CHAPTER 6

MODEL TESTING ON RAINFALL: II. ANNUAL TOTALS

6.1 RESULTS AND DISCUSSION

6.1.0 Preamble

The methodology was also tested on the three rainfall gauging stations, i.e. 0228495, 0228170 and 0228458 as depicted in Figure A3 of Appendix A and in section 3.4 of Chapter 3. The regime considered for the time-series was the annual total rainfall (1924-1989). These data are useful for design-oriented studies such as reservoirs, bridges, etc.

As said so far, 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 rainfall data at each gauge was tested by computing the value of the first order serial correlation of the annual rainfall (totals) at each gauge. The computed values of the first order serial correlation were found to be –0.158, -0.1659 and –0.2360 for stations 0228170, 0228458 and 022495 respectively. It was concluded that the independence assumption could not be rejected statistically at 95 % confidence interval. By just assuming a threshold value of 0.5 for the correlation coefficient (below which the series is considered to be insignificantly correlated) as mentioned by Elshorbagy et al. (2000b), it can be concluded that the annual total rainfall at these rainfall stations are assumed to be independent. The second assumption, normality, is not always met, as real data is rarely normal, especially for a set of annual rainfall (totals). One way of meeting this assumption was to transform strategically the data to follow approximately the normal distribution. It was found in this study the Box-Cox family of transformations to be suitable. The annual rainfall (totals) series for the three gauges were approximated normally distributed with a square root transformation. Makhuvha et al. (1997a) found this transformation suitable in his study on rainfall data patching for South African sites.

6.1.1 Selection of base/subject station

The entropy computations for the respective rainfall gauges are shown in Tables 6.1 and the values of the cross-correlation of lag zero for the site pairs, i.e. 0228170-0228458, 0228170-0228495 and 0228458-0228495, were found to be 0.874, 0.925 and 0.888 respectively.

Table 6. 1a Marginal entropy for rainfall gauges

	0228170	0228458	0228495
Marginal entropy (napiers)	3.7968	3.8075	3.7992

Table 6.1b Informational matrix (e.g. T) of annual total rainfall for different site-pairs.

	0228170	0228458	0228495
0228170	1	0.7223	0.9704
0228458	0.7223	1	0.7753
0228495	0.9704	0.7753	1

Table 6.1c Informational matrix (e.g. DIT) of annual total rainfall for different site-pairs.

	0228170	0228458	0228495
0228170	1	0.1902	0.2556
0228458	0.1902	1	0.2036
0228495	0.2556	0.2036	1

The threshold value (e.g. Threshold1) for DIT was just assumed to be 0.20 (above which the first station in a given site-pair is considered as potential information predictor of the other station). Thus, the following station pairs were selected 0228495 (base)-0228170 (subject), 0228495 (base)-0228458 (subject), 0228170 (base)-0228495 (subject) and 0228458 (base)-0228495 (subject). From the selected gauge pairs, it could be said that the potential predictor (base) rainfall station for the predicted (subject) rainfall station 0228170 is only 0228495. Similarly, 0228170 and 0228458 are potential base stations for station 0228495. For the rainfall gauge 0228458, the only selected potential base station is 0228495. The association between gauges 0228458 and 0228170 was weak since the entropy criterion was not satisfied. Hence, the station pair 0228458-0228170 could not be selected.

6.1.2 Training and assessment of rainfall data infilling techniques

Since the rainfall data exhibited no gaps, different proportions of missing values were created at the target (subject) station for each selected pair, starting from year 1935. The different values for gap duration of 7.6 %, 13.6, 19.7% and 30.3 % were considered and then infilled (interpolated) by the selected potential base station.

For station pairs 0228458 (base)-0228495 (subject) and 0228495 (base)-0228458 (subject), the only results discussed are for gap duration of 7.6 % and 13.7 % since the rainfall data infilling techniques satisfied the entropy criterion. Nonetheless, the rest of gap duration (e.g. 20 % and 30.3 %) could be also considered to investigate the relationship between gap duration and accuracy of the estimated annual total rainfall values.

For station pairs 0228495 (base)-0228170 (subject) and 0228170 (base)-0228495 (subject), the results were not discussed but only presented in Appendix C as these results were similar to those obtained from other pairs. The entropy criterion for the above-mentioned station pairs was successfully fulfilled for the whole range of gap duration, i.e. between 7.6 % and 30.3 %.

A three-layered ANN was used as in Zealand et al. (1999). The learning rate was set to the range between 0.15 and 0.45 for quite reasonable results, unless stated otherwise. The number of nodes in the hidden layer was set to the range between 2 and 6 for reasonable results. The ANN techniques were then trained on the concurrent parts of the observed values (e.g. annual total rainfall) and the parameters obtained (weights, etc.) were subsequently used to estimate the missing values. This is done similarly to Kuligowsky and Barros (1998) who used ANNs to estimate the rainfall missing data at the target gauge using rainfall data from the nearby gauges. Training comprises of the annual total rainfall 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 and hyperbolic tangent) is defined and the network is assigned values to the interconnected weights. A sigmoid function could be 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 remain intact throughout, only the missing values were repeatedly estimated until convergence.

The different techniques were first assessed among themselves and in turn compared with the other stream of techniques. Based on the results of these applications, the proposed and existing techniques were assessed predominantly through entropy criterion. The graphical and statistical criteria were also done to verify the results from entropy criterion. The graphical representation is done 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.

6.1.3 Infilling annual total rainfall at gauge 0228495 with base gauge 0228458

6.1.3.1 Using 7.6 % of missing annual total rainfall at gauge 0228495

6.1.3.1.1 Selection of ANN and EM techniques for rainfall infilling at 0228495 using base station 0228458

The results of performance for the different rainfall data infilling techniques at 0228495 are summarized in Table 6.2 and Figures 6.1 (a-k). Rainfall station 0228458 is the base station for interpolating (filling in) the missing annual total rainfall at the subject station 0228495. Table 6.2 summarizes entropy calculations and statistics at 0228495 using ANN and EM techniques.

The techniques presented in Table 6.2 are thought to have performed well, i.e. satisfying the entropy criterion. At least 30 % of uncertainty in the annual total rainfall data at 0228495 can be removed by the application of these techniques. The values for DIT (T) range from 0.3460 (2.0716 napiers) to 0.3804 (2.278 napiers). This range falls under the entropy criterion. It could be said that the standard BP performed as well as its variants. The same could be noticed for the standard EM and its variants.

The statistical criteria (Table 6.2) and the graphical plot (Figures 6.1 (a-k)) for each respective technique were just made to crosscheck results found from entropy calculations. The statistical indicators viz. RMSEp and RMEp range respectively from 41.842 to 52.562 mm of rainfall and 0.0783 to 0.0985. The values for RMEp, which are dimensionless, are relatively small (e.g. close to zero). This can eventually confirm the results obtained from entropy criterion. 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. This also confirms the results from entropy calculations.

Thus, the above, the selected techniques are thought to be good estimators of the missing annual total rainfall at gauging station 0228495 using the base station 0228458.

Table 6.2 Entropy calculations and statistics at the target gauge 0228495 (7.6 % missing annual total rainfall) using the base gauge 0228458

	Description	T (napiers)	DIT	RMSEp (mm)	RMEp	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	2.0716	0.3460	52.562	0.0985	0.0854	0.9941	0.993
	McL1BP (sigmoid hidden layer)	2.179	0.3640	46.714	0.0849	0.0709	0.9992	0.9946
	GenerBP (sigmoid hidden layer), s = 5	2.142	0.3577	48.529	0.0937	0.0680	0.9959	0.9949
	QBP (sigmoid hidden layer), acc = 0.015, lr = 0.85, Weight Cond	2.278	0.3804	41.842	0.0783	0.0580	1.00196	0.9957
	GoldSBP (sigmoid hidden layer), lr = 0.25	2.104	0.3514	50.661	0.0966	0.0774	0.9949	0.9941
	VLR (sigmoid hidden layer)	2.189	0.3656	45.964	0.0911	0.0575	0.9970	0.9957
EM Techniques	Standard EM	2.202	0.3677	45.35	0.0893	0.0554	0.9975	0.9959
	MEM1	2.202	0.3677	45.35	0.0893	0.0554	0.9975	0.9959
	MEM2-3	2.201	0.3676	45.383	0.0894	0.0552	0.9974	0.996
	ECM1-2	2.202	0.3677	45.35	0.0893	0.0554	0.9975	0.9959
	ECME1-2 (k = 2), 3	2.202	0.3677	45.35	0.0893	0.0554	0.9975	0.9959

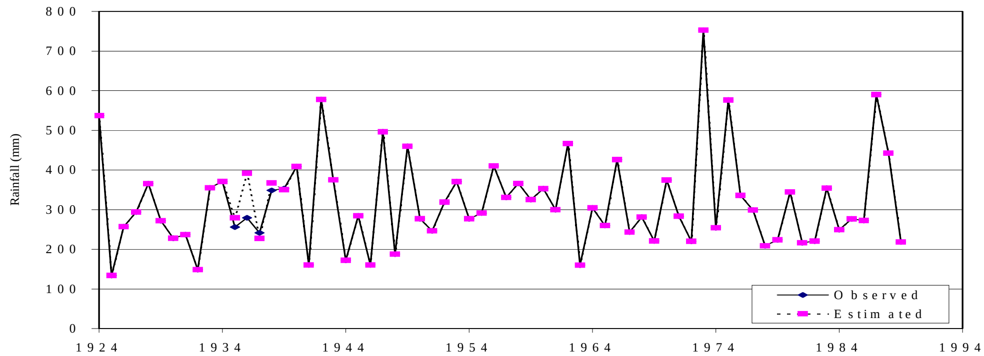


Figure 6.1a Annual total rainfall at 0228495 (7.6 % missing from 1935) using standard BP with base gauge 0228458

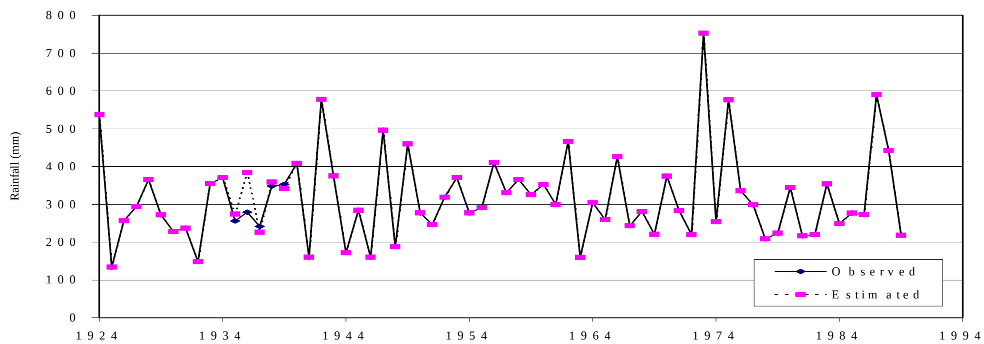


Figure 6.1b Annual total rainfall at 0228495 (7.6 % missing from 1935) using GenerBP with base gauge 0228458

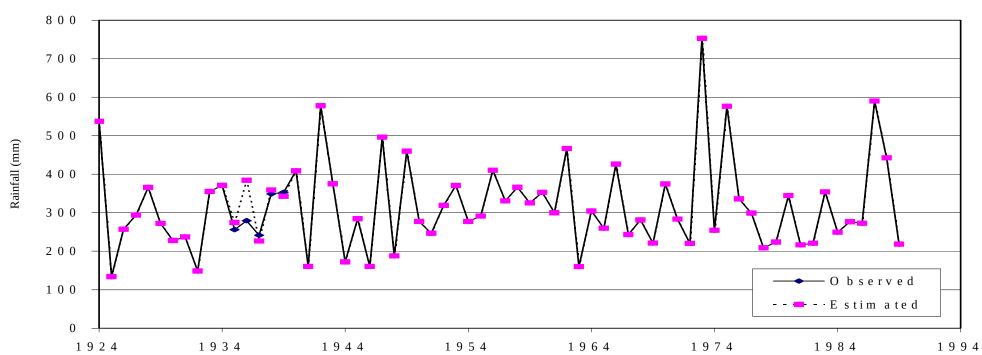


Figure 6.1c Annual total rainfall at 0228495 (7.6 % missing from 1935) using McL1BP with base gauge 0228458

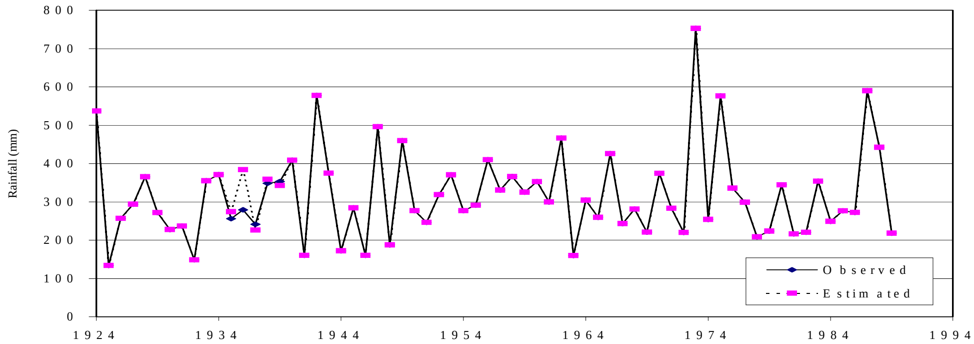


Figure 6.1d Annual total rainfall at 0228495 (7.6 % missing from 1935) using GoldSBP with base gauge 0228458

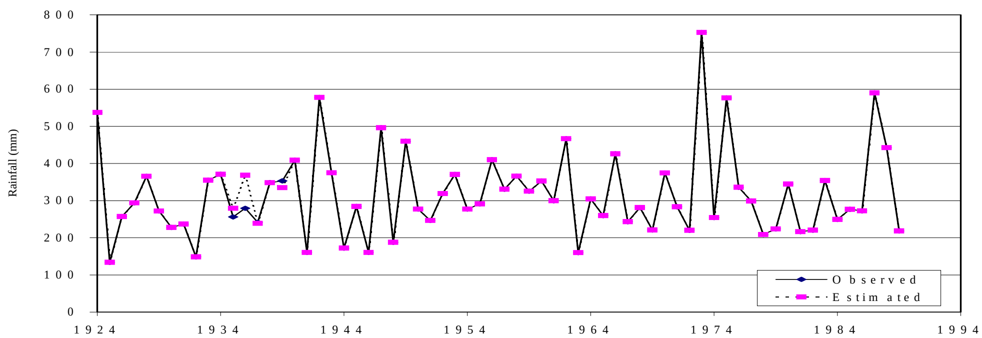


Figure 6.1e Annual total rainfall at 0228495 (7.6 % missing from 1935) using QBP with base gauge 0228458

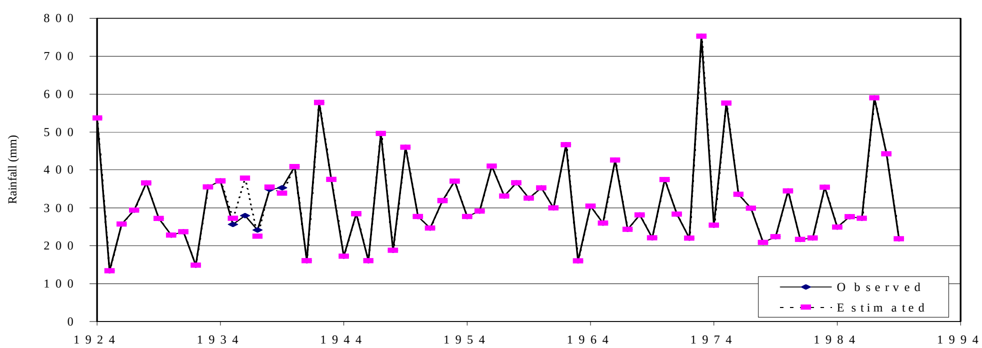


Figure 6.1f Annual total rainfall at 0228495 (7.6 % missing from 1935) using VLRBP with base gauge 0228458

