.3593	40.123	0.115	-0.029	0.980	0.9985
	BP (sigmoid hidden layer), u = 0.05	2.257	0.3553	41.23	0.121	0.0849	0.995	1.000
	Mac Laurin 1 BP (sigmoid hidden layer)	2.084	0.3280	48.12	0.14	0.105	1.002	0.997
	GenerBP (sigmoid hidden layer), s = 2	2.315	0.3644	38.30	0.10	0.07	1.05	0.999
	BP (hyperbolic tangent hidden layer)	2.396	0.3771	33.76	0.152	0.0143	1.03	1.00
	QBP (hyperbolic tangent hidden layer), acc = 0.025, lr = 0.25, Weight Cond	2.453	0.3861	32.95	0.12	-0.013	1.014	0.985
	VLR (sigmoid hidden layer)	2.444	0.3847	32.28	0.136	-0.013	0.993	0.999
EM Techniques	Standard EM	2.452	0.3859	33.04	0.126	-0.013	1.013	1.00
	MEM1(u = 0.95), 3	2.452	0.3859	33.04	0.126	-0.013	1.013	1.00
	MEM2 (u = 0.95, lr = 0.01)	2.438	0.3837	32.78	0.108	0.013	1.012	1.00
	ECM1-2	2.452	0.3859	33.04	0.126	-0.013	1.013	1.00
	ECME1, 2(k=3), 3 (u = 0.90)	2.452	0.3859	33.04	0.126	-0.013	1.013	1.00

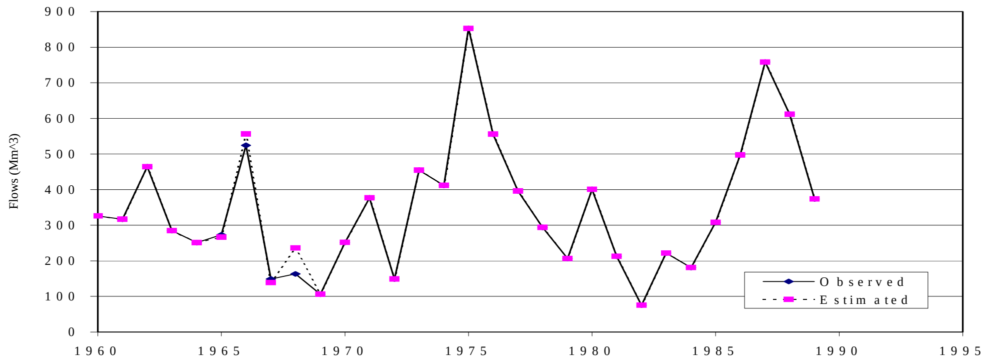


Figure 4.27a Annual mean flows at D1H009 (13.4 % missing from 1965) using standard BP with base gauge D1H003

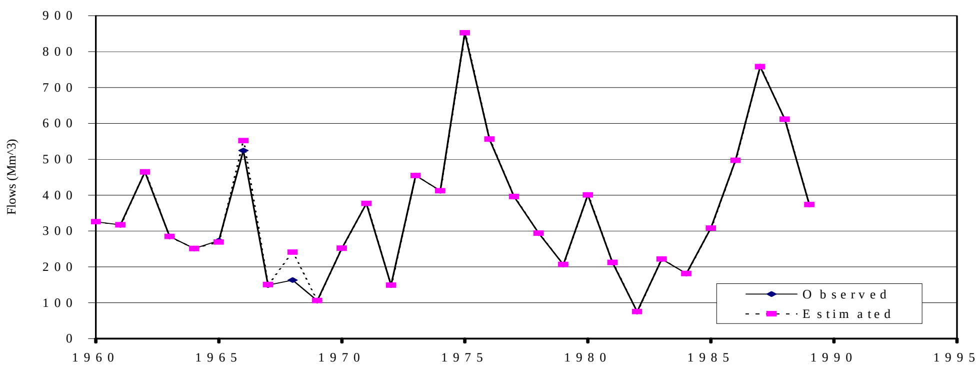


Figure 4.27b Annual mean flows at D1H009 (13.4 % missing from 1965) using MBP with base gauge D1H003

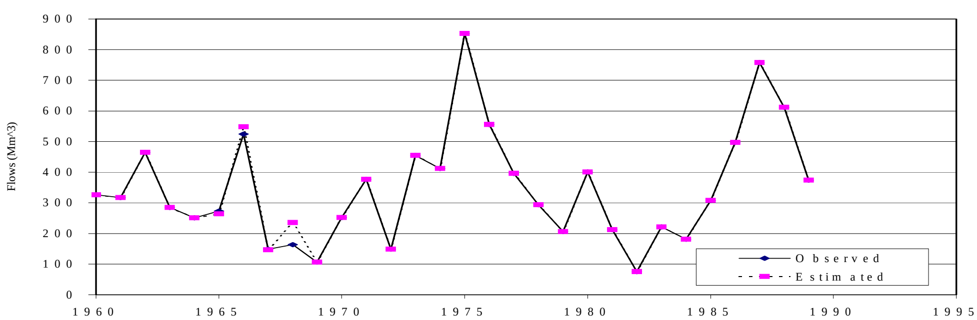


Figure 4.27c Annual mean flows at D1H009 (13.4 % missing from 1965) using GenerBP with base gauge D1H003

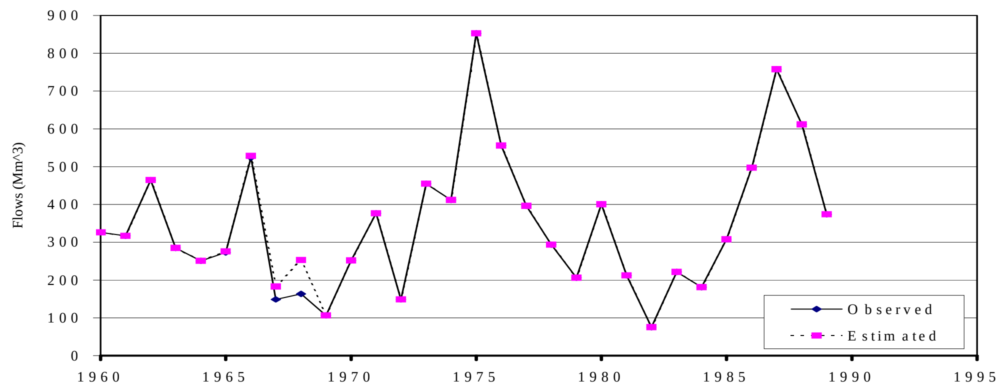


Figure 4.27d Annual mean flows at D1H009 (13.4 % missing from 1965) using McL1BP with base gauge D1H003

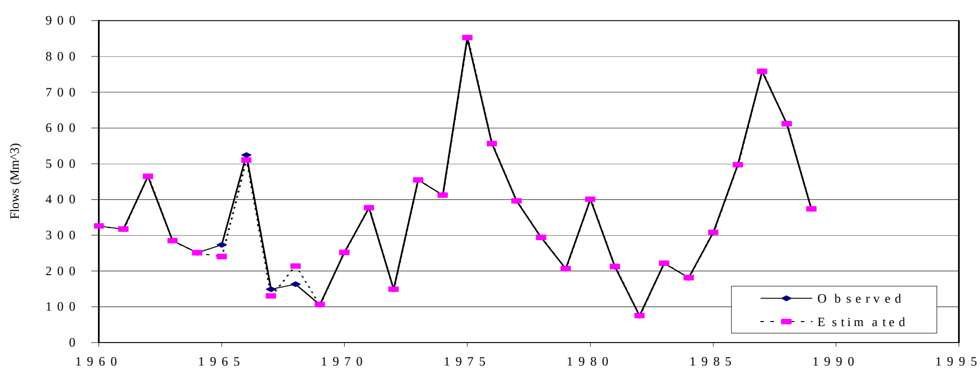


Figure 4.27e Annual mean flows at D1H009 (13.4 % missing from 1965) using BP (hyperbolic tangent hidden layer) at gauge D1H003

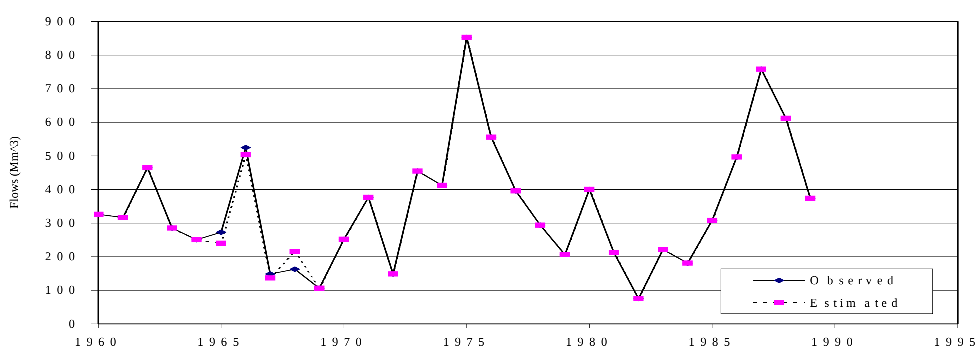


Figure 4.27f Annual mean flows at D1H009 (1.3% missing from 1965) using QBP with base gauge D1H003

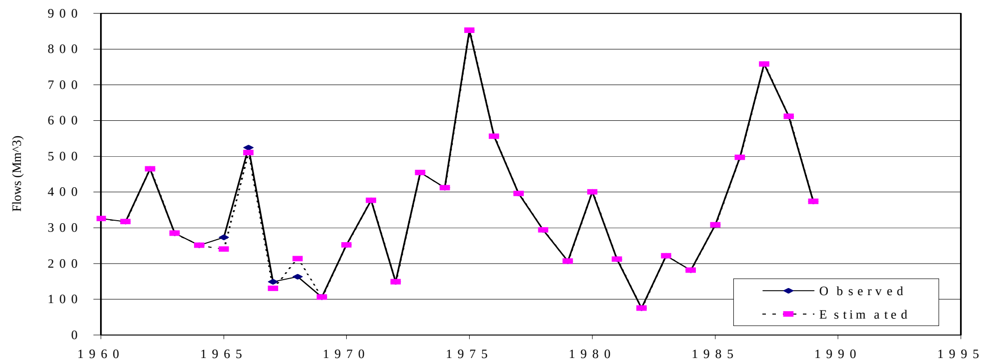


Figure 4.27g Annual mean flows at D1H009 (1.3.4 % missing from 1965) using VLRBP with base gauge D1H003

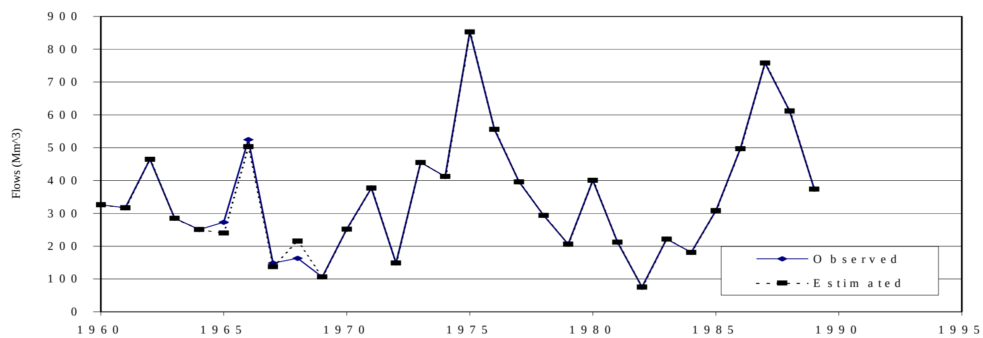


Figure 4.27h Annual mean flows at D1H009 (1.3.4 % missing from 1965) using standard EM with base gauge D1H003