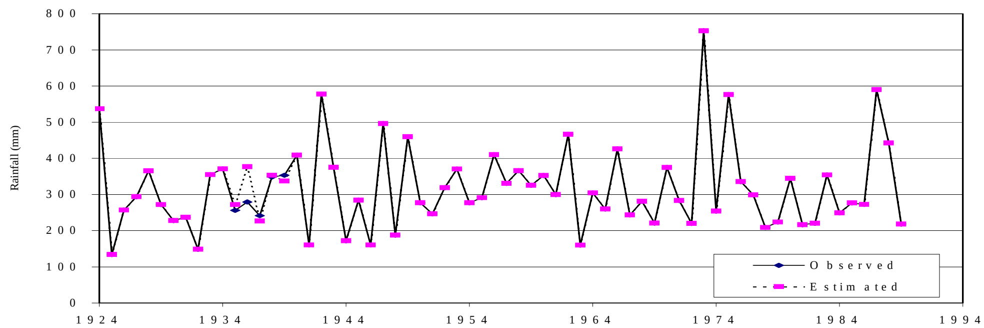


Figure 6.1g Annual total rainfall at 0228495 (7.6 % missing from 1935) using standard EM with base gauge 0228458

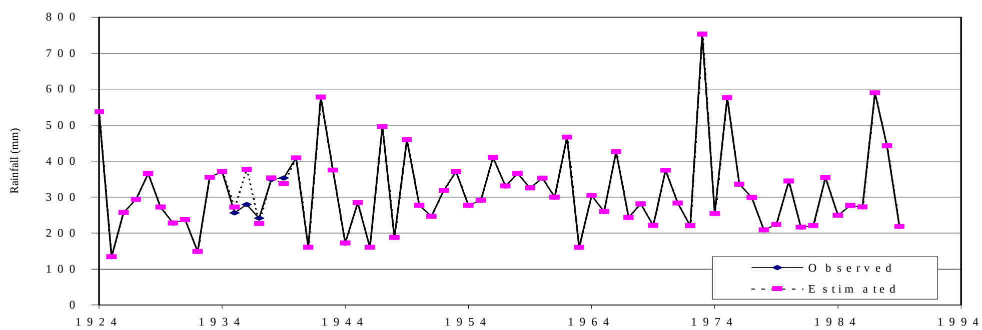


Figure 6.1h Annual total rainfall at 0228495 (7.6 % missing from 1935) using MEM1 with base gauge 0228458

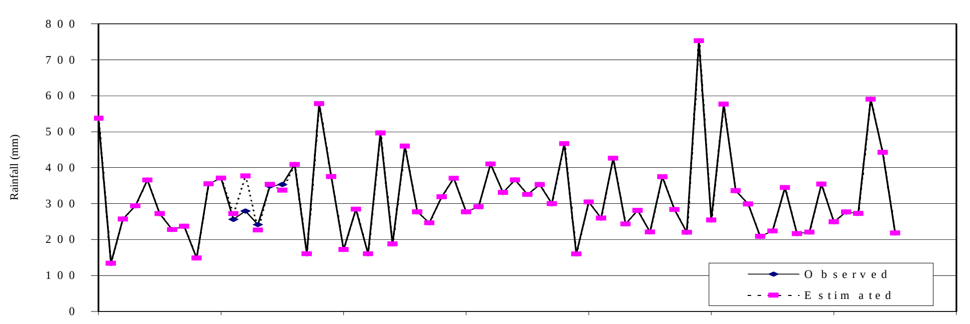


Figure 6.1i Annual total rainfall at 0228495 (7.6 % missing from 1935) using MEM2-3 with base gauge 0228458

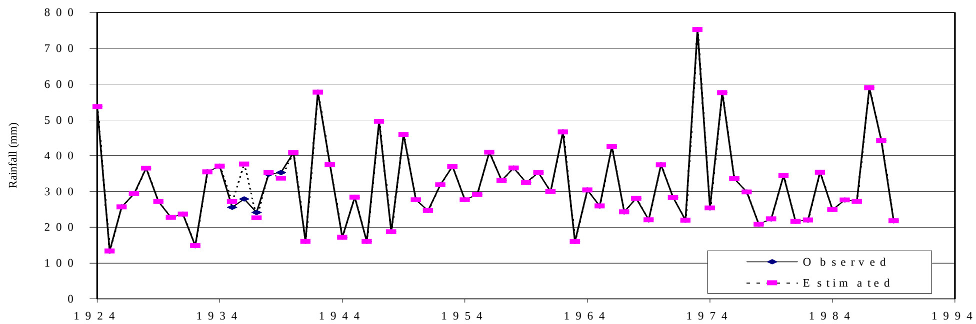


Figure 6.1j Annual rainfall at 0228495 (7.6 % missing from 1935) using ECM1-2 with base gauge 0228458

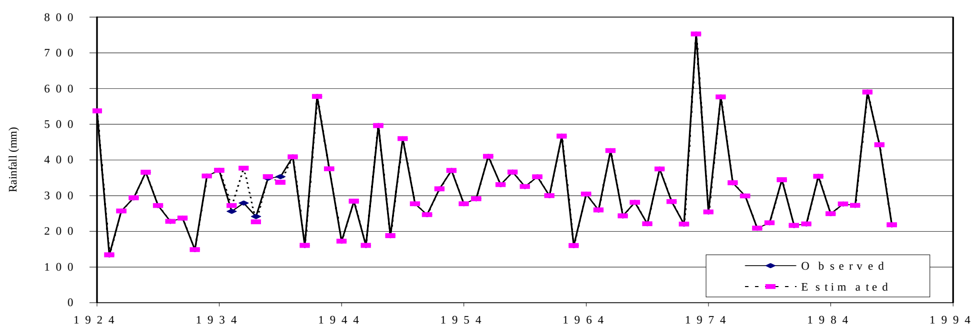


Figure 6.1k Annual total rainfall at 0228495 (7.6 % missing from 1935) using ECME1-2-3 with base gauge 0228458

6.1.3.1.2 Comparison of performance of ANN techniques at the subject gauge 0228495 using base gauge 0228458

Based on the entropic values for the selected techniques, model performance assessment was then made for the estimation of missing annual total rainfall at gauge 0228495. Considering 0228458 as base station and for gap duration of 7.6 % at 0228495, the performance (in a decreasing order) of the different ANN techniques in terms of DIT (T) is as follows:

-QBP, acc = 0.015, lr = 0.85 and Weight cond.

-VLR

-McL1 BP

-GenerBP, lr = 0.25, s = 5

-GoldSBP

-Standard BP

From Table 6.2 and Figure 6.2 (a-b), the values of DIT (T) range from 0.346 (2.0716 napiers) to 0.3804 (2.278 napiers). Thus, the first best model (i.e. QBP) among the selected ANN techniques has the highest entropic value of 0.3804 (2.278 napiers) for DIT (T). In other words, 38.04 % of uncertainty can be removed from the annual total rainfall data at station 0228495 via the QBP technique. This can also be seen as the proportion of information transferred by the knowledge of the estimated series (through QBP) into the process to make the annual total rainfall data at 0228495 better defined. The QBP technique is thought to be relatively more capable (than the rest of ANNs) of estimating the missing annual total rainfall at 0228495.

The variants of the standand BP were shown to be relatively superior to the standard BP according to entropy criterion. The standard BP technique is the last best among the ANNs and has the lowest entropic value of 0.346 (2.0716 napiers) for DIT (T). In other words, 34.6 % of information can be physically inferred about gauge 0228495 via the standard BP. Nonetheless, this technique is thought, in this study, to be a good estimator of missing annual total rainfall data at 0228495 as long as it satisfies the entropy criterion.

On the other hand, the values for RMSEp range from 41.842 mm to 52.562 mm of rainfall. These two values correspond to the QBP and standard BP respectively. The values of RMEp ranged 0.0783 to 0.0985. The QBP technique has the lowest values of the statistical indicators while the standard BP has the highest values for the same statistical indicators. These results just confirm the conclusion drawn from the entropy calculations.

Figure 6.3 summarizes the hydrographs of the different ANN techniques. This figure shows that the hydrographs are close.

In general these techniques appear to be good estimators of the missing annual total rainfall (7.6 %) at gauging station 0228495. For 7.6 % missing annual total rainfall, the statistical parameters for 0228495 such as mean and variance did not vary considerably as shown (in Table 6.2) 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.

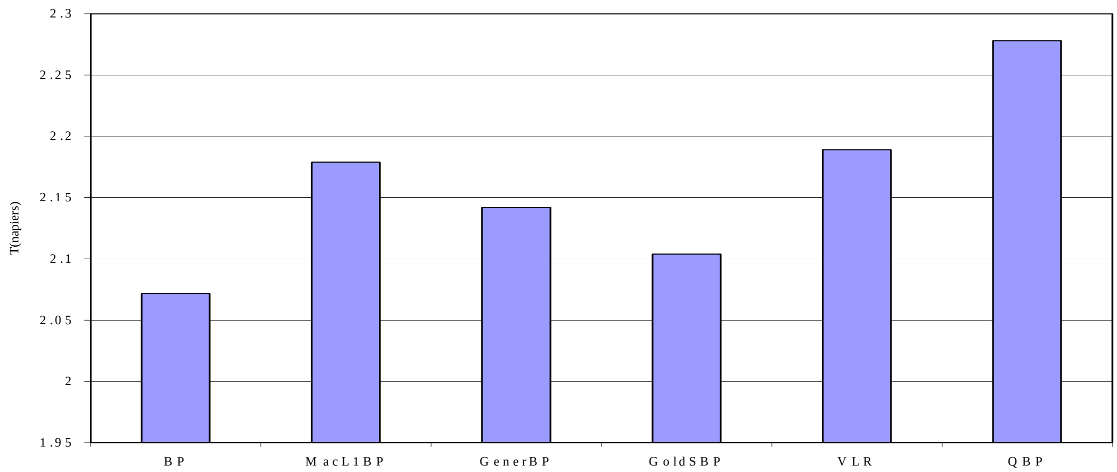


Figure 6.2a Comparison of performance of ANNs in terms of T (7.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

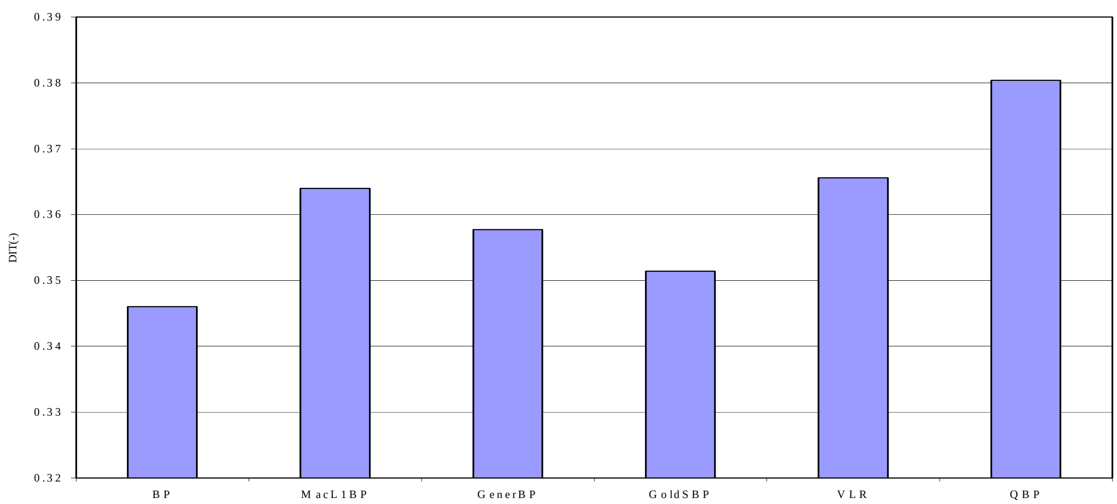


Figure 6.2b Comparison of performance of ANNs in terms of DIT (7.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

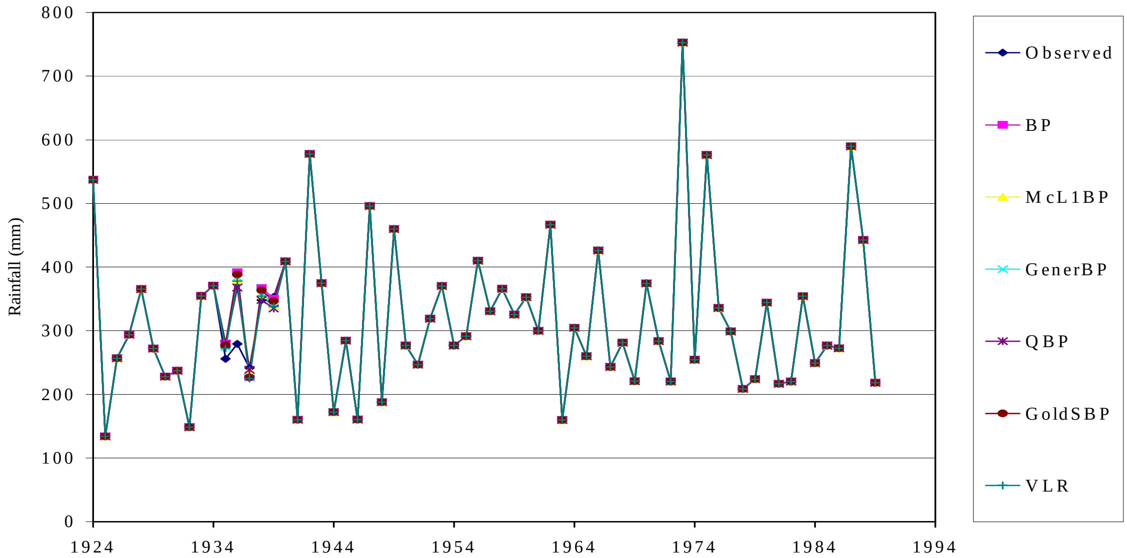


Figure 6.3 Comparison of hydrographs for ANNs (7.6 % missing annual total rainfall from 1935 at 0228495) using base 0228458

6.1.3.1.3 Comparison of performance of EM techniques at gauge 0228495 using base gauge 0228458

The selected EM techniques are thought to be promising estimators of the missing annual total rainfall at 0228495 for gap duration of 7.6 % as long as they satisfy the entropy criterion. From the results shown in Table 6.2 and Figures 6.4 (a-b), it can be seen that the standard EM technique and its variants have identical entropic values. In other words, the proportion of the uncertainty in the annual total rainfall data at 0228495, which has been removed by the application of the standard EM and its variants, is almost the same. This shows that the standard EM leads to the same solution as its variants in some cases.

It can be concluded from entropy calculations that both the standard EM and its variants perform likely equally in estimating the missing annual total rainfall (e.g. 7.6 %) at gauge 0228495. It was noticed some differences in terms of iterations between the standard method and its variants, although the computational efficiency wasn't the primary purpose of this study.

The statistical indicators were also almost identical for all EM techniques, except for MEM2-3.

Figure 6.5 summarizes the hydrographs of the different EM techniques. From a visual observation, all the estimated hydrographs were almost identical. This, again, confirms 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 total rainfall at 0228495 for gap duration of 7.6 %.

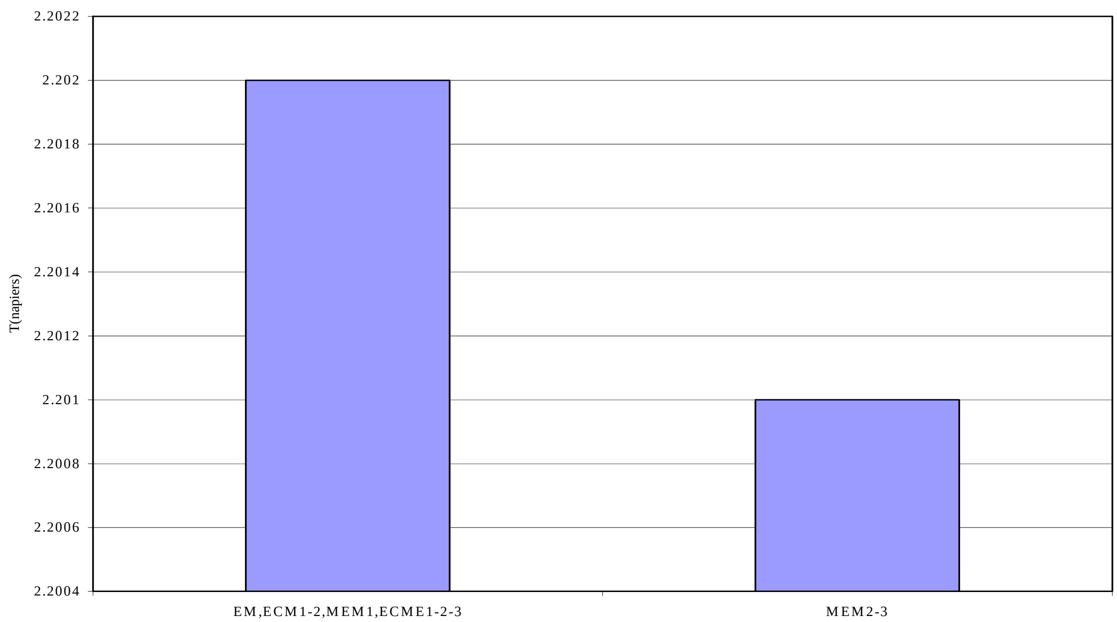


Figure 6.4a Comparison of EM techniques in terms of T (7.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

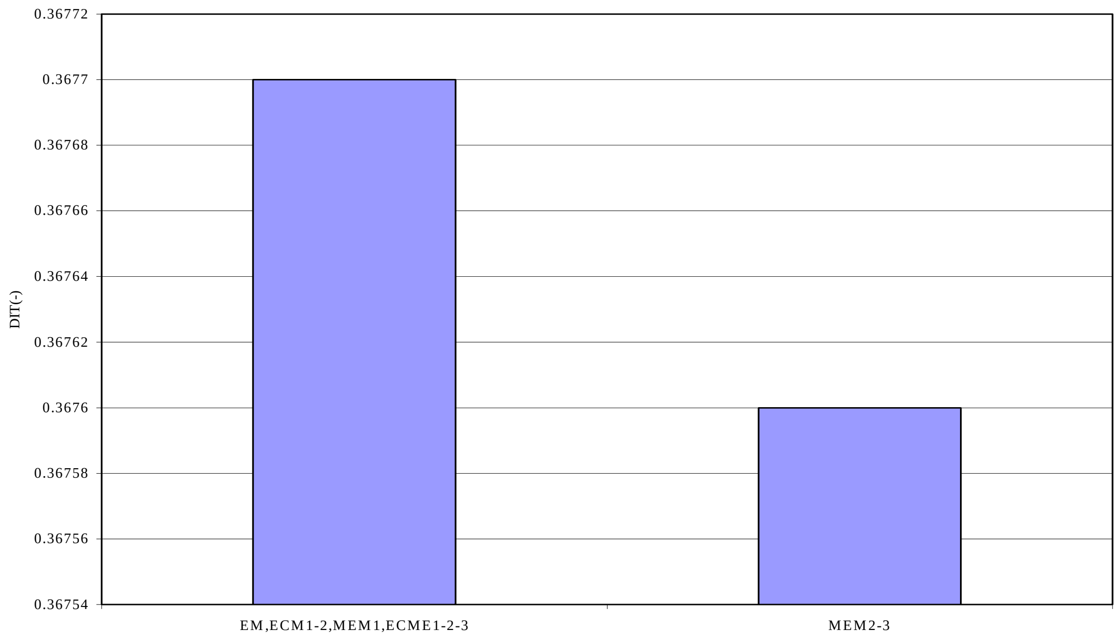


Figure 6.4b Comparison of EM techniques in terms of DIT (7.6 % missing annual total rainfall from 1935 at 0228495) using base 0228458

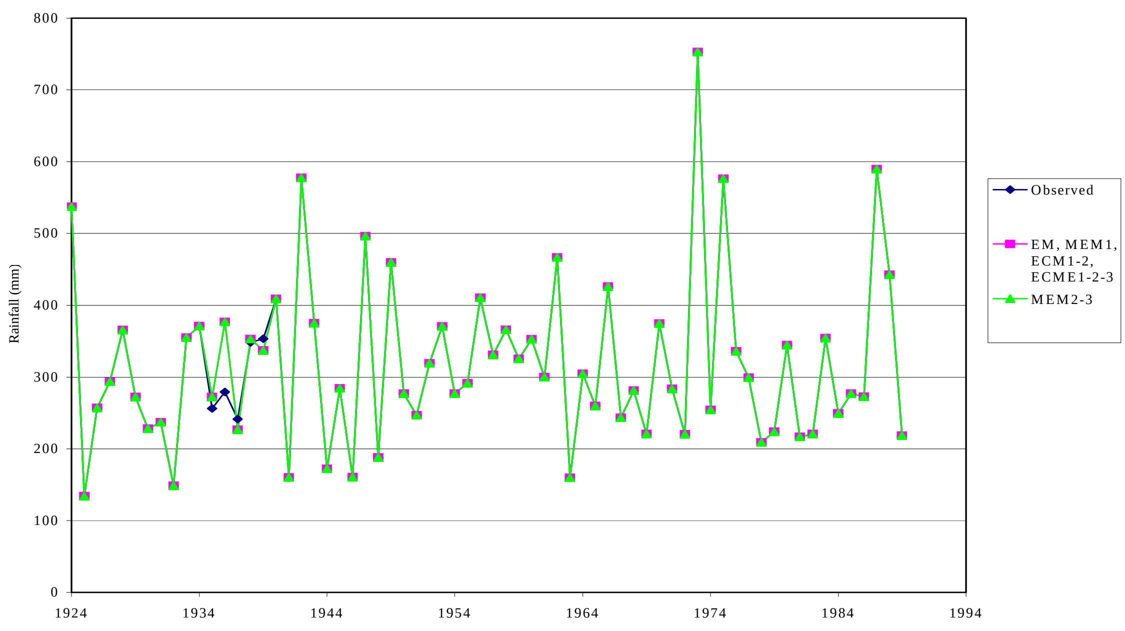


Figure 6.5 Comparison of EM techniques in terms of hydrographs (7.6 % missing total annual rainfall from 1935 at 0228495) using base gauge 0228458

6.1.3.1.4 Comparison of performance of ANN and EM techniques at gauge 0228495 using base gauge 0228458

Comparison between ANN and EM techniques is exhibited in Figures 6.6 (a-b) and Figure 6.7. For visual convenience, only the first best two ANNs and EM techniques are presented here. The EM techniques come in the second position after the QBP and show good estimation capabilities for the missing annual total rainfall at 0228495. This may be due to the presence of linear relationship (in the rainfall data series) between gauges 0228495 and 0228458 since the EM techniques used in were developed within the linear regression context. Khalil et al. (2001) made also a similar observation, when using linear regression methods.

The difference in entropic calculations between the QBP and the EM techniques was very small. In other words, the QBP technique is thought to have slightly a higher capability (than the EM techniques) of mapping the non-linear characteristics in the missing annual total rainfall through the use of a sigmoid 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 EM and first best two ANN techniques for the gauge 0228495 is exhibited in Figure 6.7. The visual observation revealed that the hydrographs were very close.

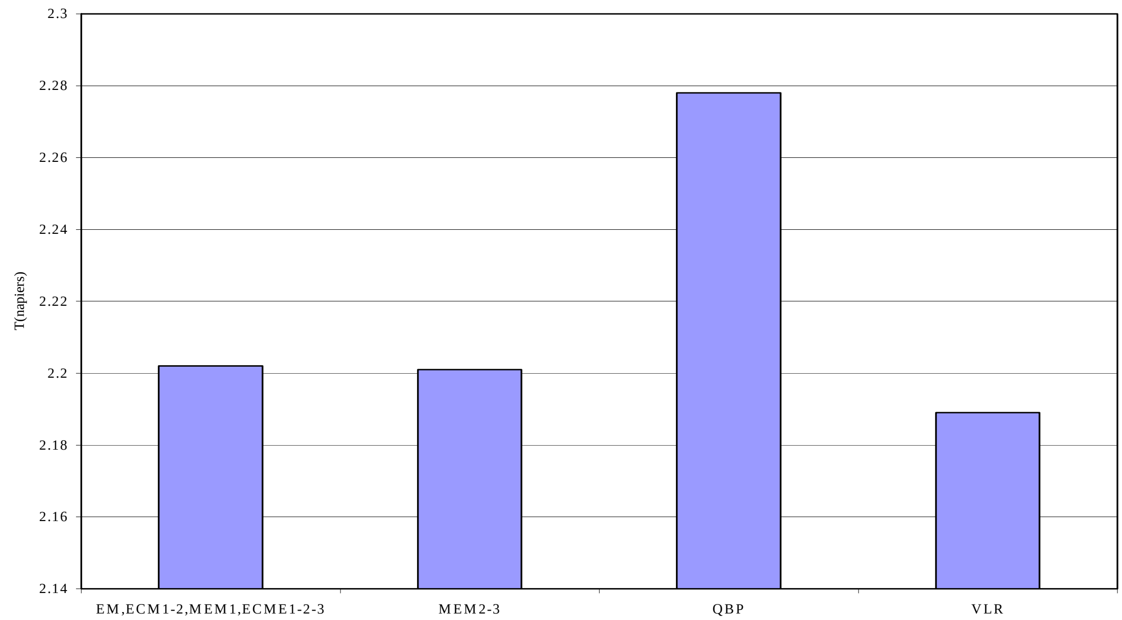


Figure 6.6a Comparison of ANNs and EM techniques in terms of T (7.6 % missing annual total rainfall from 1935 at 0228495) using base 0228458

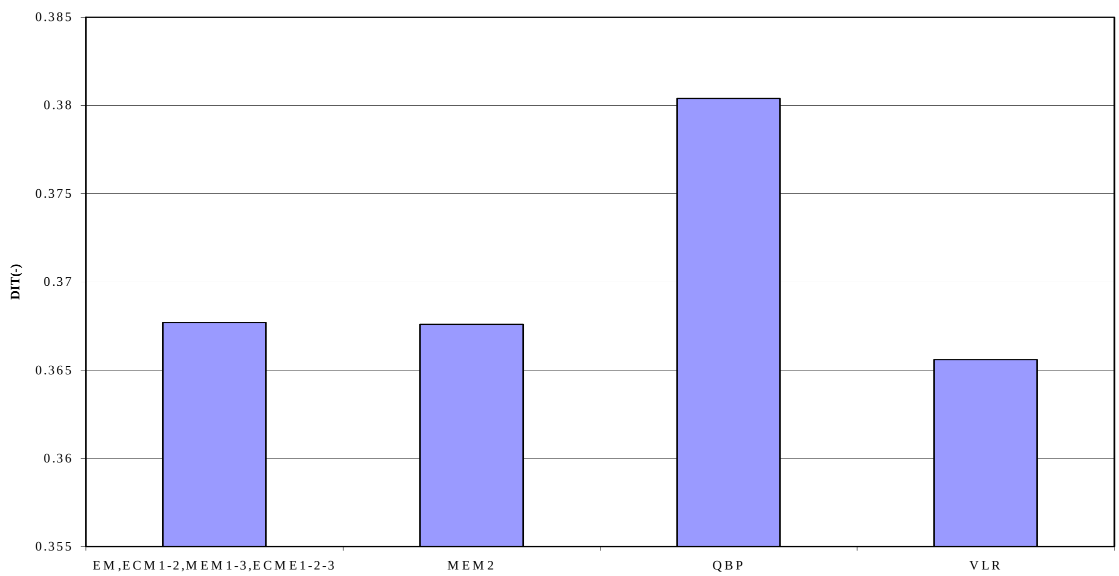


Figure 6.6b Comparison of ANNs and EM techniques in terms of DIT (7.6 % missing annual total rainfall from 1935 at 0228495) using base 0228458

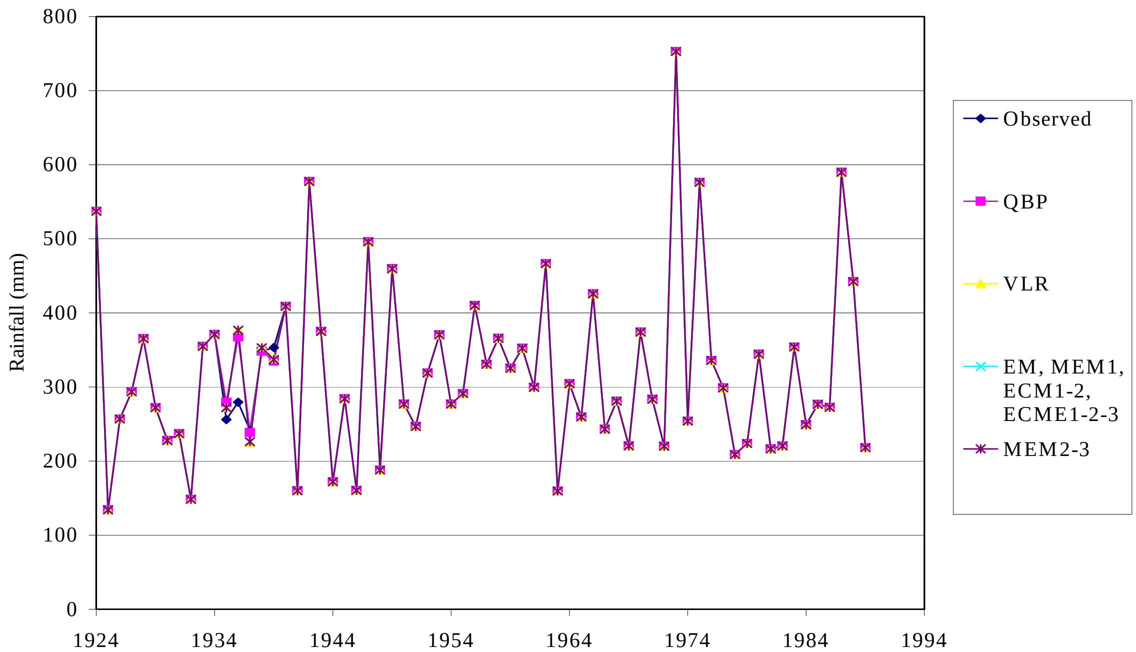


Figure 6.7 Comparison of ANNs and EM techniques in terms of hydrographs (7.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

6.1.3.2 Using 13.6 % of missing annual total rainfall at gauge 0228495

6.1.3.2.1 Selection of ANN and EM techniques for rainfall infilling at gauge 0228495 using base gauge 0228458

The results of performance for the different rainfall data infilling techniques (i.e. ANN and EM) at 0228495 are summarized in Table 6.3 and Figures 6.8 (a-j). Table 6.3 summarizes entropy calculations and statistics at 0228495 using ANN and EM techniques.