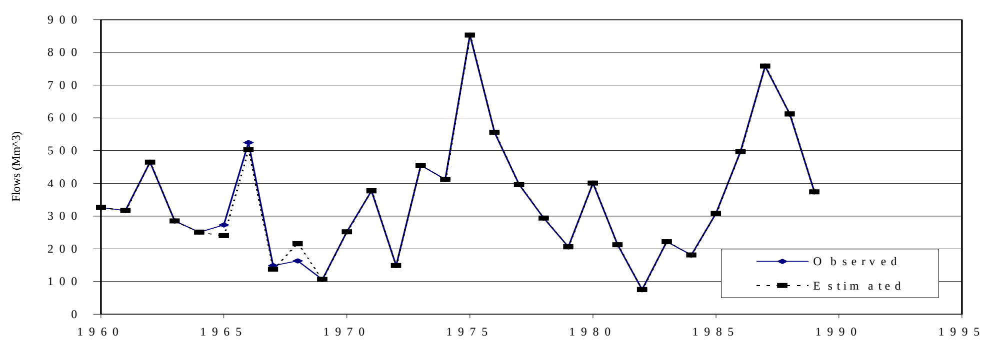


Figure 4.27i. Annual mean flows at D1H009 (1.3.4 % missing from 1965) using MEM1-3 base gauge D1H003

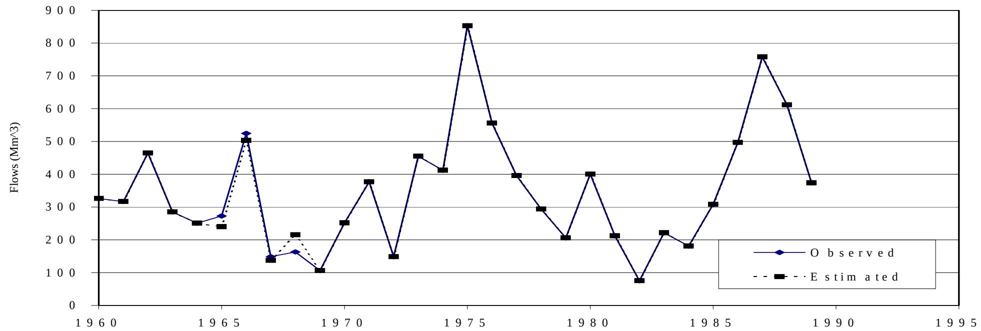


Figure 4.27j Annual mean flows at D1H009 (1.3.4 % missing from 1965) using MEM2 with base gauge D1H003

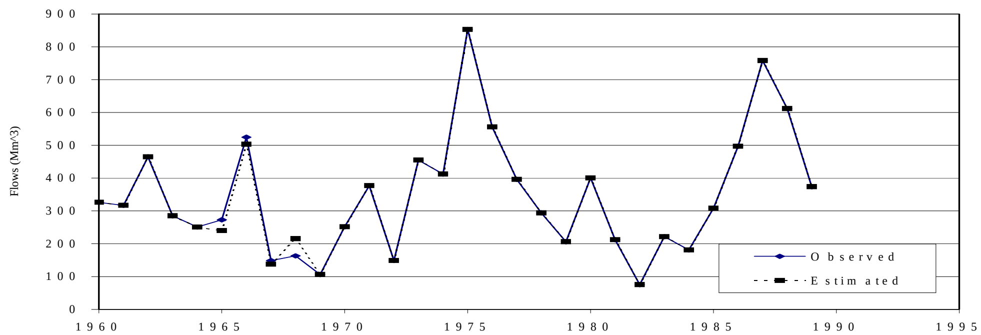


Figure 4.27k Annual mean flows at D1H009 (1.3.4 % missing from 1965) using ECM1-2 with base gauge D1H003

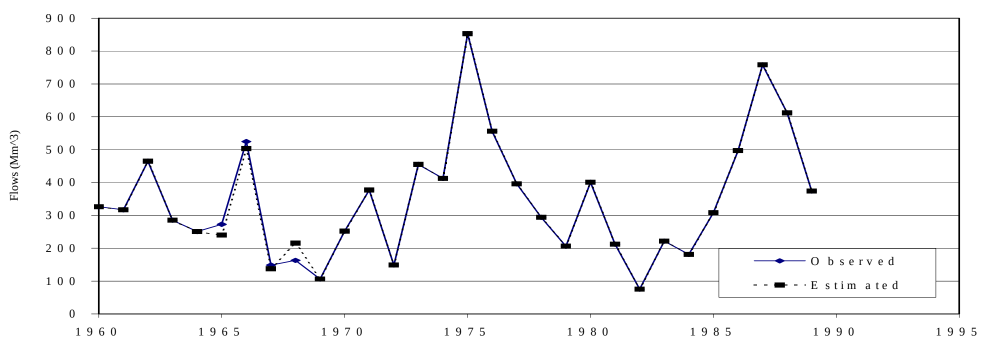


Figure 4.27l Annual mean flows at D1H009 (1.3.4 % missing from 1965) using ECME1-2-3 with base gauge D1H003

4.1.4.2.2 Comparison of performance of ANN techniques at gauge D1H009 using base gauge D1H003

For gap duration of 13.4 % at D1H009, the performance (in a decreasing order) of the different ANN techniques in terms of DIT (T) values was as follows:

-QBP

-VLR

-BP (hyperbolic tangent hidden layer)

-GenerBP

-BP (sigmoid hidden layer)

-MBP

-McL1BP

From Table 4.6 and Figure 4.28, the values of DIT (T) range from 0.3280 (2.084 napiers) to 0.3861 (2.453 napiers). Thus, the first best technique (i.e. QBP) among the selected ANN techniques has the highest entropic value of 0.3861 (2.453 napiers) for DIT (T). In other words, 38.61 % of uncertainty can be removed from the annual mean flow series at gauge D1H009 via the QBP technique. The QBP technique is thought to be relatively more capable (than the rest) of estimating the missing annual mean flows at gauge D1H009. The VLR technique was shown to be also better (than the standard BP) in the estimation of the missing annual mean flows at D1H009. However, the standard BP could perform better than GenerBP, MBP and McL1BP techniques.

It can be noticed that 37.71 % of information can be physically inferred about gauge D1H009 via the standard BP (hyperbolic tangent hidden layer). The last best model among these selected ANN techniques (i.e. McL1BP) has the lowest entropic value of 0.3280 (2.084 napiers) for DIT (T). It can be said that 32.80 % of uncertainty can be removed from the data series at gauge D1H009 via McL1 BP. Figure 4.28 shows the comparison of entropic values for the different ANNs.

On the other hand, the values for RMSEp range from 32.28 to 48.12 Mm^3 of flows. The values of RMEp range from 0.12 to 0.14. The values of EV range from -0.013 to 0.105.

The QBP technique has the lowest values of RMSEp, RMEp and EV while McL1BP has the highest value for RMSEp, RMEp and EV. The statistical indicators confirm the conclusions drawn from entropy calculations.

Figure 4.29 summarizes the comparison of the performance of the different ANN techniques in terms of their hydrographs. This figure shows that the results are close. This also confirms the results from entropy calculations.

In general, the techniques presented above appear to be good estimators of the missing annual mean flows (13.4 %) at gauging station D1H009. For this gap duration, the above-selected ANNs preserved the mean and the variance of the estimated series compared to the observed one at D1H009 (see Table 4.6).

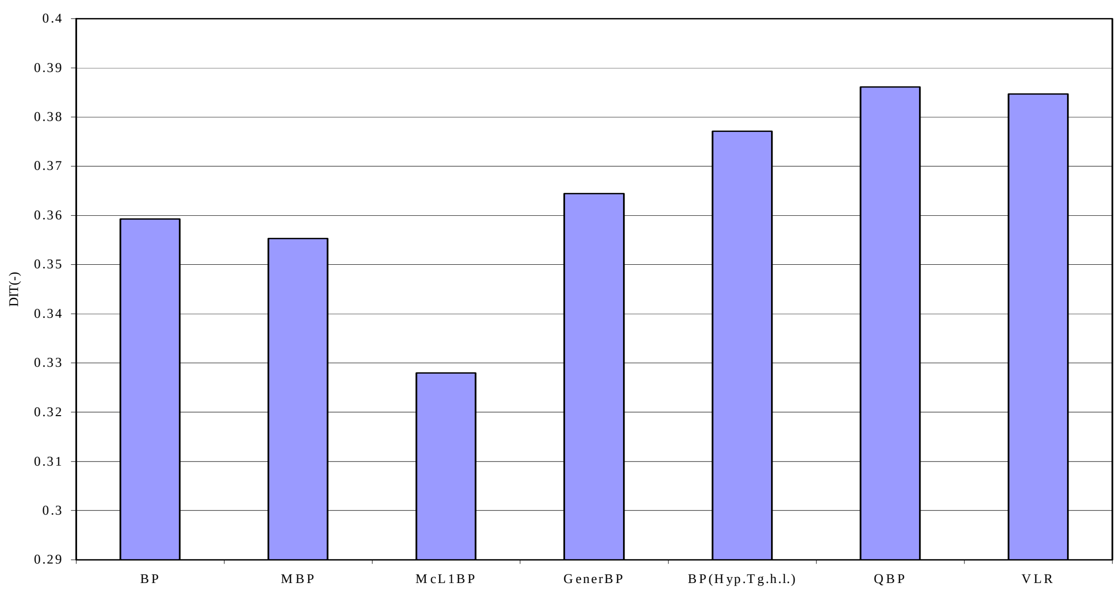


Figure 4.28 Comparison of ANN in terms of DIT (13.4 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

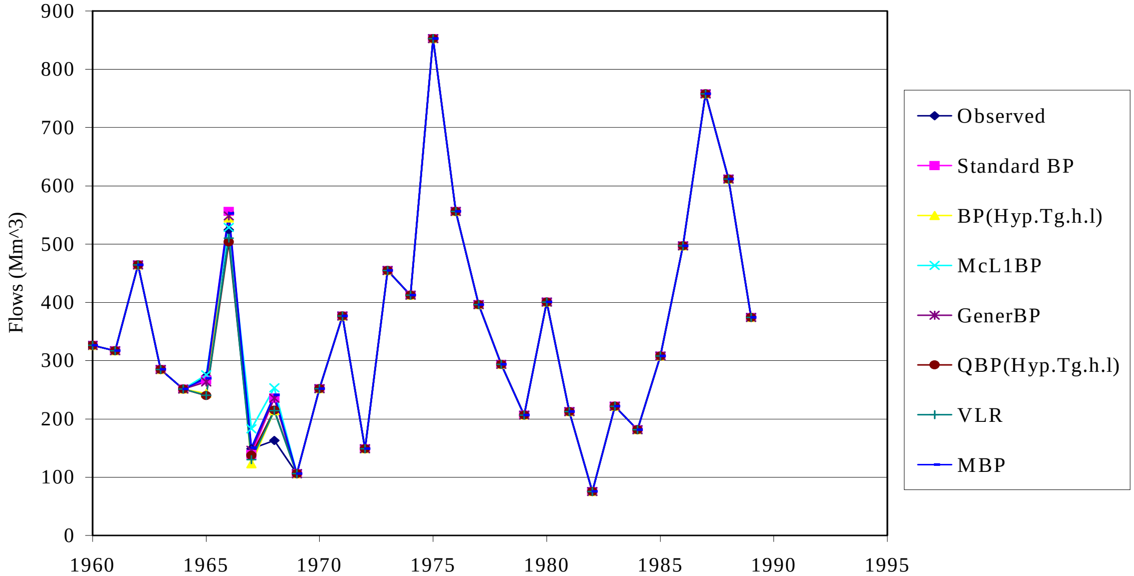


Figure 4.29 Comparison of ANN in terms of hydrographs (13.4 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

4.1.4.2.3 Comparison of performance of EM techniques at gauge D1H009 using base gauge D1H003

From the results shown in Table 4.6 and Figure 4.30, it can be seen that the standard EM technique and its variants have almost identical entropic values. This shows that the standard EM leads to the same solution as its variants in some cases.