The techniques presented in the above-mentioned table are thought to have performed well, i.e. satisfying the entropy criterion. Thus, at least 30% of uncertainty in the annual total rainfall can be removed by applying these techniques. The values for DIT (T) range from 0.3142 (1.882 napiers) to 0.3250 (1.946 napiers). These two limits (i.e. 31.42 % and 32.50 %) are close and fall under the entropy criterion. It can be seen from the results in Table 6.3 that the standard BP performed as well as its variants. The same was noticed for the standard EM and its variants. It is believed that the standard BP leads to some solution in most cases as noticed by Minns and Hall (1996).

The statistical indicators (Table 6.3) and the graphical plots (Figures 6.2 (a-j)) for each respective technique were made just to crosscheck results found from entropy calculations.

The statistical indicators viz. RMSEp and RMEp range respectively from 44.264 to 46.653 mm of rainfall and from 0.0991 to 0.1000. The values for RMEp, which are dimensionless, are relatively small (e.g. close to zero). 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 (Figure 6.8 (a-j)). Both statistical indicators and graphical plots confirm the results from entropy calculations. Thus, the techniques in Table 6.3 are thought to be good estimators of the missing annual total rainfall at gauge 0228495.

Table 6.3 Entropy calculations and statistics at the subject gauge 0228495 (13.6 % missing annual total rainfall) using the base gauge 0228458

	Description	T (napiers)	DIT	RMSEp (mm)	RMEp	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	1.905	0.3181	46.653	0.1000	0.0359	1.0106	0.9947
	McL1 BP (sigmoid hidden layer)	1.882	0.3142	46.682	0.1053	0.0438	1.0078	0.993
	GenerBP (sigmoid hidden layer), s=5	1.931	0.3225	45.094	0.0991	0.0262	1.0114	0.9961
	GoldSBP (sigmoid hidden layer), lr = 0.25	1.905	0.3181	46.602	0.0999	0.0369	1.00965	0.994
	VLR (sigmoid hidden layer)	1.940	0.3240	44.876	0.1011	0.0140	1.0237	0.9980
EM Techniques	Standard EM	1.946	0.3250	44.264	0.1003	0.0134	1.0214	0.998
	MEM1-3	1.946	0.3250	44.264	0.1003	0.0134	1.0214	0.998
	MEM2, u = 0.89, lr = 0.001	1.944	0.3246	44.348	0.1004	0.0158	1.0208	0.9977
	ECM1-2	1.946	0.3246	44.264	0.1003	0.0134	1.0214	0.998
	ECME1-2,(k = 2), 3	1.946	0.3246	44.264	0.1003	0.0134	1.0214	0.998

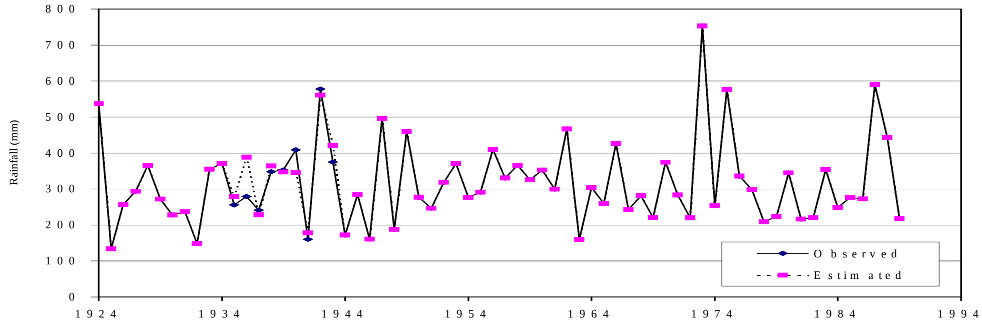


Figure 6.8a Annual total rainfall at 0228495 (13.6 % missing from 1935) using standard BP with base gauge 0228458

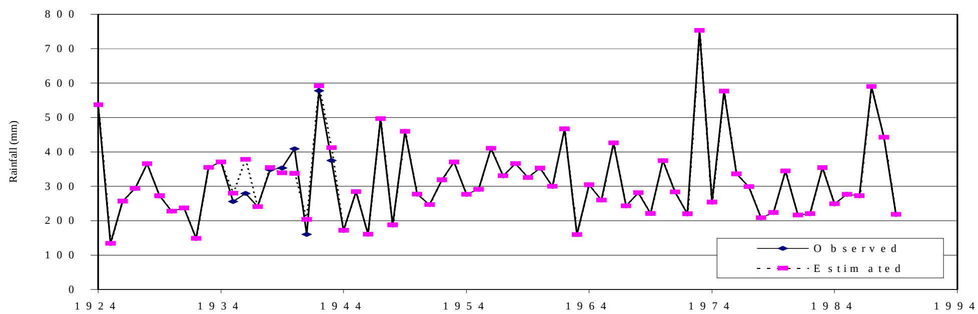


Figure 6.8b Annual total rainfall 0228495 (13.6 % missing from 1935) using McL1BP with base gauge 0228458

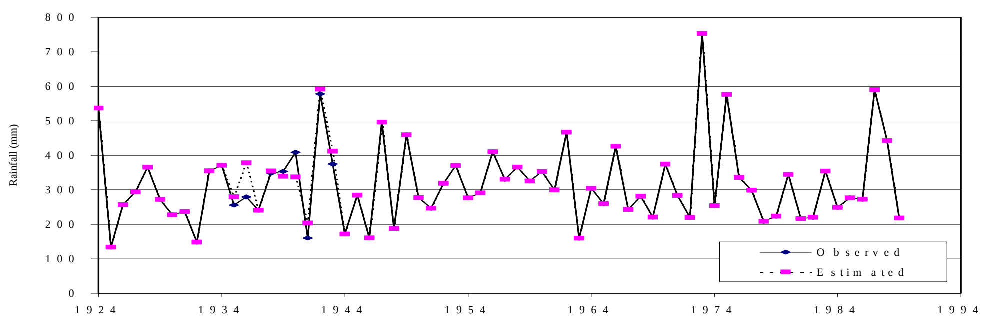


Figure 6.8c Annual total rainfall at 0228495 (13.6 % missing from 1935) using GenerBP with base gauge 0228458

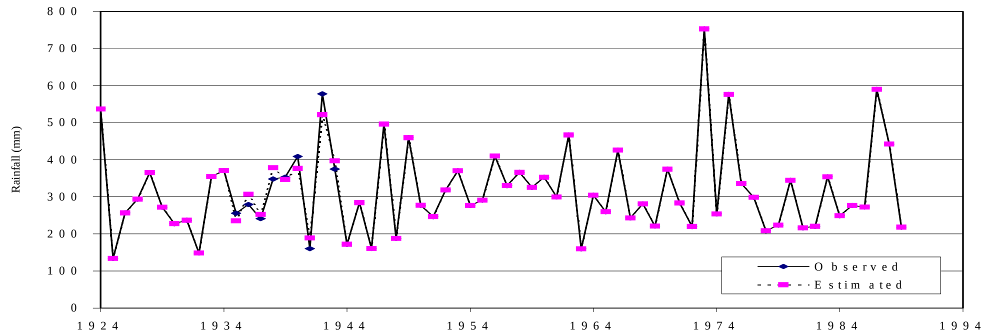


Figure 6.8d Annual total rainfall at 0228495 (13.6 % missing from 1935) using GoldSBP with base gauge 0228458

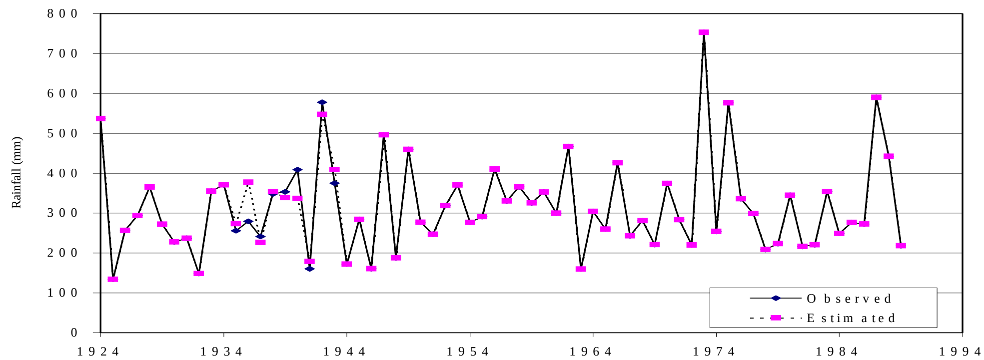


Figure 6.8e Annual total rainfall at 0228495 (13.6 % missing from 1935) using VLRBP with base gauge 0228458

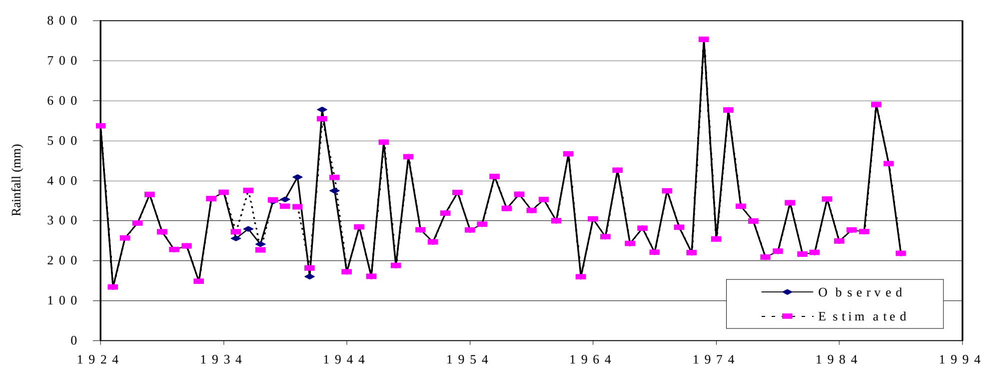


Figure 6.8f Annual total rainfall at 0228495 (13.6 % missing from 1935) using standard EM with base gauge 0228458

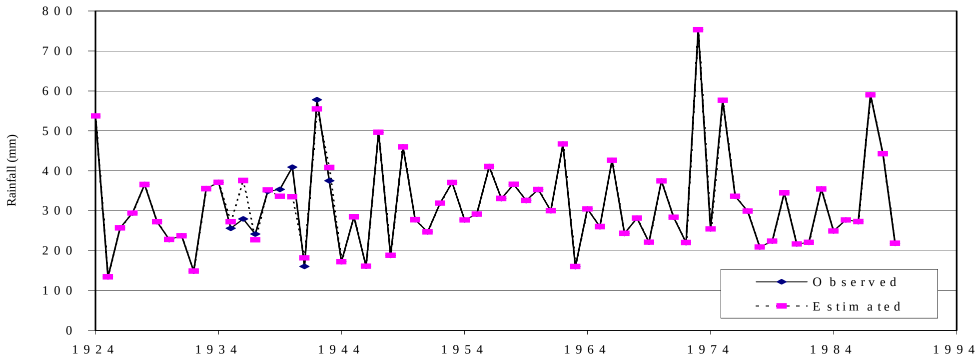


Figure 6.8g Annual total rainfall at 0228495 (13.6 % missing from 1935) using MEM1-3 with base gauge 0228458

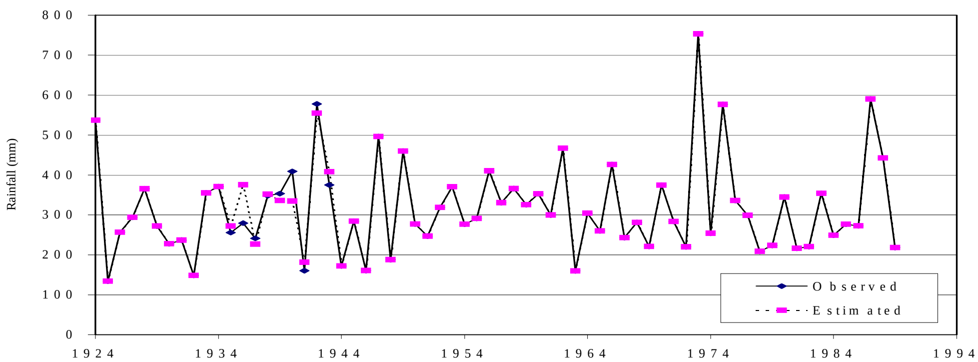


Figure 6.8h Annual total rainfall at 0228495 (13.6 % missing from 1935) using MEM2 with base gauge 0228458

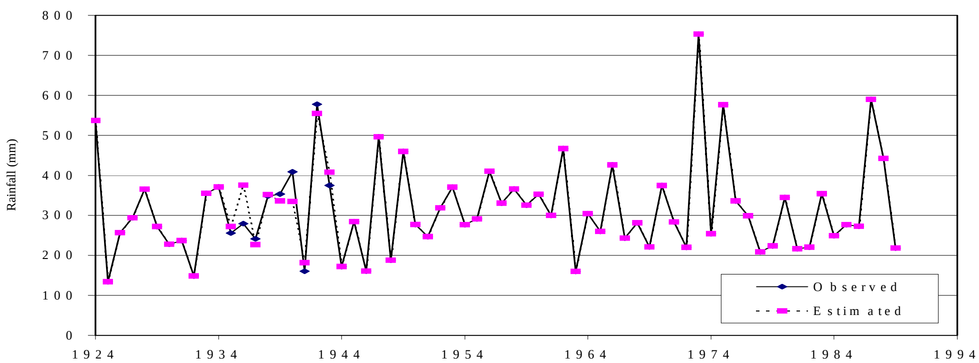


Figure 6.8i Annual total rainfall at 0228495 (13.6 % missing from 1935) using ECM1-2 with base gauge 0228458

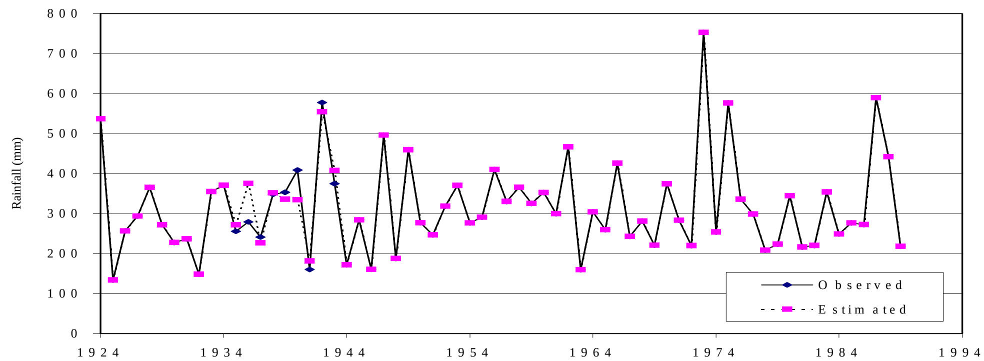


Figure 6.8j Annual total rainfall at 0228495 (13.6 % missing from 1935) using ECME1-2-3 with base gauge 0228458

6.1.3.2.2 Comparison of performance of ANN techniques at gauge 0228495 using base gauge 0228458

Based on the entropic values for the selected techniques, model performance assessment was then made for the estimation of missing annual total rainfall at gauge 0228495. For gap duration of 13.6 % at gauging station 0228495, the performance in a decreasing order of the different ANN techniques in terms of DIT (T) values was as follows:

-VLR

-GenerBP, s = 5

-Standard BP (hyperbolic tangent hidden layer)

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

-GoldSBP

-McL1BP

From Table 6.3 and Figures 6.9 (a-b), the values for DIT (T) range from 0.3142 (1.882 napiers) to 0.3240 (1.940 napiers). Thus, the first best model (i.e. VLR) among the selected ANN techniques have the highest entropic value of 0.3240 (1.940 napiers) for DIT (T). In other words, 32.40 % of uncertainty can be removed from the annual total rainfall at gauge 0228495 via the VLR technique. The VLR technique is thought to have relatively a higher capability (than the rest of ANNs) of estimating the missing annual rainfall at 0228495. The McL1BP technique has the lowest value of entropy 0.3181

(1.905 napiers) for DIT (T). It can be noticed that 31.93 % of information can be inferred about gauge 0228495 via the standard BP. This value is not very different from the VLR as shown in Figures 6.9 (a-b). This technique is thought, in this study, to be a good estimator of missing data at 0228495.

The values of RMSEp range from 44.876 to 46.653 mm of rainfall for the ANNs and those of RMEp range from 0.0991 to 0.1000. The VLR technique has the lowest values of RMSEp and RMEp while the McL1BP has the highest value for RMSEp and RMEp. These results enhance the conclusions drawn from the entropy calculations.

Figure 6.10 shows that the results are very close. Thus, this figure did not show any substantial difference between the ANN techniques.

In general, the techniques presented above appear to be good estimators of the missing annual total rainfall (13.6 %) at gauge 0228495. For this gap duration, the above selected ANNs preserve the statistical parameters (i.e. mean and variance) for 0228495.

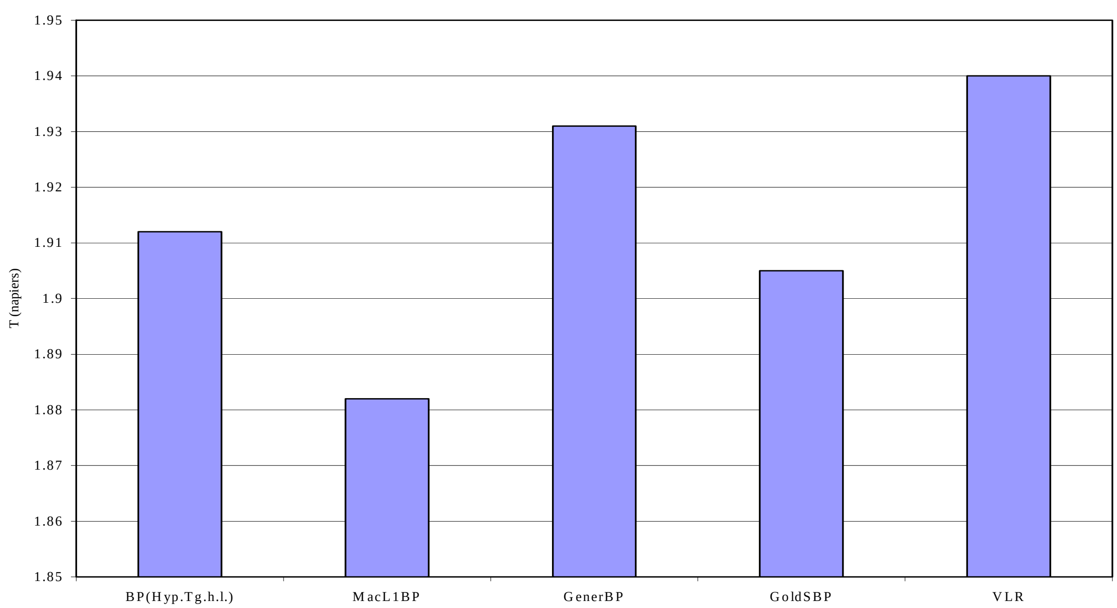


Figure 6.9a Comparison of ANNs in terms of T (13.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

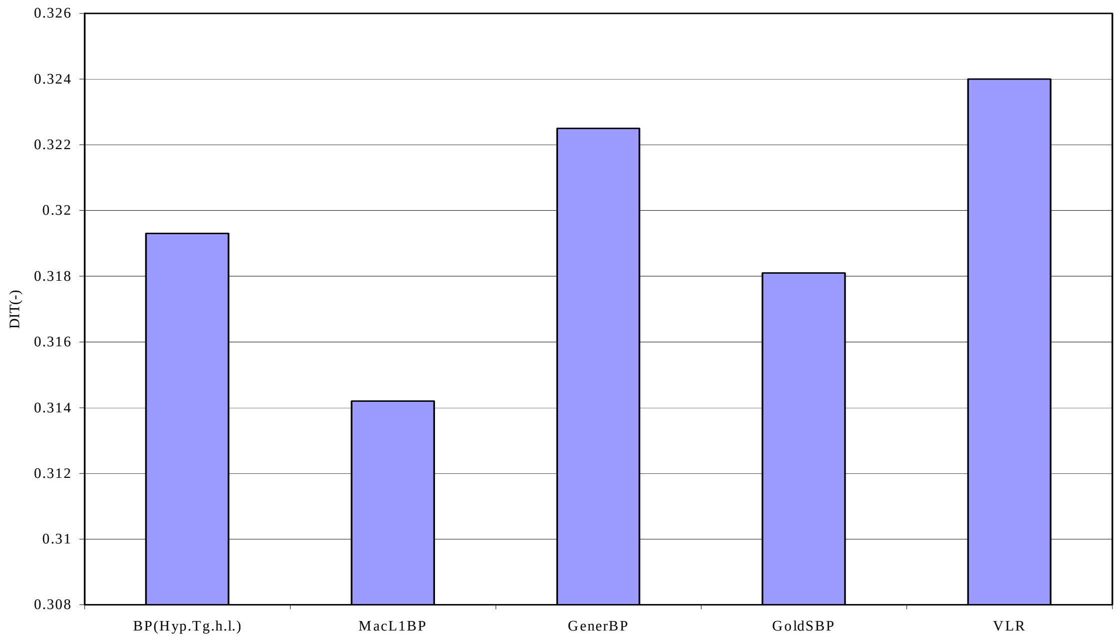


Figure 6.9b Comparison of ANNs in terms of DIT (13.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

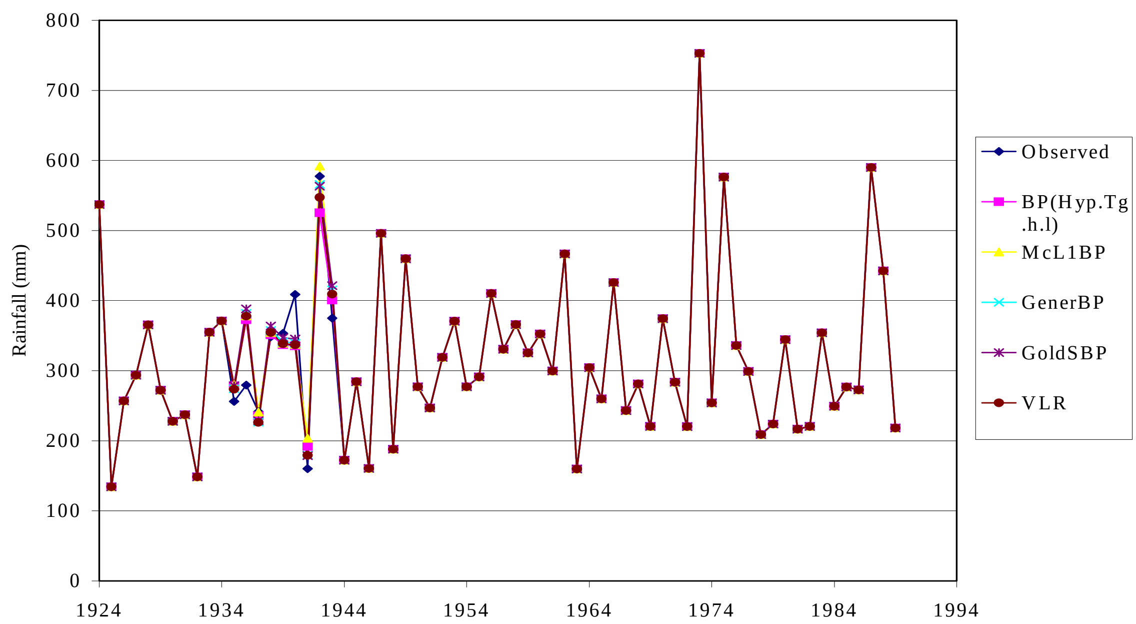


Figure 6.10 Comparison of ANNs in terms of hydrographs (13.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

6.1.3.2.3 Comparison of performance of EM techniques at gauge 0228495 using base gauge 0228458

The selected EM techniques are thought to be promising estimators of the missing annual rainfall at the subject station 0228495 for gap duration of 13.6 % as long as they satisfy the entropy criterion. From the results presented in Table 6.3 and Figures 6.11 (a-b), it can be seen that the standard EM technique and its variants have almost identical entropic values. It is believed 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 total rainfall values (e.g. 13.6 %) at gauge 0228495. Again, this could be due to the form of the maximum likelihood estimation of parameters (for EM techniques) considered in this thesis.

The statistical indicators were also almost identical for all the EM techniques. Figure 6.12 summarizes the comparison of the performance of the different EM techniques. From a visual observation, all the estimated hydrographs were almost identical.

Both the graphical and statistical comparison between the different EM techniques confirm the results obtained from the entropy calculations.

In an overall assessment, the EM techniques and its variants lead almost to the same solution and are thought to be promising estimators of the missing annual total rainfall at 0228495 for gap duration of 13.6 %.

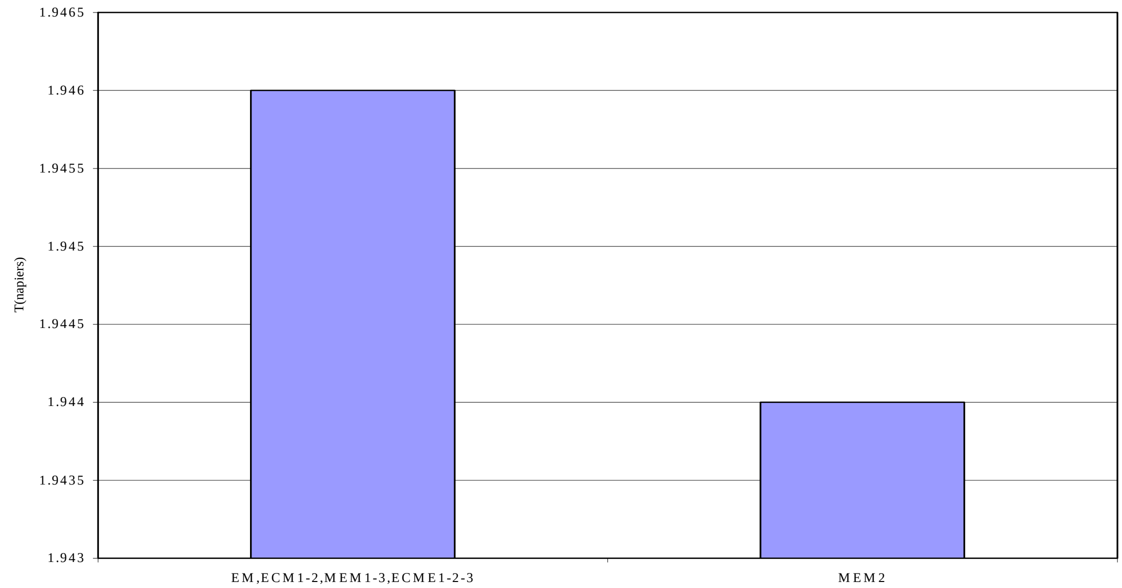


Figure 6.11a Comparison of EM techniques in terms of T (13.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

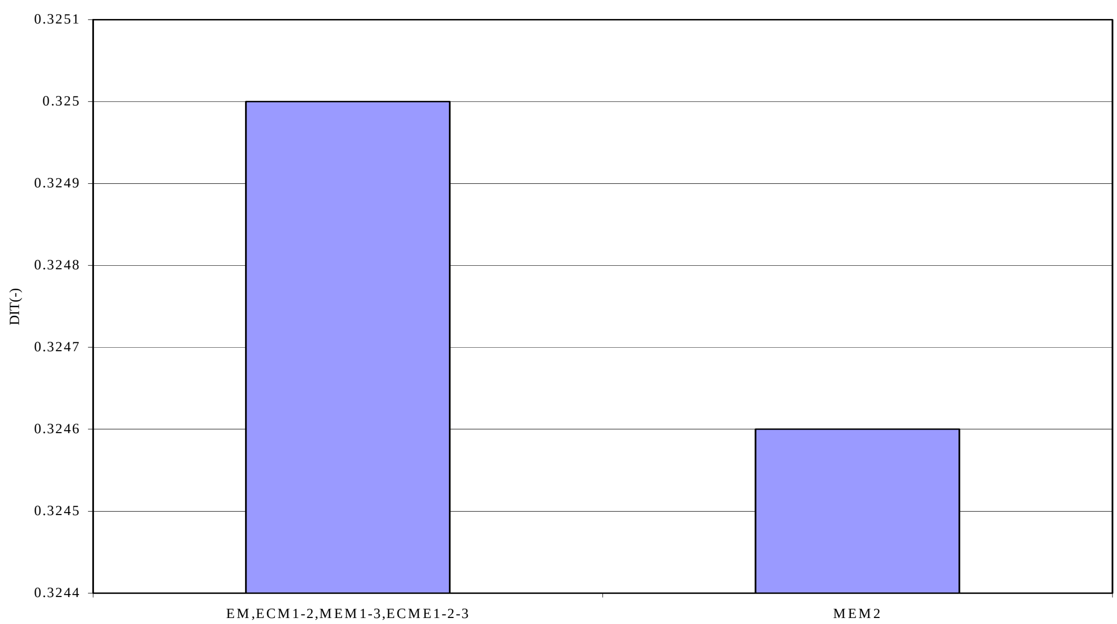


Figure 6.11b Comparison of EM techniques in terms of DIT (13.6 % missing annual total rainfall from 1935 at the target 0228495) using the base gauge 0228458

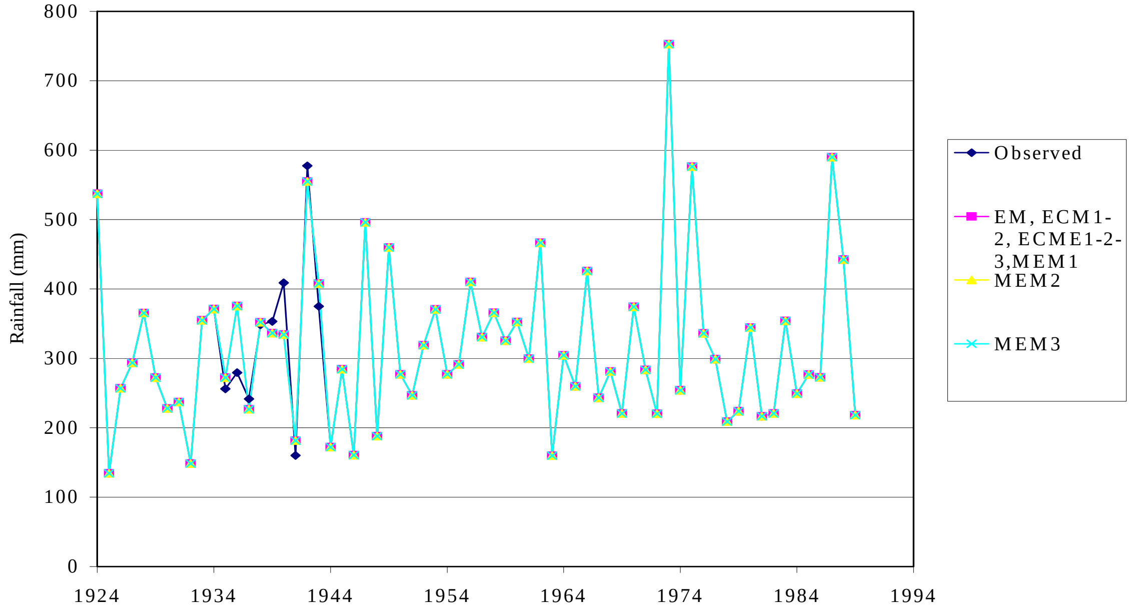


Figure 6.12 Comparison of EM techniques in terms of hydrographs (13.6 % missing annual total rainfall from 1935 at the target 0228495) using the base gauge 0228458

6.1.3.2.4 Comparison of performance of ANN and EM techniques at gauge 0228495 using base gauge 0228458

Comparison between ANN and EM techniques is exhibited in Figures 13 (a-b). For visual convenience, only the best first two ANN techniques and all the EM techniques are presented here. Accordingly, the EM techniques come in first position before the ANNs. Thus, the EM techniques show higher estimation capabilities (than ANNs) for the missing annual total rainfall at 0228495. This may be due to the presence of a strong linear relationship in the rainfall data series between gauges 0228495 and 0228458 since the EM techniques used here were developed within the linear regression context. Recall that a similar observation was also made by Khalil et al. (2001), when using linear regression methods.