From entropy calculations, it can be concluded that both the standard EM and its variants perform likely equally in estimating the missing annual mean flows (e.g. 13.4 %) at gauge D1H009.

The statistical indicators were almost identical for all the EM techniques. The EM techniques preserved the mean and the variance for D1H009 as shown in Table 4.6.

From a visual observation, all the estimated hydrographs were almost identical (Figure 4.31). Again, this is a confirmation of the results obtained from the entropy calculations.

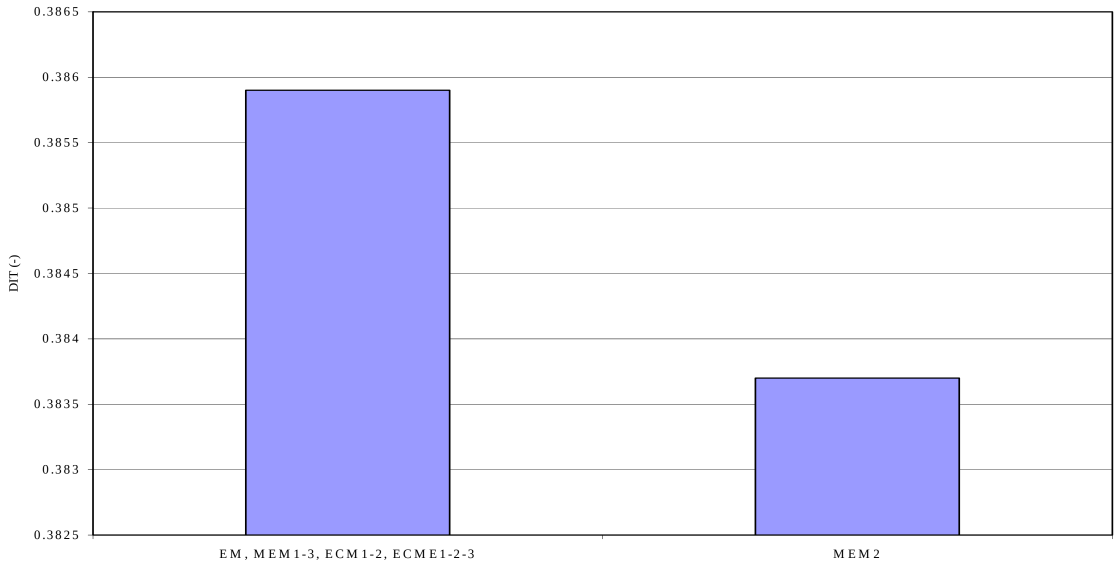


Figure 4.30 Comparison of EM in terms of DIT (13.4 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

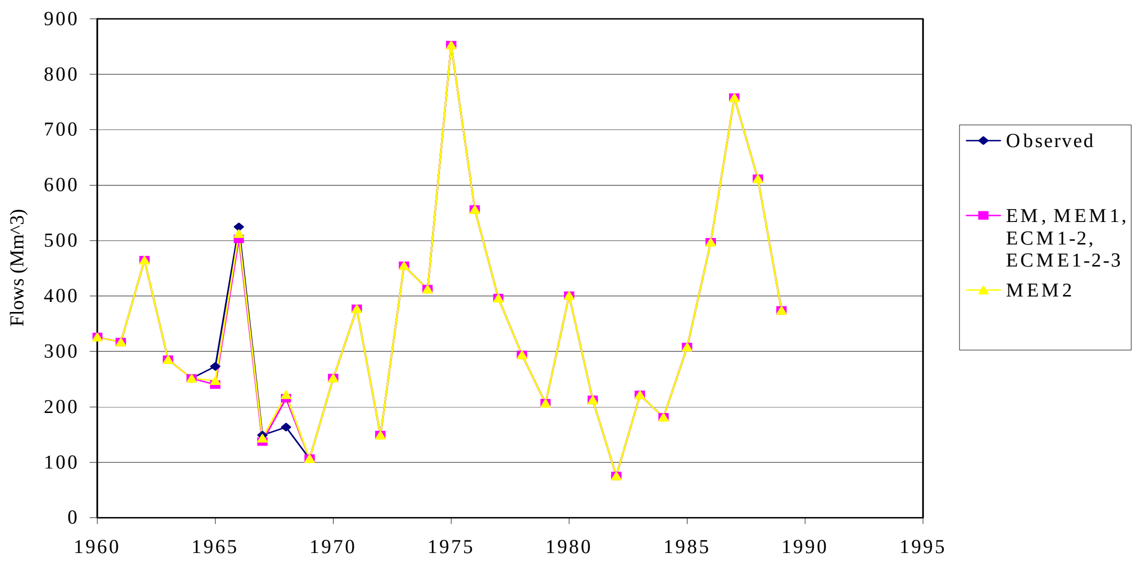


Figure 4.31 Comparison of EM techniques in terms of hydrographs (13.4 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

4.1.4.2.4 Comparison of performance of ANN and EM techniques at gauge D1H009

The results for the comparison between ANNs and EM techniques are depicted in Figure 4.32 and Figure 4.33. For visual convenience, only the best first three ANNs and all the EM techniques are presented here. The EM techniques come in the second position after the QBP technique, however the results are very close. The EM techniques show good estimation capabilities for the missing data at gauging station D1H009. This may be due to the presence of linear relationship between gauges D1H003 and D1H009.

The first best ANN technique, QBP was found to be marginally better than the EM technique according to the entropy criterion. The difference in entropic calculations between this ANN and the EM techniques was very small. In other words, the above-mentioned ANN is thought to have slightly a higher capability of mapping the non-linear characteristics in the missing values through the use of a hyperbolic hidden layer. On the other hand, the EM techniques were shown to perform slightly better than the rest of ANNs in the estimation of missing annual mean flows at D1H009.

The graphical comparison (hydrographs) between the EM and first three ANN techniques for the gauge D1H009 is exhibited in Figure 4.33. The visual observation reveals that the hydrographs were close.

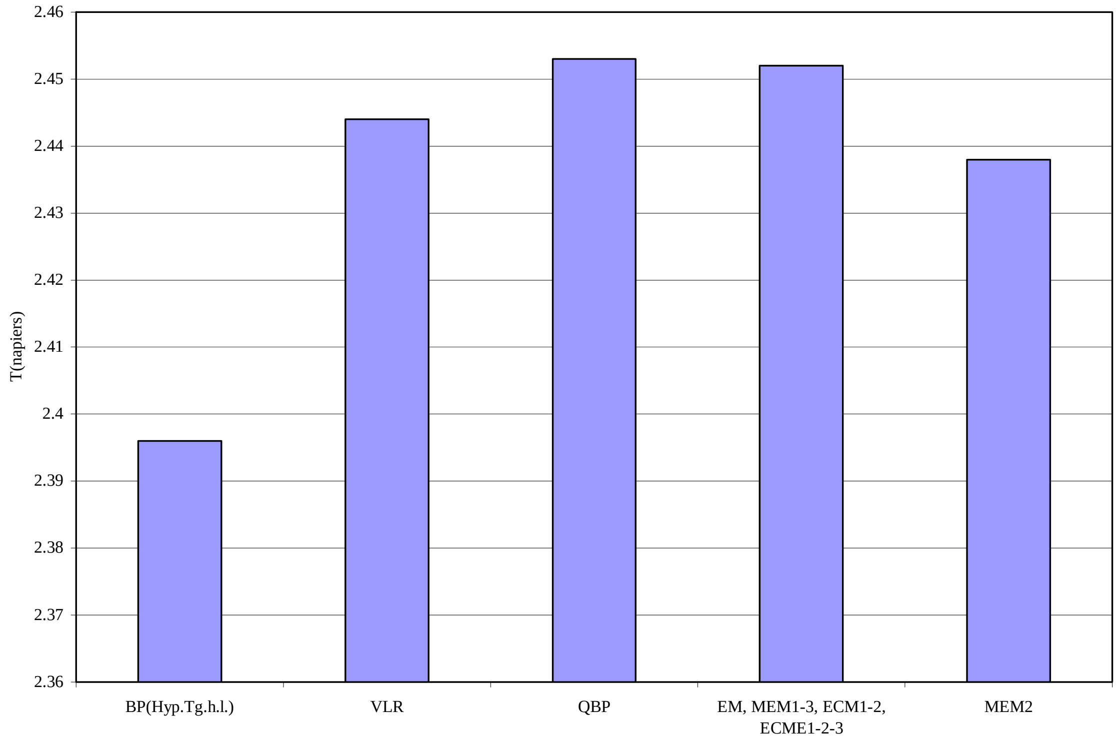


Figure 4.32 Comparison of ANN and EM in terms of T (13.4 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

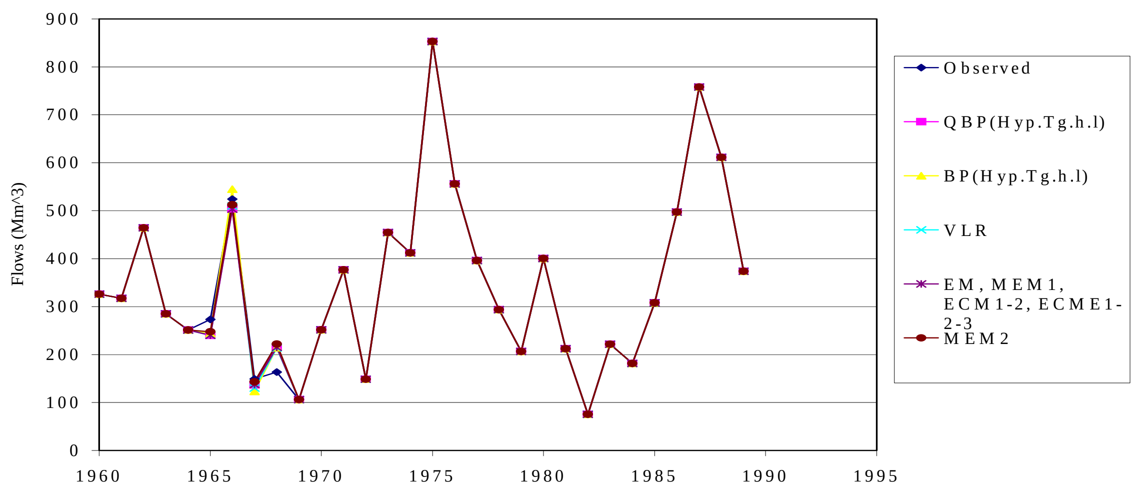


Figure 4.33 Comparison of ANN and EM techniques in terms of hydrographs (13.4 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

4.1.4.3 Using 20 % of missing annual mean flows at D1H009 gauge

4.1.4.3.1 Selection of ANN and EM techniques for flow infilling at D1H009

The results of performance for the different streamflow data interpolation (infilling) techniques (e.g. ANN and EM) at the subject gauge D1H003 are summarized in Table 4.7 and Figures 4.34 (a-k.). Table 4.7 summarizes the entropy calculations and statistics at D1H009 using ANN and EM techniques.

The values of DIT (T) for the selected techniques range from 0.3006 (1.910 napiers) to 0.3493 (2.220 napiers). These two limits satisfy the entropy criterion. It could be said that the standard BP technique performed as well as its variants. The same could be noticed for the standard EM and its variants.

The values of statistical indicators viz. RMSEp, RMEp and EV range respectively from 36.00 to 48.51 Mm^3, from 0.101 to 0.136 and from 0.995 to 1.02 (Tale 4.7). The values for RMEp and EV, which are dimensionless, are relatively small. The two hydrographs (i.e. observed and estimated) were found to be close for each respective technique. Thus, the selected techniques are thought to be good estimators of the missing annual mean flows at gauge D1H009. Both statistical indicators and graphical plot confirm the results from entropy calculations.

Table 4.7 Entropy calculations and statistics at the target gauge D1H009 (20 % missing annual mean flows) using the base gauge D1H003

	Description	T (napiers)	DIT	RMSEp (Mm^3)	RMEp (n = 6)	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	1.971	0.3102	48.51	0.125	0.13	0.985	0.995
	BP (hyperbolic tangent hidden layer)	1.992	0.3135	45.86	0.110	0.125	0.998	0.990
	GenerBP (sigmoid hidden layer), s=5	1.91	0.3006	44.28	0.136	0.138	1.02	1.000
	QBP (sigmoid hidden layer), acc = 0.025, lr = 0.85, Weight Cond	2.168	0.3412	38.44	0.105	0.115	0.995	0.962
	GoldSBP(sigmoid hidden layer), lr = 0.25	1.999	0.3132	46.87	0.102	0.104	0.996	0.963
	VLR (sigmoid hidden layer)	2.136	0.3362	35.95	0.1	0.116	1.03	0.997
EM Techniques	Standard EM	2.219	0.3493	36.052	0.101	0.995	1.02	0.995
	MEM1-2	2.219	0.3493	36.052	0.101	0.995	1.02	0.995
	MEM3	2.220	0.3494	36.00	0.101	0.995	1.021	0.995
	ECM1-2	2.219	0.3493	36.052	0.101	0.995	1.02	0.995
	ECME1, 2 (k = 2)	2.219	0.3493	36.052	0.101	0.995	1.02	0.995

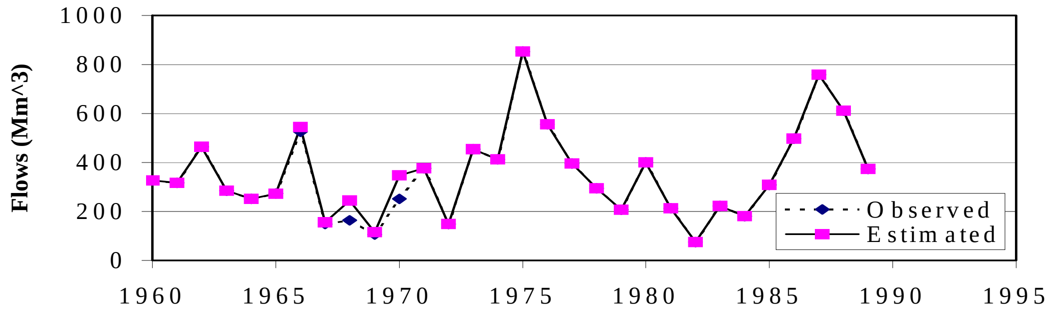


Figure 4.34a Annual mean flows at D1H009 (20 % missing from 1965) using standard BP with base gauge D1H003

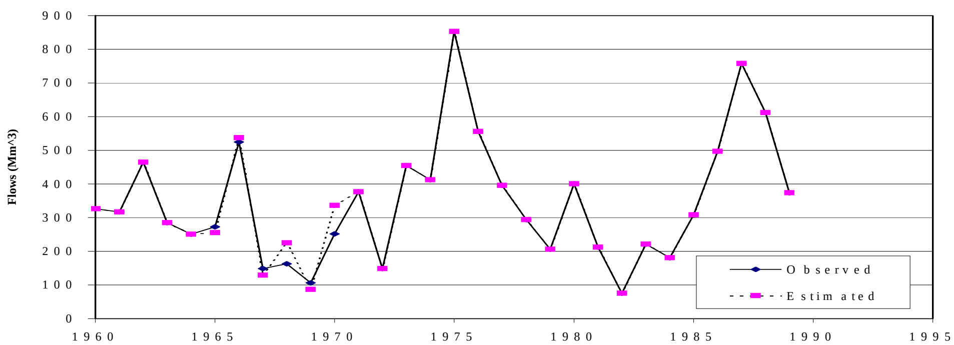


Figure 4.34b Annual mean flows at D1H009 (20 % missing from 1965) using BP (hyperbolic tangent hidden layer) with base gauge D1H003

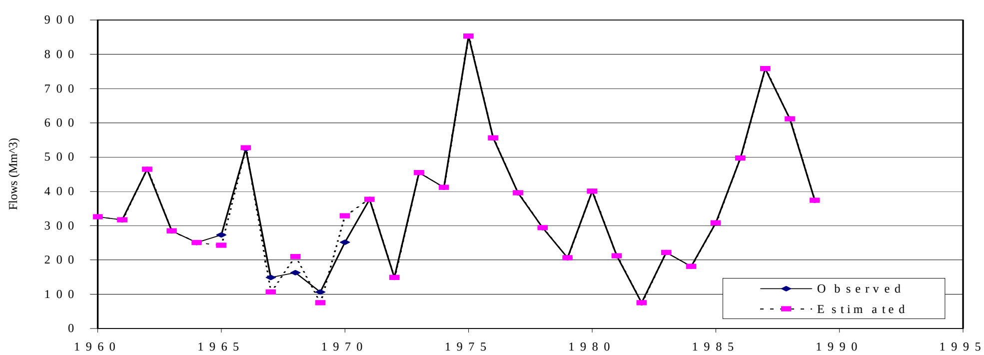


Figure 4.34c Annual mean flows at D1H009 (20 % missing from 1965) using QBP with base gauge D1H003

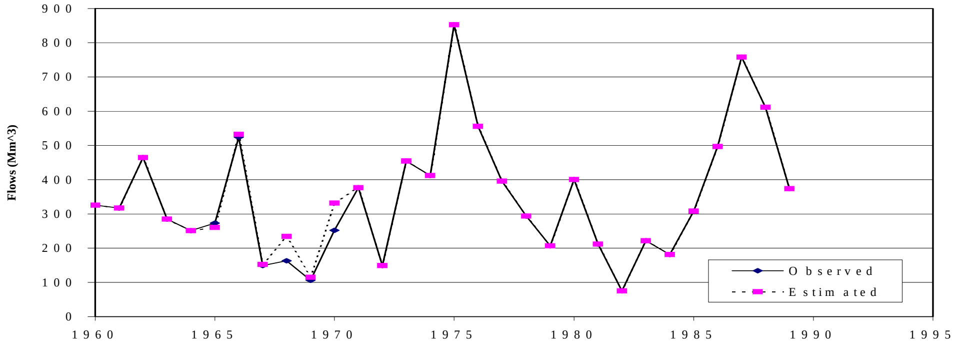


Figure 4.34d Annual mean flows at D1H009 (20 % missing from 1965) using GenerBP with base gauge D1H003

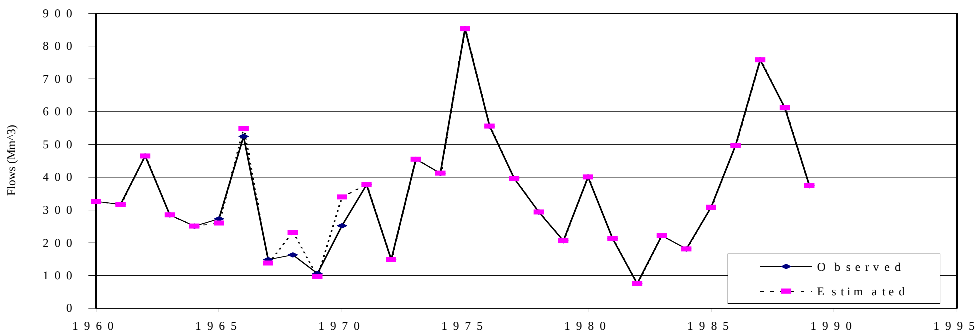


Figure 4.34e Annual mean flows at D1H009 (20 % missing from 1965) using GoldSBP, lr = 0.25 with base gauge D1H003

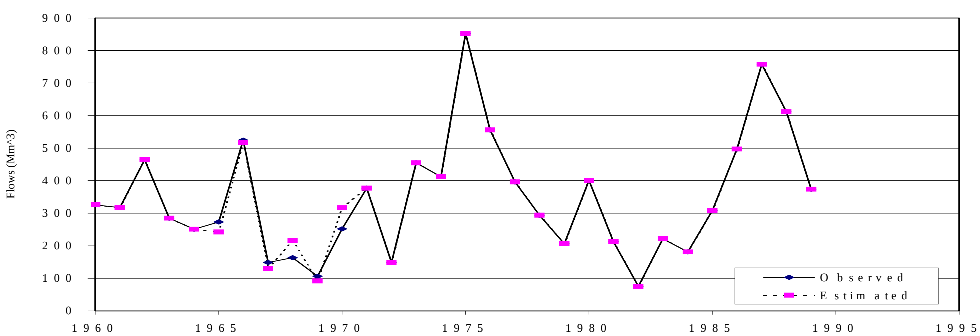


Figure 4.34f Annual mean flows at D1H009 (20 % missing from 1965) using VLR with base gauge D1H003

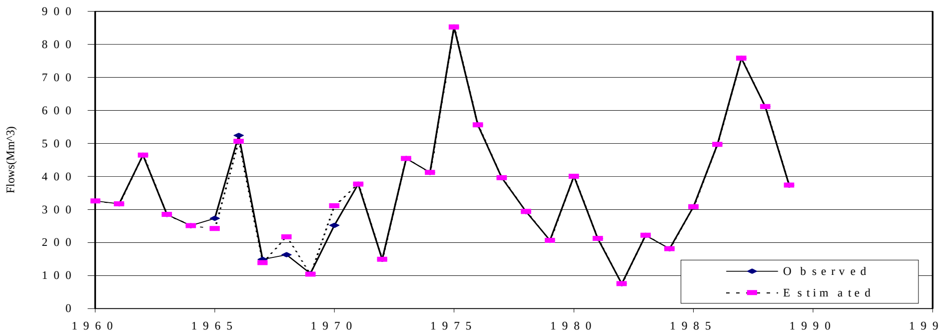


Figure 4.34g Annual mean flows at D1H009 (20 % missing from 1965) using standard EM with base gauge D1H003