The difference in entropic values between these ANNs and the EM techniques were small.

The graphical comparison (hydrographs) between the EM and best first two ANNs techniques for gauge 0228495 is exhibited in Figure 6.14. A visual observation reveals that the hydrographs were very close.

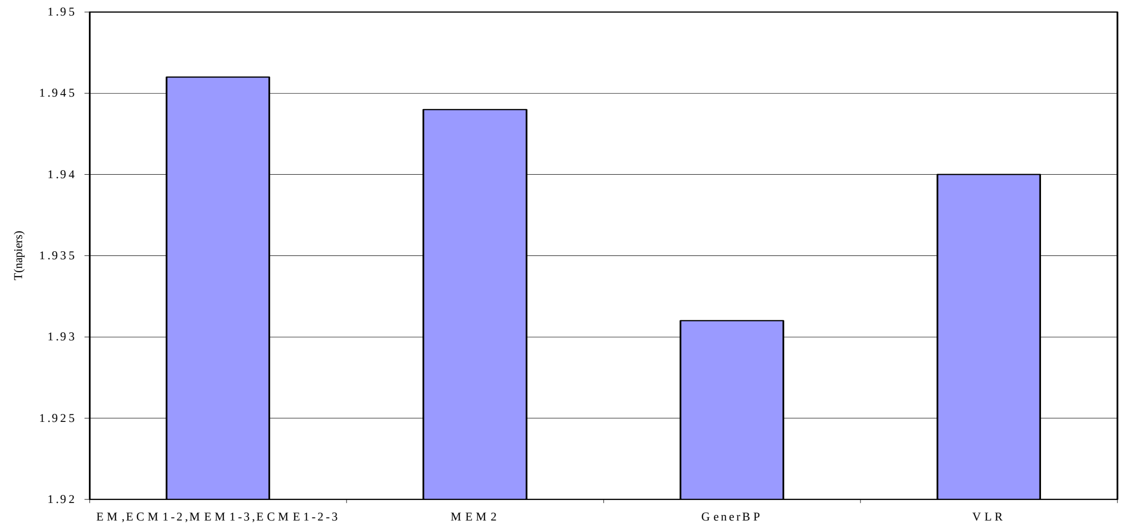


Figure 6.13a Comparison of ANN and EM techniques in terms of T (13.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

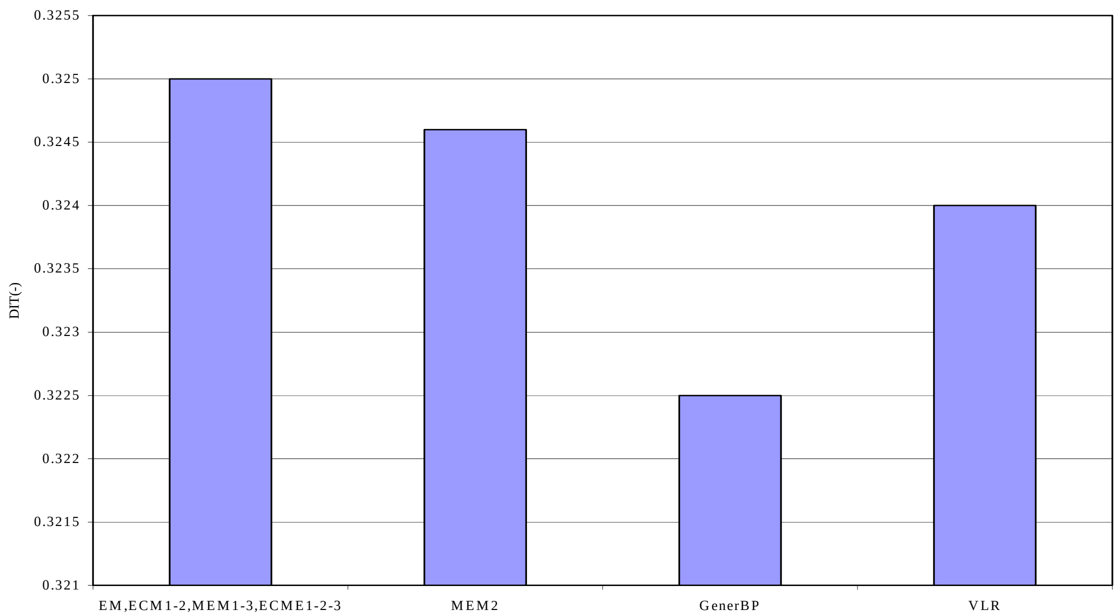


Figure 6.13b Comparison of ANN and EM techniques in terms of DIT (13.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

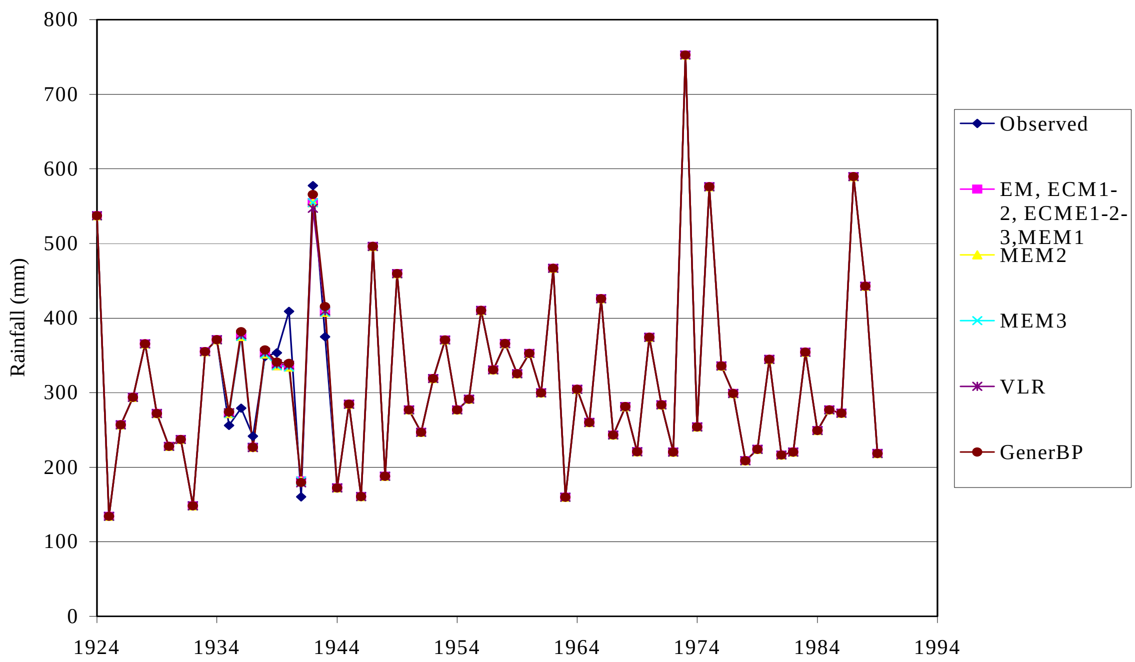


Figure 6.14 Comparison of ANNs and EM techniques in terms of hydrographs (13.6 % missing annual total rainfall from 1935 at 0228495) using base gauge 0228458

6.1.4 Infilling annual total rainfall at gauge 0228458 with base gauge 0228495

6.1.4.1 Using 7.6 % of missing annual total rainfall at gauge 0228458

6.1.4.1.1 Selection of ANN and EM techniques for rainfall infilling at gauge 0228458 using base gauge 0228495

The results of performance for the different rainfall data infilling techniques (e.g. ANN and EM) at station 0228458 are summarized in Table 6.4 and Figures 6.15 (a-i). Table 6.4 summarizes entropy calculations and statistics at the target station 0228458 using ANN and EM techniques, with base station 0228495.

The techniques presented in the above-mentioned table are thought to have performed well, i.e. satisfying the entropy criterion. The values for DIT (T) range from 0.3370 (2.018 napiers) to 0.3579 (2.143 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 can be noticed for the standard EM technique and its variants.

The statistical indicators (Table 6.4) and the graphical plot (Figures 6.15 (a-j)) for each respective technique was just made to crosscheck the results found from entropy calculations. The statistical indicators viz. RMSEp and RMEp range respectively from 46.132 mm to 52.563 mm of rainfall and from 0.120 to 0.1524. The values for RMEp, which are dimensionless, are relatively small. This can eventually confirm the results obtained from entropy criterion. 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 selected technique. Both the statistical indicators and the graphical plot confirm the results from entropy calculations. Hence, the selected techniques are thought to be good estimators of the missing annual total rainfall values at 0228458 for gap duration of 7.6 %.

Table 6.4 Entropy calculations and statistics at the target gauge 0228458 (7.6 % missing annual rainfall) using the base gauge 0228495

	Description	T (napiers)	DIT	RMSEp (mm)	RMEp	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	2.143	0.3579	46.132	0.1203	-0.0576	1.0088	1.0041
	McL1BP (sigmoid hidden layer)	2.018	0.3370	52.563	0.1524	-0.1049	1.0094	1.0071
	GenerBP (sigmoid hidden layer), s=5	2.127	0.3552	46.819	0.1224	-0.0614	1.0094	1.0043
	QBP (sigmoid hidden layer), acc = 0.015, lr = 0.85, Weight Cond	2.142	0.3577	45.885	0.118	-0.0511	1.011	1.0036
	GoldSBP(sigmoid hidden layer), lr = 0.25	2.135	0.3565	46.49	0.1214	-0.0594	1.0091	1.004
	VLR (sigmoid hidden layer)	2.140	0.3574	46.20	0.120	-0.0055	1.0092	1.0039
EM Techniques	Standard EM	2.127	0.3552	46.822	0.1225	-0.0611	1.0093	1.0043
	MEM1-2-3	2.127	0.3552	46.822	0.1225	-0.0611	1.0093	1.0043
	ECM1-2	2.127	0.3552	46.822	0.1225	-0.0611	1.0093	1.0043
	ECME1-2-3	2.127	0.3552	46.822	0.1225	-0.0611	1.0093	1.0043

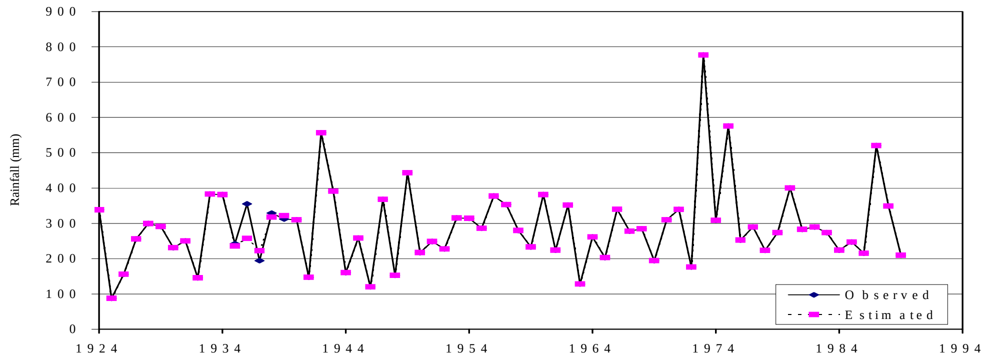


Figure 6.15a Annual total rainfall at 0228458 (7.6 % missing from 1935) using standard BP with base gauge 0228495

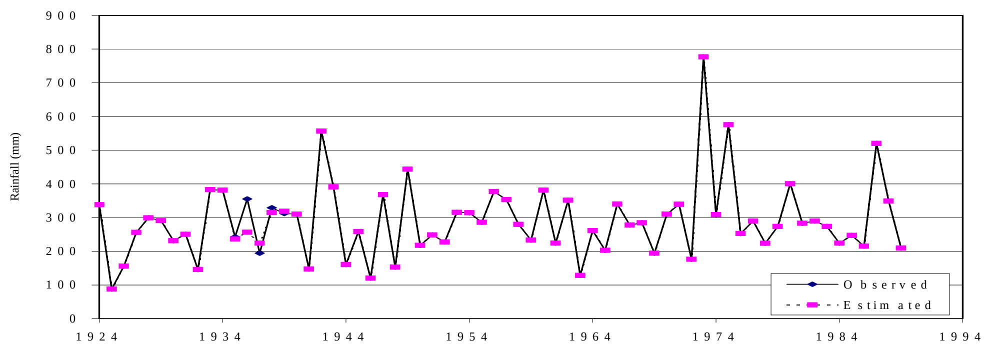


Figure 6.15b Annual total rainfall at 0228458 (7.6 % missing from 1935) using GenerBP with 0228495

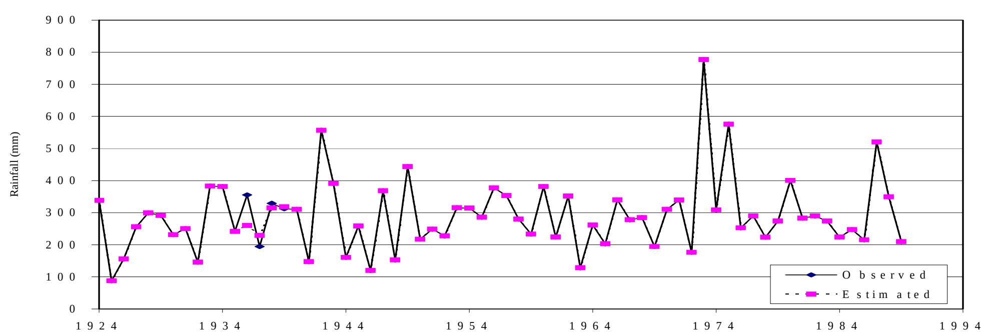


Figure 6.15c Annual total rainfall at 0228458 (7.6 % missing from 1935) using QBP with base gauge 0228495