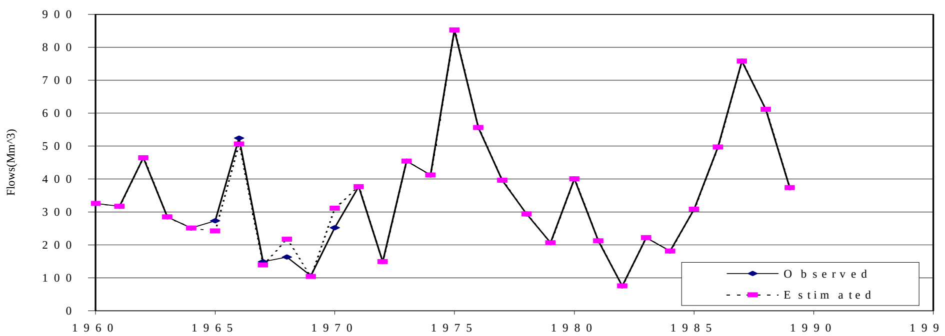


Figure 4.34h Annual average flows at D1H009 (20 % missing from 1965) using MEM1-2 with base gauge D1H003

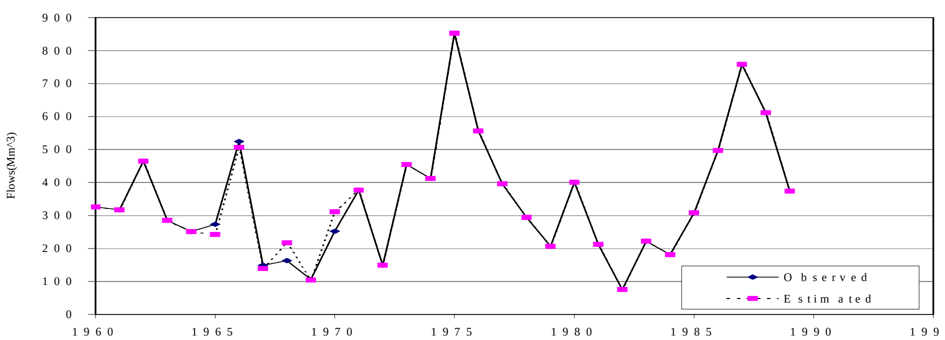


Figure 4.34i. Annual mean flows at D1H009 (20 % missing from 1965) using MEM3 with base gauge D1H003

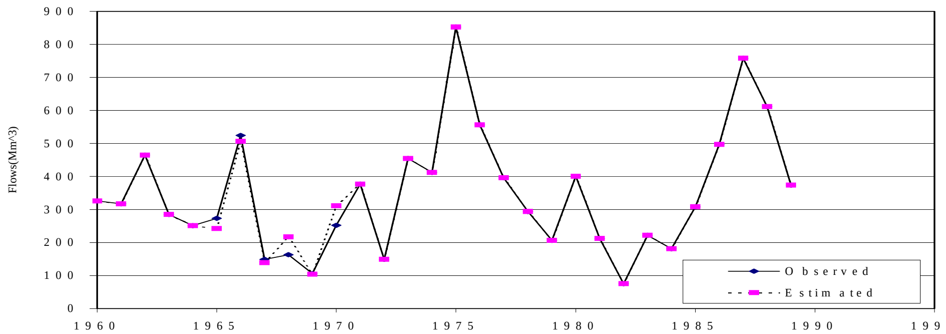


Figure 4.34j Annual mean flows at D1H009 (20 % missing from 1965) using ECM1-2 with base gauge D1H003

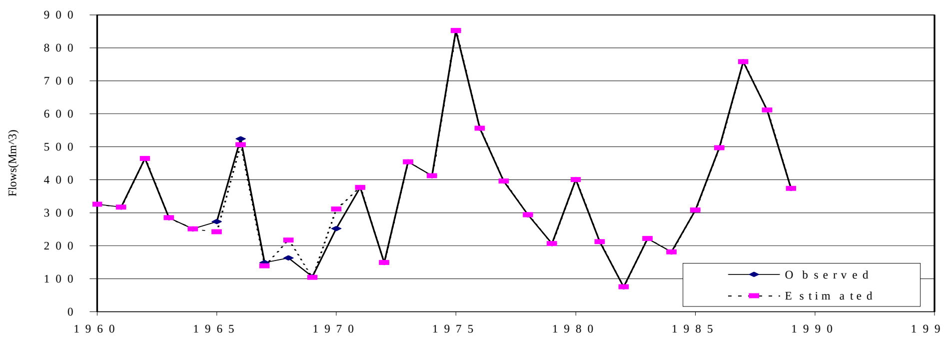


Figure 4.34k Annual average flows at D1H009 (20 % missing from 1965) using ECME1-2-3 with base gauge D1H003

4.1.4.3.2 Comparison of performance of ANN techniques at gauge D1H009 using base gauge D1H003

For a gap duration of 20 % at D1H009, the performance (in a decreasing order) of the different ANN techniques in terms of DIT (T) values was then possible and is as follows:

-QBP

-VLR

-BP (hyperbolic activation function)

-GoldSBP

- Standard BP

-GenerBP

From Table 4.7 and Figure 4.35, the values of DIT (T) range from 0.3006 (1.910 napiers) to 0.34124 (2.168 napiers). Thus, the first best technique (i.e. QBP) among ANNs has the highest entropic value of 0.34124 (2.168 napiers) for DIT (T). In other words, 34.12 % of uncertainty can be removed from the annual mean flow data at gauge D1H009 via QBP. The QBP technique is thought to have relatively higher capabilities (than other ANNs) of estimating the missing annual mean flows at D1H009.

It can be noticed that 31.02 % of information can be physically inferred about gauge D1H003 via the standard BP (hyperbolic tangent hidden layer). The standard BP is also thought to be a good estimator of missing annual mean flows at D1H009. The last best technique among ANNs (i.e. GenerBP) has the lowest entropic value of 0.3006 (1.910 napiers) for DIT (T). Thus, 30.06 % of uncertainty can be removed from the data series at gauge D1H009 via the GenerBP technique. In this case, the generalization parameter did not improve the results compared to the standard BP

On the other hand, the values for RMSEp range from 38.44 to 48.51 Mm^3. The values of RMEp range from 0.100 to 0.136. The values of EV range from 0.104 to 0.14. The GenerBP technique has the lowest values of statistical indicators.

Figure 4.36 summarizes the hydrographs of the different ANN techniques. This figure shows that the results are close. Hence, both statistical indicators and hydrographs enhance the results obtained from entropy calculations.

In general, the selected ANN techniques appear to be good estimators of the missing annual mean flow values (i.e. 20 %) at gauge station D1H009. For 20 % missing values, the statistical parameters for D1H009 such as mean, variance did not vary considerably as shown (in Table 4.7).

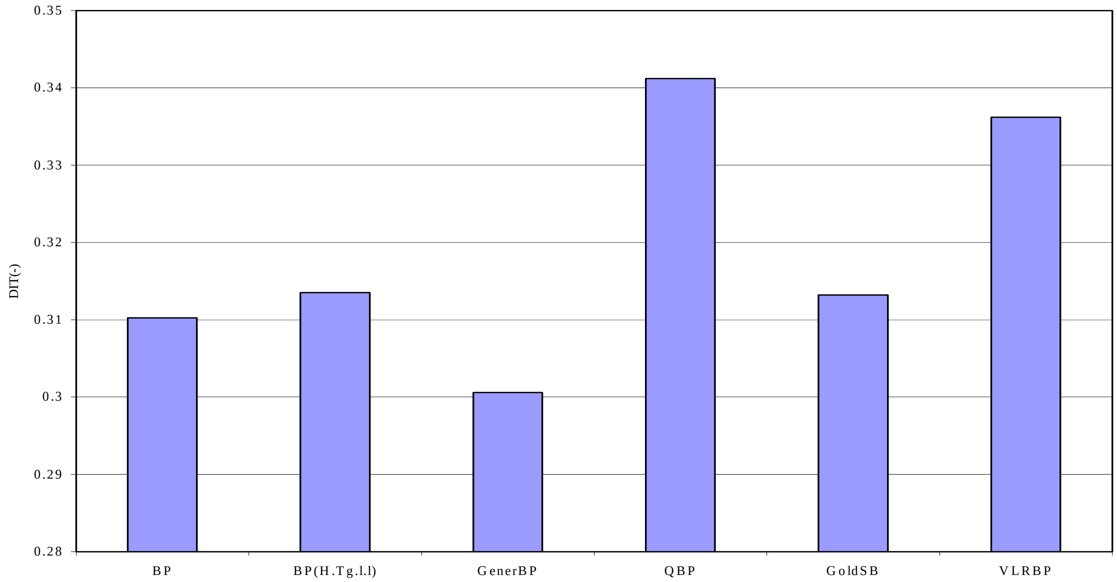


Figure 4.35 Comparison og ANN in terms of DIT (20 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

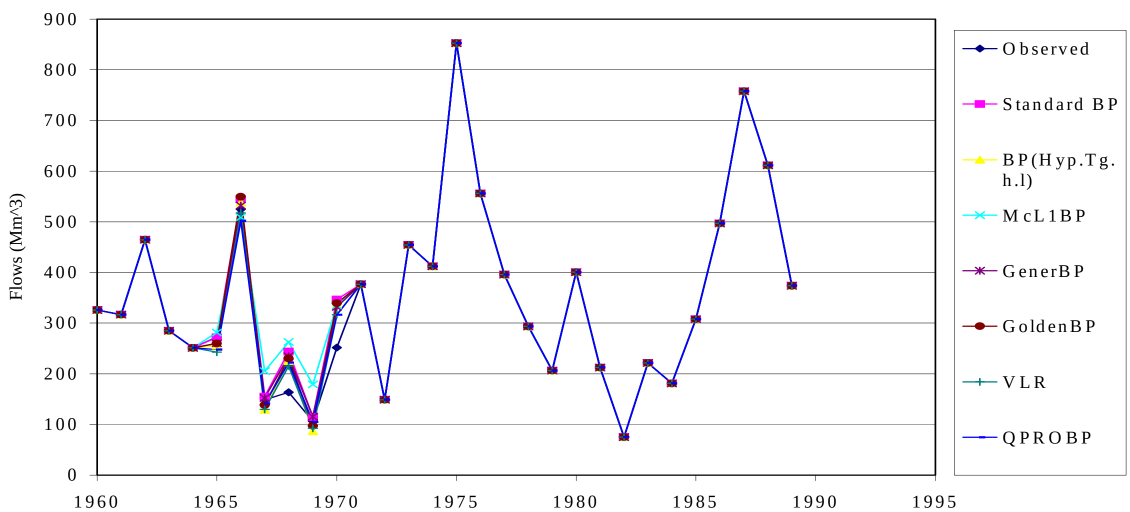


Figure 4.36 Comparison of ANN in terms of hydrographs (20 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

4.1.4.3.3 Comparison of performance of EM techniques at gauge D1H009 using base gauge D1H003

From the results shown in Table 4.7 and Figure 4.37, it can be seen that the standard EM technique and its variants have identical entropic values. Hence, the values for DIT (T) were 0.3493 (2.219 napiers) everywhere, except for the MEM3 technique, which just gave 0.3494 (2.220 napiers). This shows that the standard EM leads to the same solution as its variants in some cases; especially for cases where the maximum likelihood estimation of parameters is not complicated.

It can be concluded from entropy calculations that both the standard EM and its variants perform likely equally in estimating the missing annual mean flows (e.g. 20 %) at gauge D1H009. Like in the previous case, the use of the momentum or the learning rate in the different versions of EM would seem to be optional since they don’t have any substantial impact on the accuracy of the estimated values.

The statistical indicators (e.g. RMSEp, RMEp and EV) were almost identical for all the EM techniques. For the standard EM and its variants, the values for these statistical indicators were revealed to be almost identical.

Figure 4.38 summarizes the comparison of performance of the different EM techniques in terms of their hydrographs. From a visual observation, all the estimated hydrographs were almost identical. Hence, hydrographs and statistical indicators confirm the results obtained from entropy calculations.

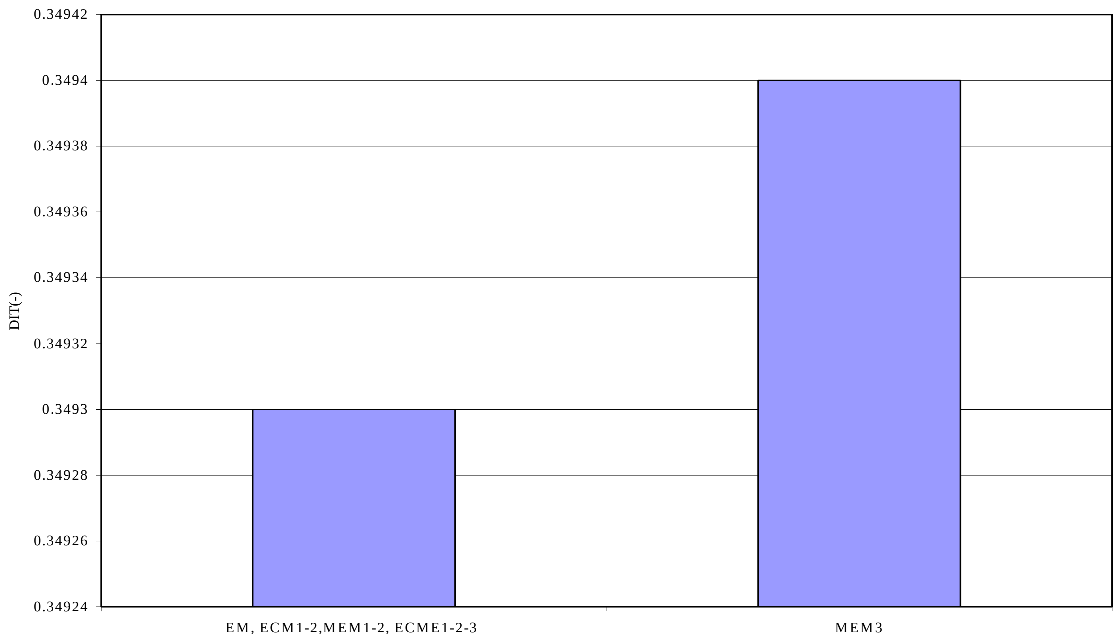


Figure 4.37 Comparison of EM in terms of DIT (20 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

4.1.4.3.4 Comparison of performance of ANN and EM techniques at gauge D1H009 using base gauge D1H003

Comparison between ANN and EM techniques is exhibited in Figure 4.38 and Figure 4.39. For visual convenience, only the first best three ANNs and all the EM techniques are presented here. The EM techniques should be ranked in the first position according to the entropy criterion. The EM techniques and its variants showed better estimation capabilities (than the ANNS) for the missing annual mean flows at D1H009. This may be due to the presence of linear relationship between gauges D1H003 and D1H009 since the EM techniques used here were developed within the linear regression context. However, the difference in entropic calculations between these ANNs and the EM techniques were small.

The graphical comparison (hydrographs) between the EM and first three ANN techniques for the gauge D1H003 is exhibited in Figure 4.39. The visual observation reveals that the hydrographs were very close. This just confirms that the difference in entropic values between ANNs and EM techniques is relatively small.

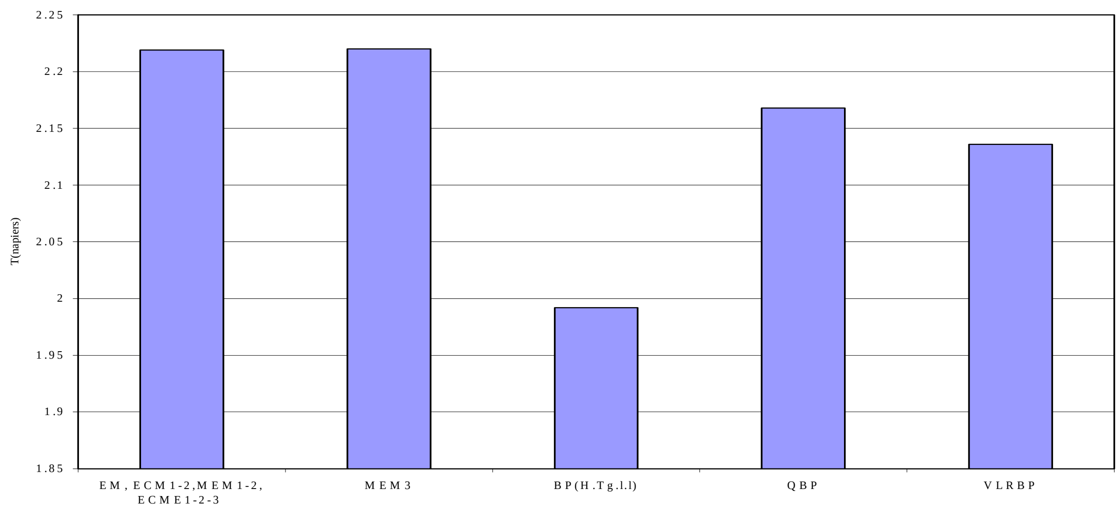


Figure 4.38 Comparison of ANN and EM in terms of T (20 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

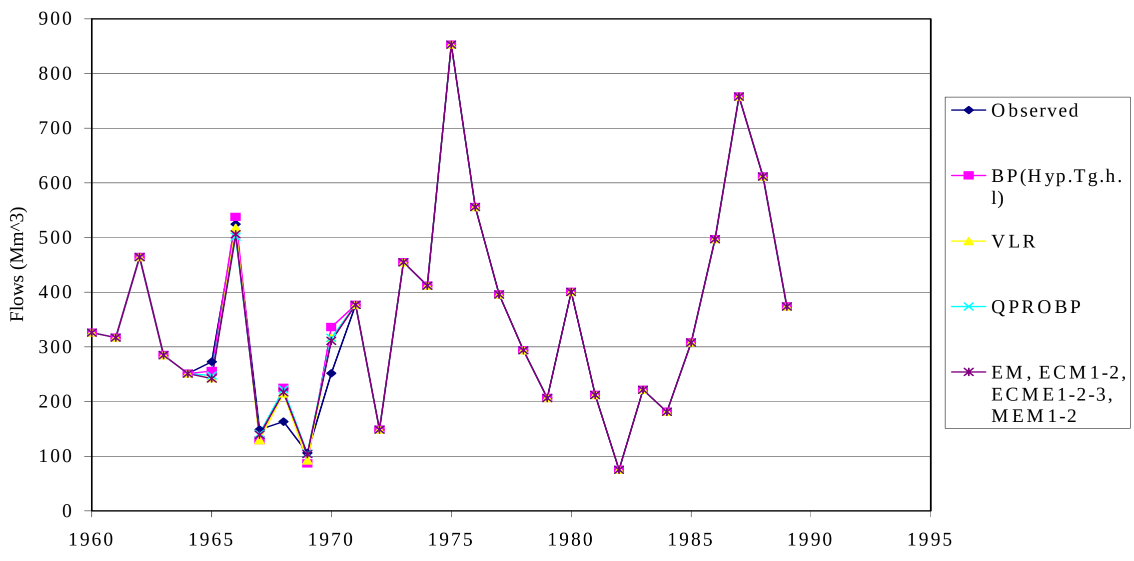


Figure 4.39 Comparison of ANN and EM in terms of hydrographs (20 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

4.1.4.4 Using 30 % of missing annual mean flows at gauge D1H009

Based on the entropy criterion, the VLR was the only ANN technique, which was retained, and all the EM techniques gave good results according to the entropy criterion (e.g. at least 30 % of uncertainty in the annual mean flows should be removed from the subject gauge D1H009). Table 4.8 shows the results obtained from entropy calculations and the statistics at gauge D1H009. Figures 4.40 (a-e) show the plot of observed and estimated hydrographs for each selected technique.

The EM techniques and its variants showed good estimation capabilities for the missing annual mean flows at D1H009 for gap duration of 30 %. This could be explained by the presence of linear relationship between gauges D1H003 and D1H009. The VLR performed just marginally better than the EM techniques as shown in Figure 4.41. The VLR could be thought of having slightly a higher capability of mapping the non-linear characteristics in the missing values through the use of a sigmoid hidden layer. Thus, 30.36 % and 30.05 % of uncertainty can be removed from the annual mean flow data at D1H009 via VLR and EM techniques respectively. In fact, these values are close. The values for RMSEp were 47.01 and 50.08 Mm^3 respectively for the VLR and EM techniques. The values for RMEp were 0.1507 and 0.1180 for the VLR and EM techniques respectively. The values of EV were 0.0525 and 0.0897 for the VLR and EM techniques respectively.

The notion of DIT could enable streamflow data infilling techniques to be compared for different catchment areas. In general, for a given gap size, the overall value of DIT for the station-pair D1H003-D1H009 is slightly higher than the one for the station-pair D1H009-D1H003. Thus, D1H003 is slightly better predictor for the estimation process of missing annual mean flows at D1H009 than when gauge D1H009 is used to fill in missing data at D1H003. Nonetheless, D1H003 could be used to interpolate (fill in) the missing annual mean flows at gauge D1H009 (and vice-versa).