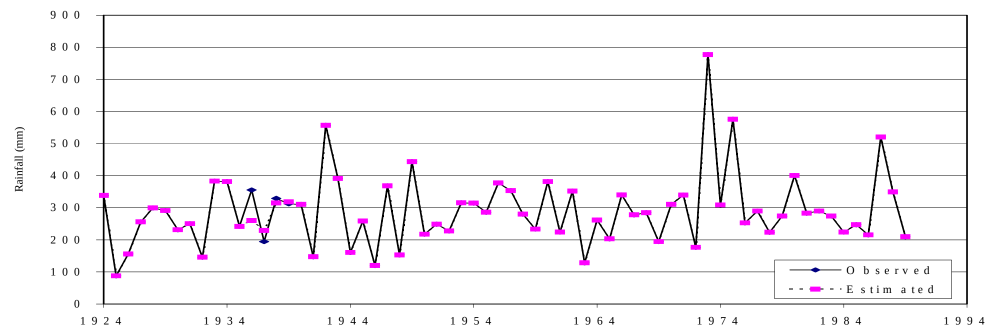


Figure 6.15d Annual total rainfall at 0228458 (7.6 % missing from 1935) using GoldSBP with base gauge 0228495

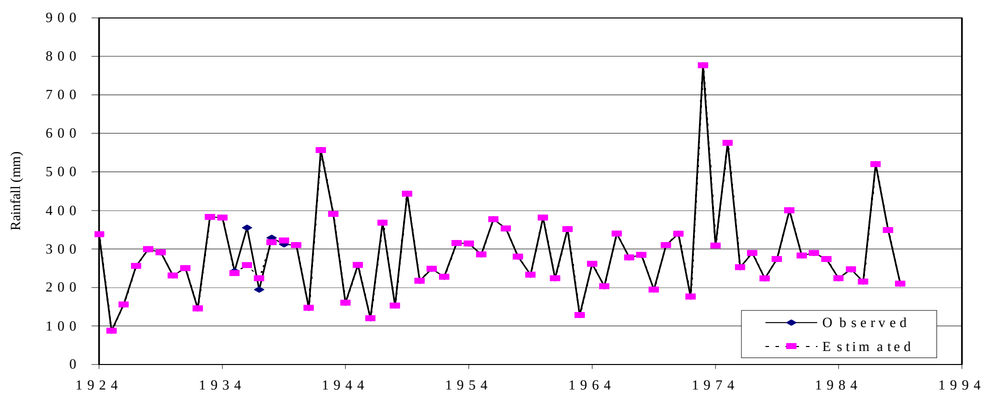


Figure 6.15e Annual total rainfall at 0228458 (7.6 % missing from 1935) using VLRBP with base gauge 0228495

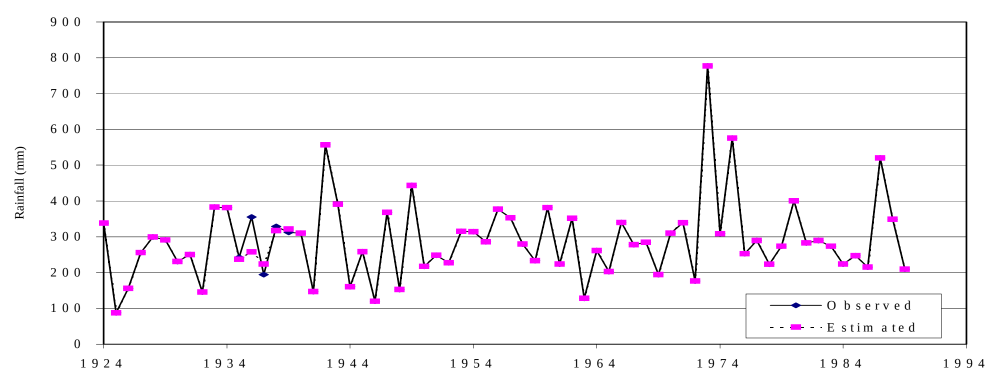


Figure 6.15f Annual total rainfall at 0228458 (7.6 % missing from 1935) using McL1BP with base gauge 0228495

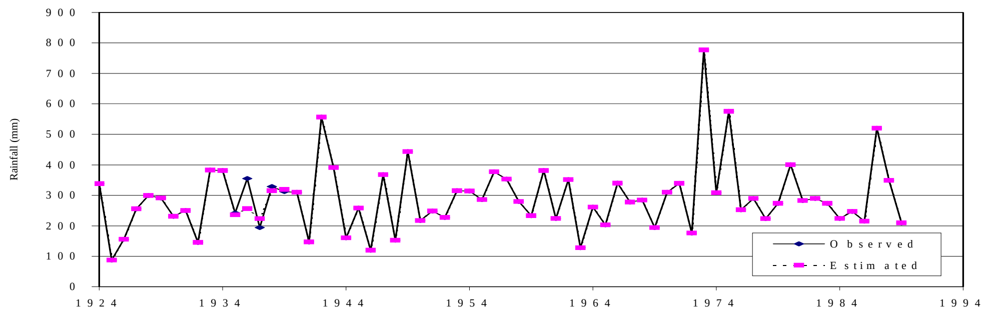


Figure 6.15g Annual total rainfall at 0228458 (7.6 % missing from 1935) using standard EM with base gauge 0228495

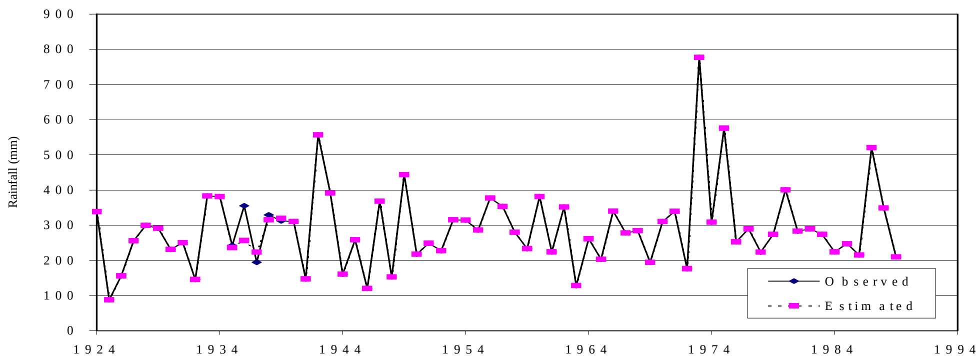


Figure 6.15h Annual total rainfall at 0228458 (7.6 % missing from 1935) using ECM1-2 with base gauge 0228495

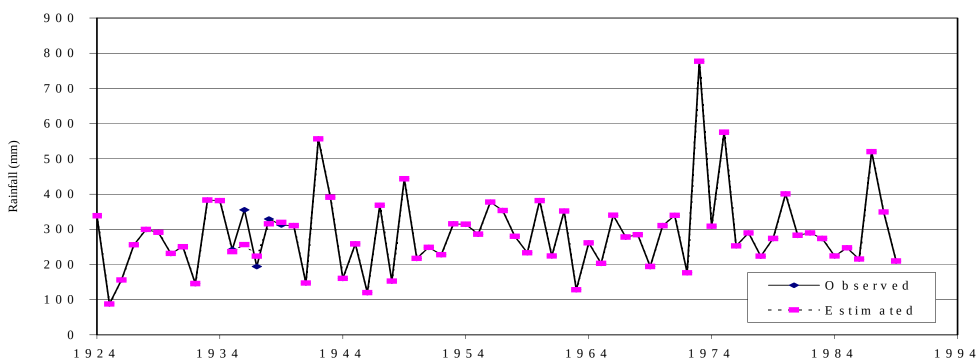


Figure 6.15i Annual total rainfall at 0228458 (7.6 % missing from 1935) using MEM1-2-3 with base gauge 0228495

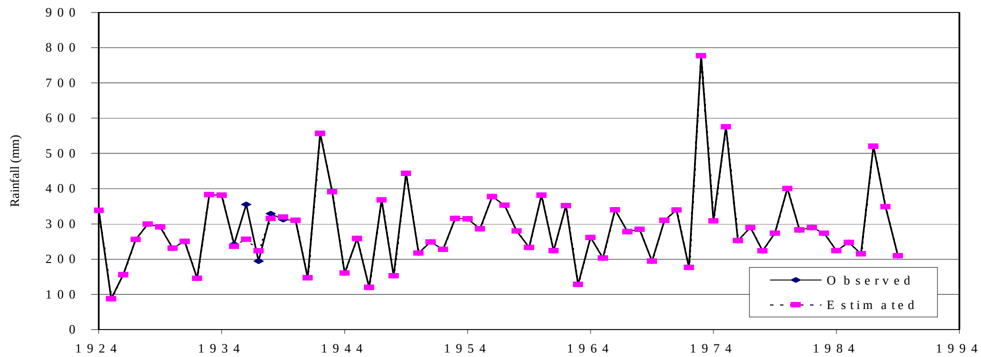


Figure 6.15j Annual total rainfall at 0228458 (7.6 % missing from 1935) using ECME1-2-3 with base gauge 0228495

6.1.4.1.2 Comparison of performance of ANN techniques at gauge 0228458 using base gauge 0228495

Based on the entropic values for the selected techniques, model performance assessment was then made for the estimation of missing annual total rainfall values at gauge 0228458. Considering 0228495 as base station and for gap duration of 7.6 % at 0228458, the performance (in a decreasing order) of the different ANN techniques in terms of DIT (T) is as follows:

-Standard BP

-QBP, acc = 0.015, lr = 0.85 and weight cond.

-VLR

-GoldSBP

-GenerBP, lr = 0.25, s = 5

-McL1BP

From Table 6.4 and Figures 6.16 (a-b), the values of DIT (T) range from 0.3370 (2.018 napiers) to 0.3579 (2.143 napiers). Thus, the first best technique (i.e. standard BP) among the selected ANN techniques has the highest entropic value of 0.3579 (2.143 napiers) for DIT (T). In other words, 35.79 % of uncertainty can be removed from the annual total rainfall data at gauging station 0228458, via the standard BP technique. The last best technique, i.e. McL1BP has the capability of removing 33.70 % of uncertainty from rainfall data at 0228458. It is noticed that the standard BP technique is thought to have relatively a higher capability (than the rest of ANNs) of estimating the missing annual total rainfall at 0228458.

Figures 6.16 (a-b) summarize the comparison of the different ANN techniques in terms of their DIT (T) values. All the values are above the threshold value of 30 %.

On the other hand, the values for RMSEp range from 46.132 mm to 52.563 mm of rainfall. These two values correspondent to the standard BP and McL1BP respectively. The values for RMEp ranged 0.120 to 0.1534. The standard BP technique has relatively low values of statistical indicators while the McL1BP has the highest values for the same statistical indicators. These results just confirm the conclusions drawn from the entropy calculations.

Figure 6.17 summarizes the comparison of hydrographs for the different ANN techniques. This figure shows that the results are very close. This plot did not show any substantial difference between the different techniques.

In general, these techniques appear to be good estimators of the missing annual total rainfall values (7.6 %) at gauge 0228458. For 7.6 % missing annual total rainfall, the statistical parameters for 0228458 such as mean and variance did not vary considerably as shown (in Table 6.4) by the ratio of observed statistic to the estimated statistic. In other words, the above selected ANNs preserved the mean and the variance of rainfall data at 0228458.

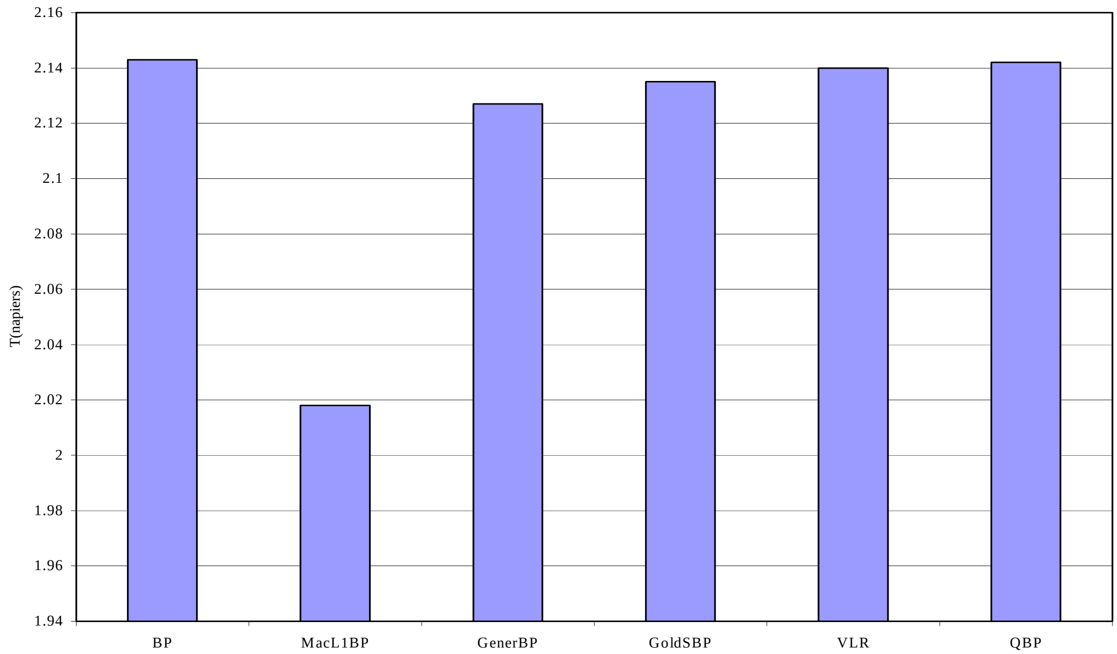


Figure 6.16a Comparison of ANNs in terms of T (7.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

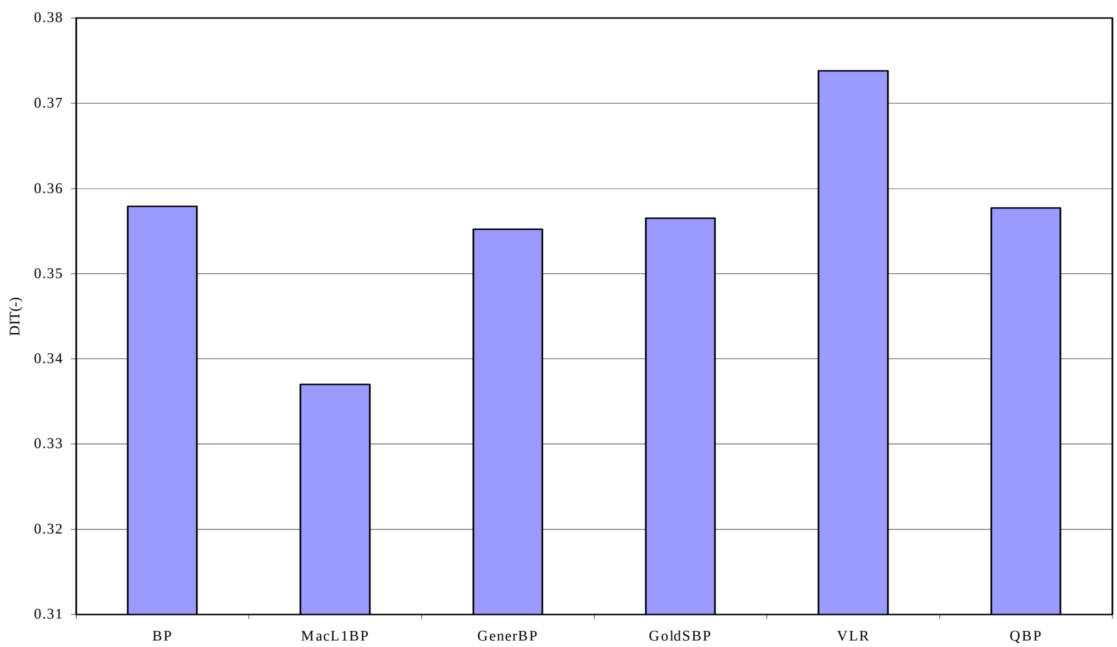


Figure 6.16b Comparison of ANNs in terms of DIT (7.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