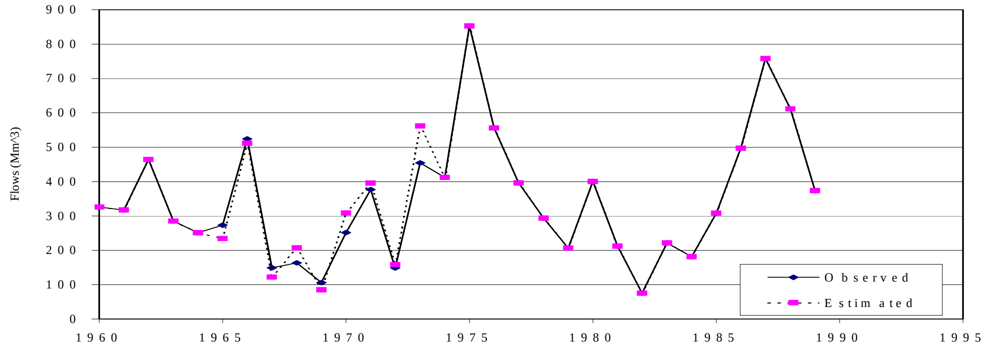


Figure 4.40a Annual mean flows at D1H009 (30 % missing from 1965) using VLR with base gauge D1H003

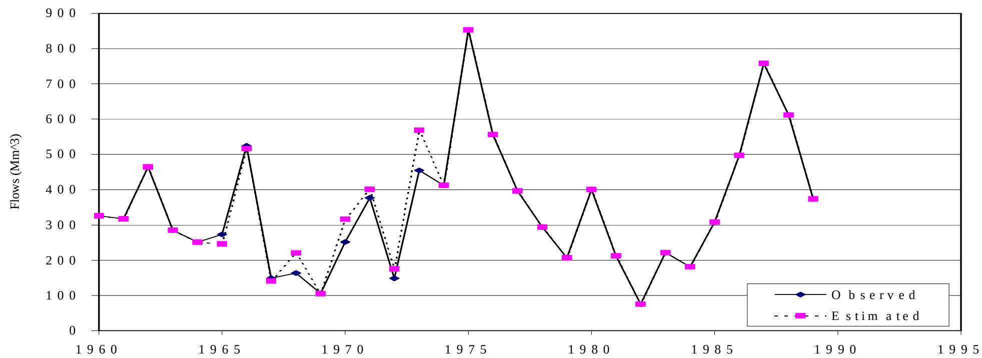


Figure 4.40b Annual mean flows at D1H009 (30 % missing from 1965) using standard EM with base gauge D1H003

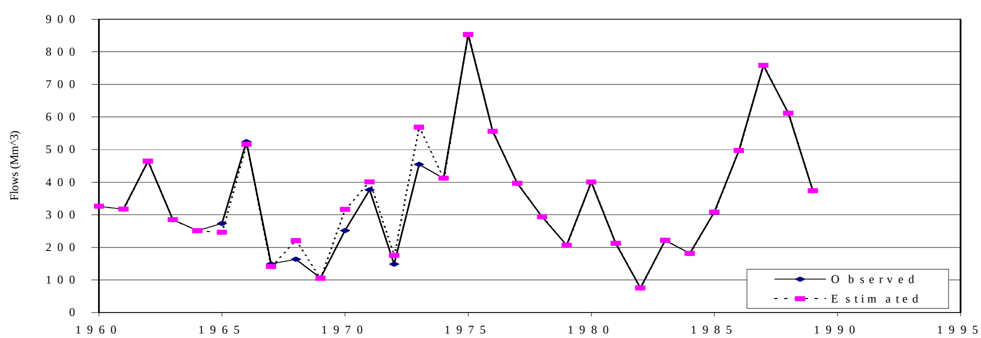


Figure 4.40c Annual mean flows at D1H009 (30 % missing from 1965) using ECM1-2 with base gauge D1H003

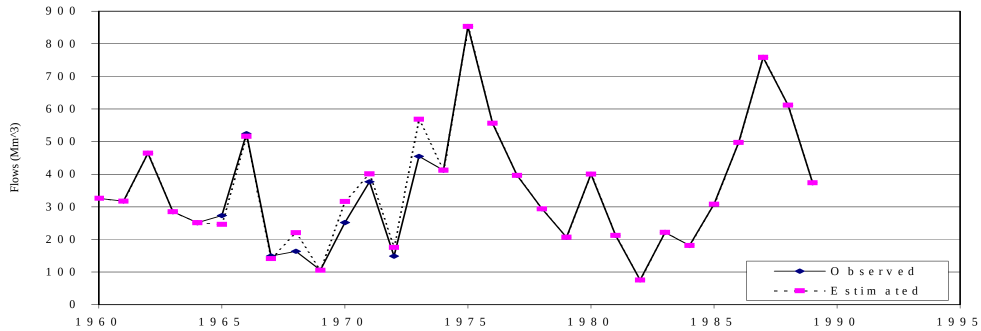


Figure 4.40d Annual mean flows at D1H009 (30 % missing from 1965) using MEM1-2-3 with base gauge D1H003

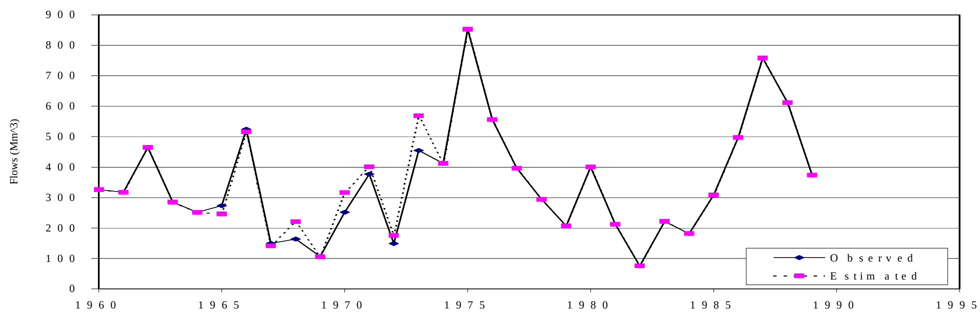


Figure 4.40e Annual mean flows at D1H009 (30 % missing from 1965) using ECME1-2-3 with base gauge D1H003

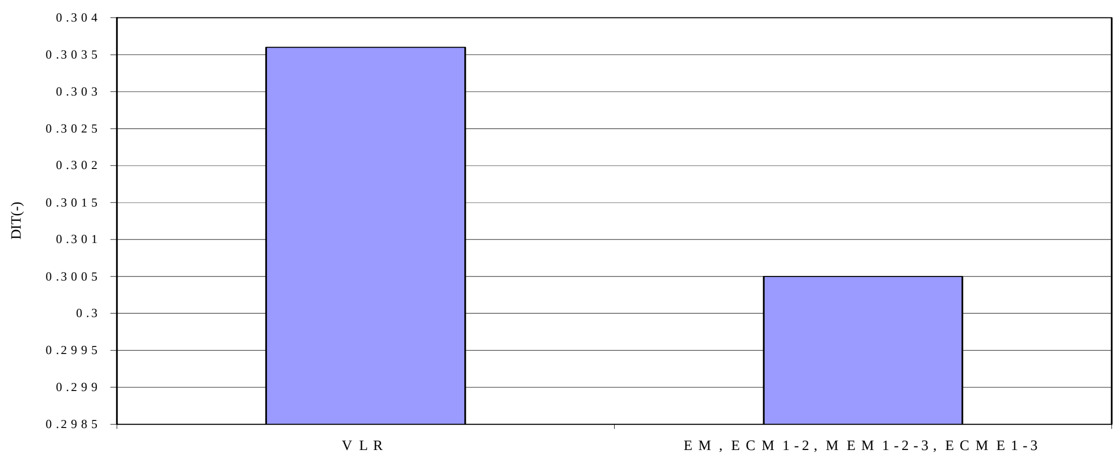


Figure 4.41 Comparison of ANN and EM in terms of DIT (30 % missing annual mean flows from 1965 at D1H009) using base gauge D1H003

Table 4.8 Entropy calculations and statistics at the target gauge D1H009 (30 % missing annual mean flows) using the base gauge D1H003

	Description	T (napiers)	DIT	RMSEp (Mm^3)	RMEp (n = 10)	EV	Ratio Variance	Ratio Mean
ANN Techniques	VLR (sigmoid hidden layer)	1.929	0.3036	47.008	0.1507	0.0525	0.96618	0.9872
EM Techniques	Standard EM	1.9091	0.3005	50.0103	0.1186	0.0897	0.993	0.9774
	MEM1, 2 (u = 0.95, lr = 0.001)	1.900	0.3000	50.638	0.1180	0.0929	0.9934	0.9766
	MEM, Option 3	1.9092	0.3005	50.077	0.1191	0.0894	0.992	0.9775
	ECM1-2	1.909	0.3005	50.0103	0.1186	0.0897	0.993	0.9774
	ECME1, 2 (k = 3), 3 (u = 0.90)	1.909	0.3005	50.0103	0.1186	0.0897	0.993	0.9774

4.2 SUMMARY

The entropy concept was shown to be a versatile through DIT. Firstly, it enables to know the information content of the streamflow gauges. Secondly, it also enables the selection of predicted/predictor gauge and it also used to assess technique performance when applied to annual mean flow interpolation (infilling).

It was shown that the gauges in the station-pairs D1H003-D1H009 and D1H009-D1H003 could infer mutually the information (contained in annual mean flow series) about one another. In other words, when one gauge could be considered as predicted (subject) station, the other one was predictor (base) station and vice-versa. The gauge D1H006 was selected as potential base station for both gauges D1H003 and D1H009. However, D1H006 was not finally considered for filling in the missing annual mean flows at gauges D1H003 and D1H009. The reason is that the entropy criterion (e.g. 30 % of uncertainty to be removed from the subject station) was not satisfied for the different streamflow data infilling techniques.

The results for gap duration of 6.7 %, 13.4 % and 20 % at the subject gauge D1H003 were the only ones to be discussed as the different techniques could perform well. Therefore, the rest of the gap duration (i.e. 30 %) was not discussed at this specific gauge. Nonetheless, the entropic values for this gap duration were used just to investigate the relationship between gap duration and accuracy of estimated values for the different techniques. Taking D1H009 as target gauge, the whole range of gap duration (e.g. from 6.7 % to 30 %) was discussed as the results for the performance of the different techniques were shown to be satisfactory.

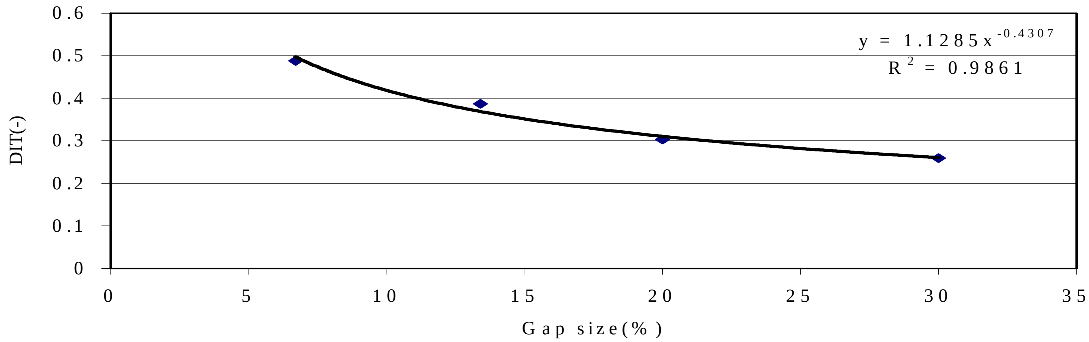
An overall assessment revealed that the standard BP and standard EM as well as their respective variants showed good capabilities of estimating the missing values at the target D1H009, when considering the base gauge D1H003 (and vice-versa). Nonetheless, the results from the EM techniques were almost identical. This could be due to the form of the maximum likelihood equations considered so far. (Refer to equations considered in the whole section 3.3.6.2 of Chapter 3). These equations were taken within the context of linear regression (e.g. without random effects on the error term).

The notion of DIT could enable streamflow infilling techniques to be also compared for different catchment areas. For a given gap size, it was shown that D1H003 was in general slightly better predictor for the estimation process of missing annual mean flows at D1H009 than when gauge D1H009 was used to fill in missing data at D1H003. Nonetheless, D1H003 could be used to fill in the missing annual mean flows at gauge D1H009 and vice-versa.

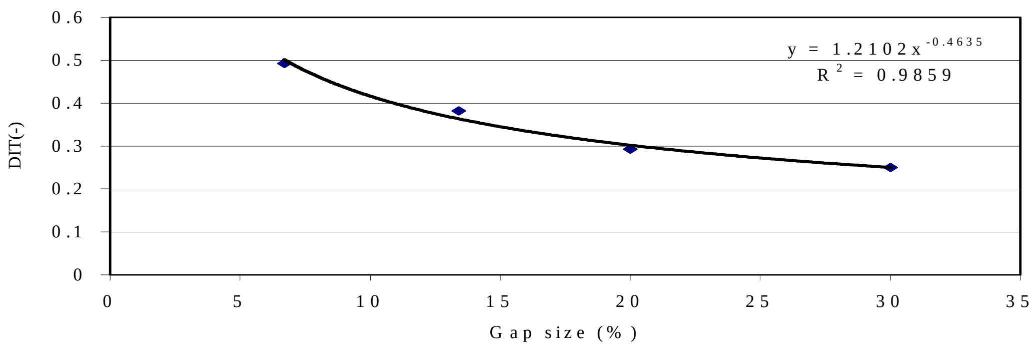
It was also noticed that the DIT between observed and estimated values generally decreases when the proportion of missing annual mean flows increases. In other words, the proportion of information physically transferred by the knowledge of the estimated series into the process to make the observed annual mean flow series better defined will decrease as the gap duration increases at the subject gauge.

Using D1H009 as subject station a decay power function could describe better (than exponential or linear) the relationship between the gap duration (e.g. range from 6.7 % to 30 %) and the technique accuracy in terms of the DIT. Considering D1H003 as subject gauge, a similar relationship could also be established. Figures 4.42 (a-d) and 4.43 (a-d) are just an illustration for the relationship mentioned here. Thus, for a given technique, it was possible to find approximately the expected accuracy of the estimated values at the subject gauge when the gap size is known (e.g. between 6.7 % and 30 %). Increasing the number of data points (e.g. up to seven values of gap sizes at the target gauge) did not sensitively affect the above-mentioned relationship. It was also noticed that an earlier start (e.g. at 1963) or later start (e.g. 1970) for the gaps created on the records of the subject gauge did not have any substantial impact on the accuracy of the estimated values.

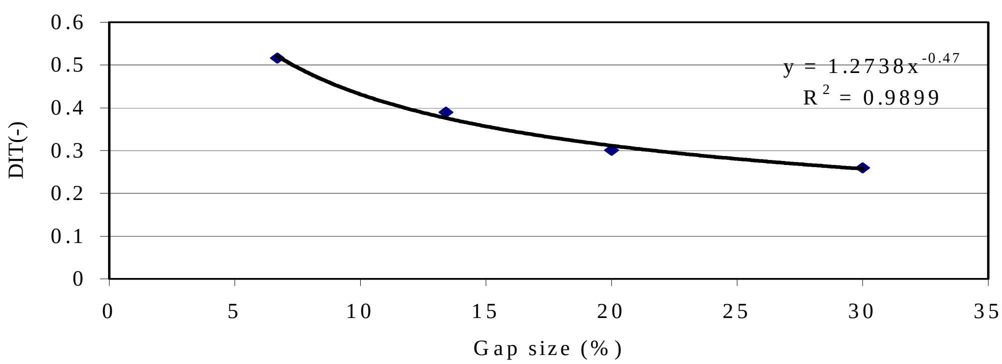
Generally, it was noticed that the performance of the different techniques depends on the gap duration at the target gauge and the station-pair involved in the estimation of missing annual mean flows.



(a)

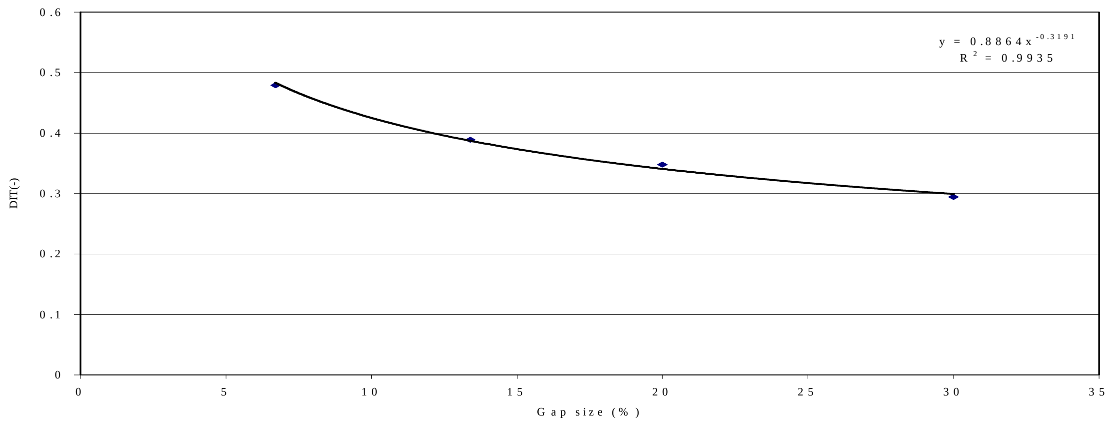


(b)



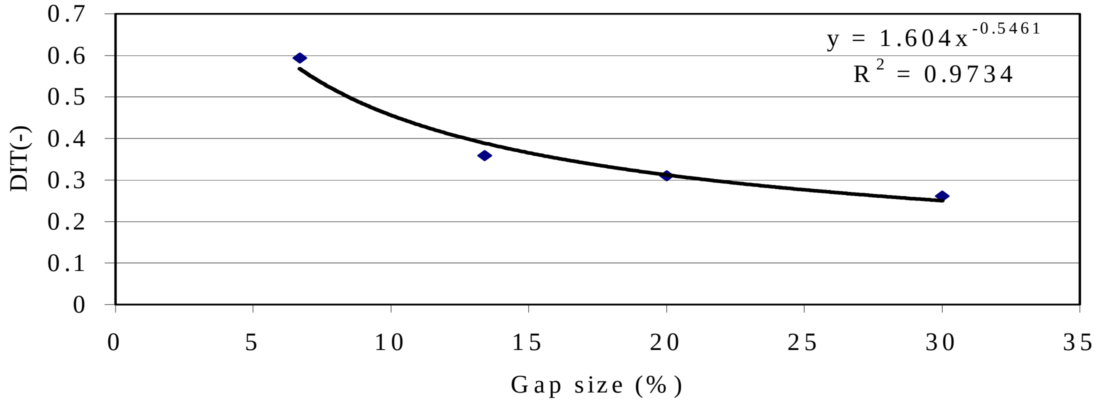
(c)

Figures 4.42 (a-c) DIT versus gap size for annual mean flows at D1H003: (a) BP, (b) McL1BP, (c) GenerBP

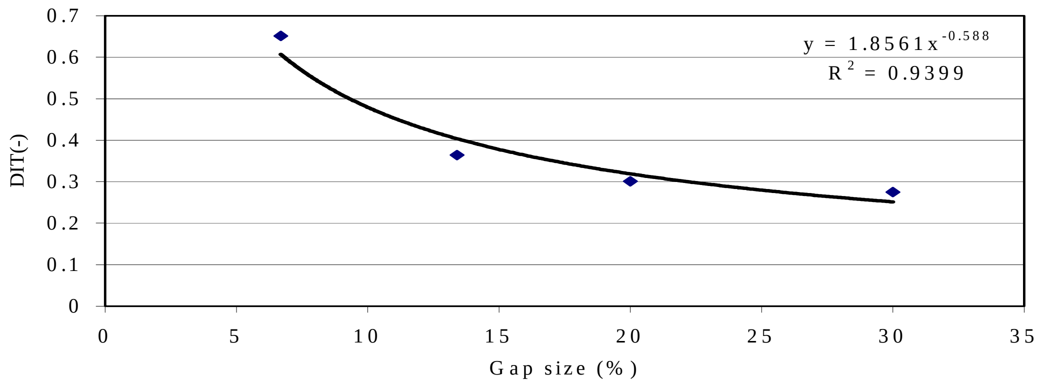


(d)

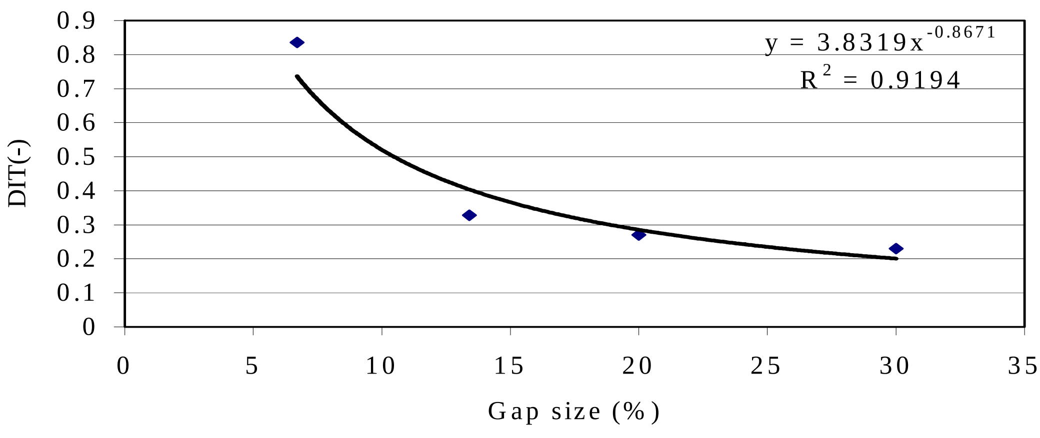
Figure 4.42d DIT versus gap size for annual mean flows at D1H003: (d) EM



(a)



(b)



(c)

Figures 4.43 (a-c) DIT versus gap size for annual mean flows at D1H009: (a) BP, (b) GenerBP, (c) McL1BP

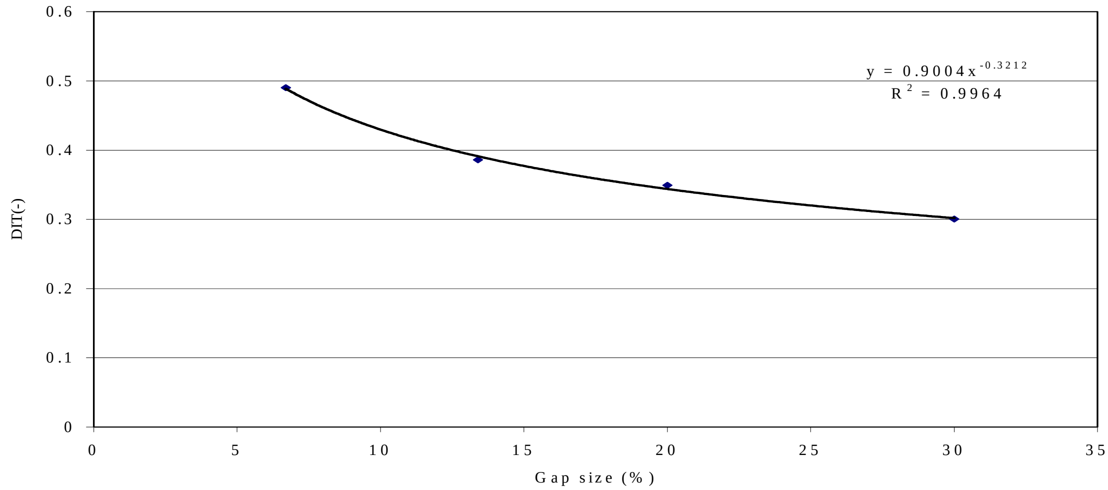


Figure 4.43d DIT versus gap size for annual mean flows at D1H009: (d) EM