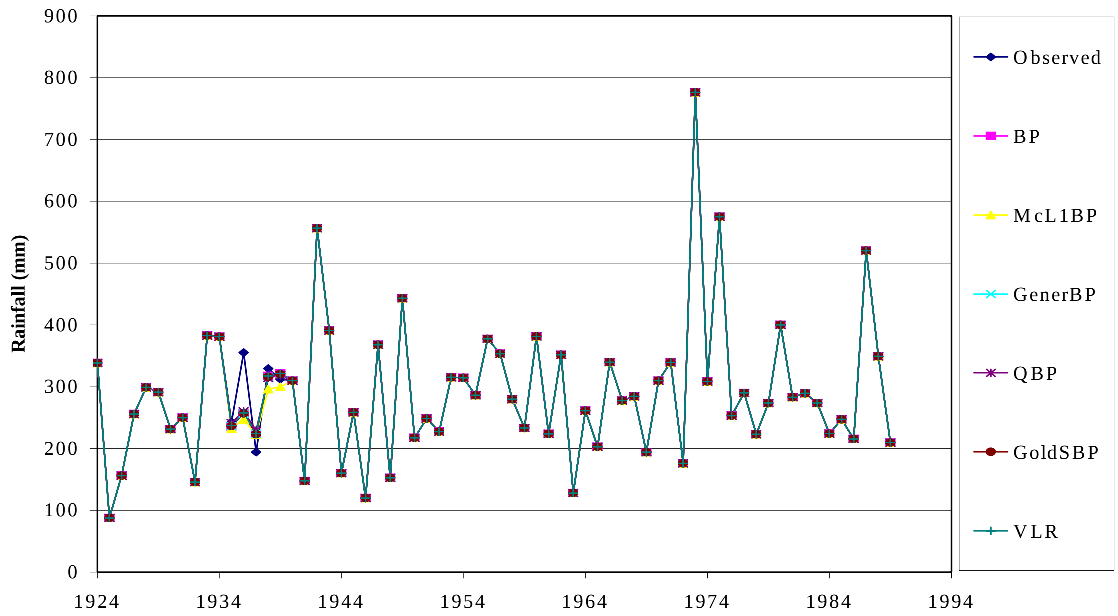


Figure 6.17 Comparison of ANN in terms of hydrographs (7.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

6.1.4.1.3 Comparison of performance of EM techniques at gauge 0228458 using base gauge 0288495

The selected EM techniques are thought to be promising estimators of the missing values at 0228495 for gap duration of 7.6 % as long as they satisfy the entropy criterion. From the results shown in Table 6.4, it can be seen that the standard EM technique and its variants have identical entropic values. In other words, the proportion of the uncertainty in the original (observed) data, which has been physically removed by the application of EM and its variants, is almost the same. This shows that the standard EM leads to the same solution as its variants in some cases. It is believed that these results could be due to the form of the maximum likelihood estimation of parameters for EM techniques, which is in fact not complicated as taken in this thesis. However, for complicated cases Meng and Xu (1993) recommended using other variants.

Although the primary purpose of this thesis was not the computational efficiency of the different techniques, it was just noticed some differences in terms of iterations between the standard method and its variants. It can be concluded from entropy calculations that both the standard EM and its variants perform likely equally in estimating the missing values (e.g. 7.6 %) at gauging station 0228458. The use of the momentum or the learning rate in the different versions as used 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 rainfall values. No graphical comparison was made since the results were identical everywhere.

In an overall assessment, the EM techniques lead almost to the same solution and are thought to be promising estimators of the missing annual total rainfall at 0228458 for gap duration of 7.6 %.

6.1.4.1.4 Comparison of performance of ANN and EM techniques at gauge 0228458 using base gauge 0228495

Comparison between ANN and EM techniques is made in terms of DIT (T) and is exhibited in Figures 6.18 (a-b). The EM techniques come in the fourth position after standard BP, QBP and GoldSBP and show, however, good estimation capabilities for the missing annual total rainfall at 0228458. This may be due to the presence of linear relationship in the rainfall data between gauges 0228495 and 0228458 since the EM techniques used in this study, were developed within the linear regression context.

The difference in entropic calculations between the above-mentioned ANNs and the EM techniques was small. In other words, the first best three ANNs are thought to have slightly higher capabilities (than the EM techniques) of mapping the non-linear characteristics in the missing annual total rainfall through the use of a sigmoid 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 ANN techniques at the subject station 0228458 is exhibited in Figure 6.19. The visual observation revealed that the hydrographs were very close.

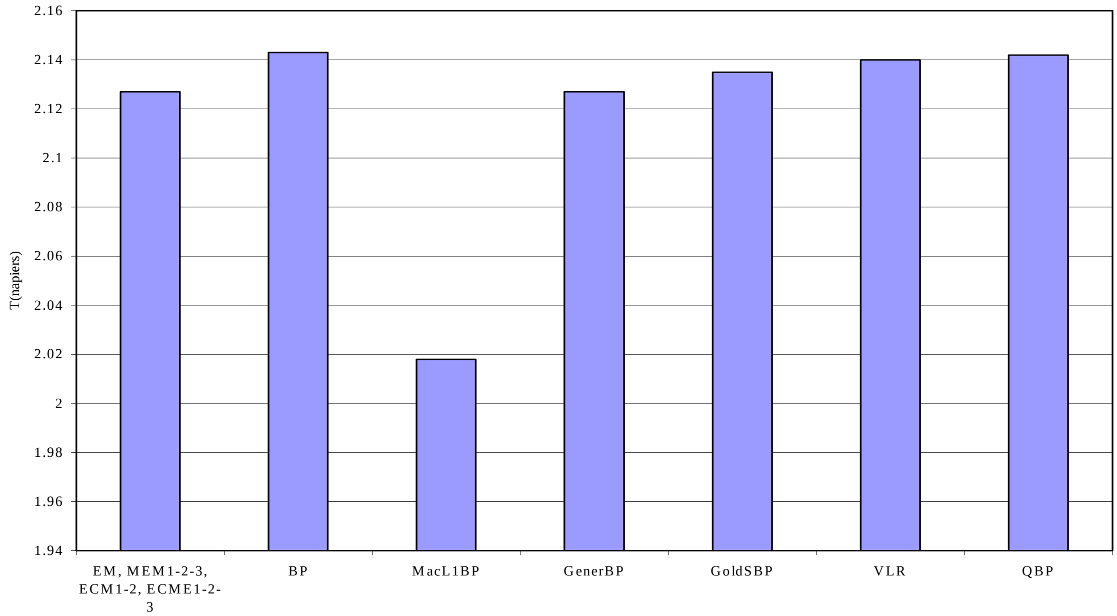


Figure 6.18a Comparison of ANN and EM in terms of T (7.6 % missing annual total rainfall from 1935 0228458) using base gauge 0228495

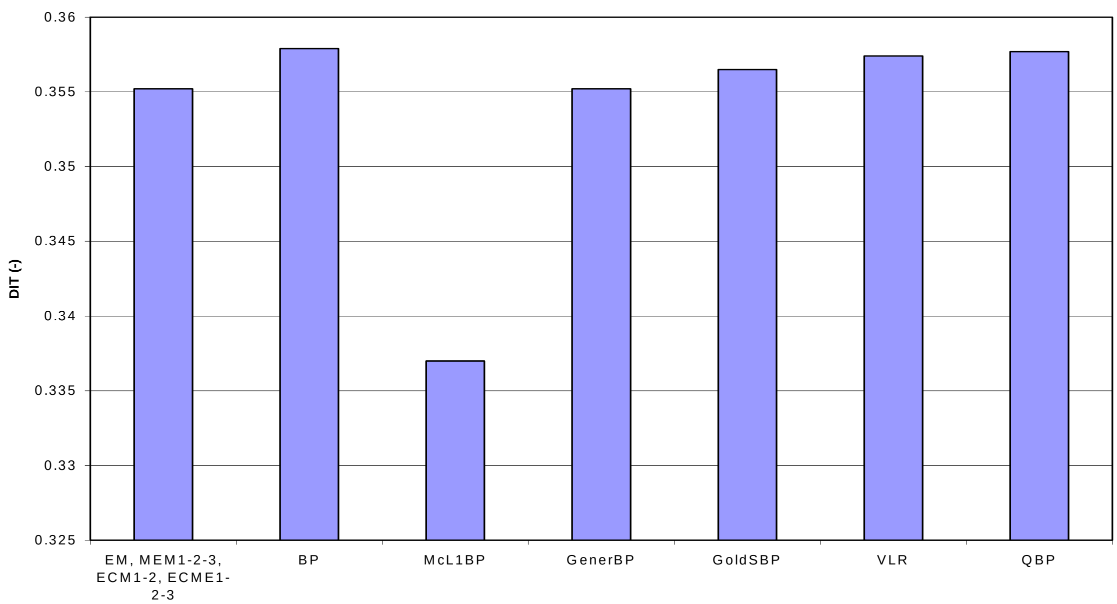


Figure 6.18b Comparison of ANN and EM in terms of DIT (7.6 % missing annual total rainfall from 1935 0228458) using base gauge 0228495

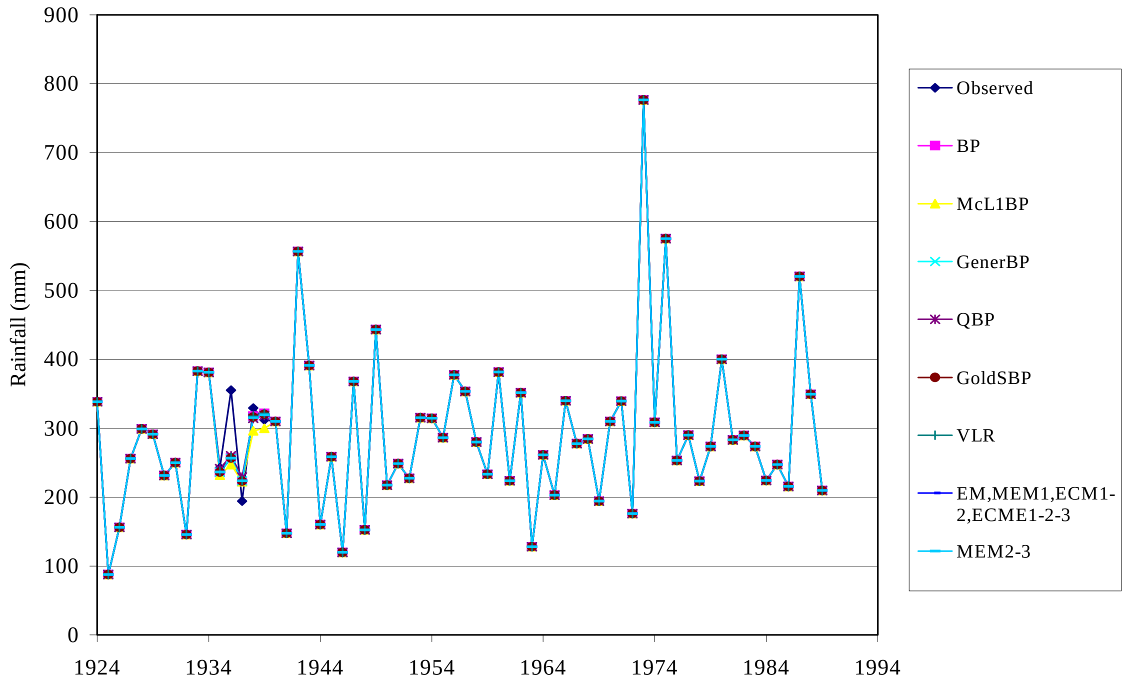


Figure 6.19 Comparison of ANN and EM techniques in terms of hydrographs (7.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

6.1.4.2 Using 13.6 % of missing annual total rainfall at gauge 0228458

6.1.4.2.1 Selection of ANN and EM techniques for rainfall infilling at gauge 0228458 using base gauge 0228495

The results of performance for the different rainfall data infilling techniques (e.g. ANN and EM) at the subject station 0228458 are shown in Table 6.5 and Figures 6.20 (a-i). 0228495 is the base station for estimating the missing annual total rainfall at 0228458. Table 6.5 summarizes entropy calculations and statistics at 0228458 using ANN and EM techniques.

The techniques presented in Table 6.5 are thought to have performed well, i.e. satisfying the entropy criterion. At least 30 % of uncertainty in the annual total rainfall data at the subject station can be removed by applying these techniques. The values for DIT (T) range from 0.3138 (1.879 napiers) to 0.3170 (1.898 napiers). These two limits, although close, fall under the entropy criterion. In terms of percentage, these two limits are 31.42 % and 32.50 % respectively. It could be said that the standard BP performed as well as its variants. It is believed that the BP leads to a reasonable solution in most cases Minns and Hall (1996).

The statistical indicators (Table 6.5) and the graphical plots (Figures 6.20 (a-i)) for each respective technique were made to crosscheck the results found from entropy calculations. The statistical indicators viz. RMSEp and RMEp range respectively from 46.51 to 48.316 mm of rainfall and from 0.1168 to 0.1190. The values for RMEp, which are dimensionless, are relatively small. 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 6.20 (a-i)).

Both statistical indicators and graphical plots confirm the conclusions from entropy computations.

Thus, the selected techniques are thought to be good estimators of the missing annual total rainfall values at gauging station 0228458, for gap duration 7.6 %.

Table 6.5 Entropy calculations and statistics at the target gauge 0228458 (13.6 % missing annual total rainfall) using the base gauge 0228495

	Description	T (napiers)	DIT	RMSEp (mm)	RMEp	EV	Ratio Variance	Ratio Mean
ANN Techniques	BP (sigmoid hidden layer)	1.879	0.3138	48.316	0.117	-0.0528	1.05	1.007
	GenerBP (sigmoid hidden layer), s=5	1.894	0.3163	46.868	0.119	-0.0453	1.048	1.0065
	GoldSBP (sigmoid hidden layer), lr = 0.25	1.883	0.3145	47.921	0.1189	-0.0520	1.05	1.007
	VLR (sigmoid hidden layer)	1.898	0.3170	46.861	0.1168	-0.0432	1.048	1.006
EM Techniques	Standard EM	1.894	0.3163	46.74	0.119	-0.0472	1.0466	1.0067
	MEM1-3	1.894	0.3163	46.74	0.119	-0.0472	1.0466	1.0067
	MEM2, u = 0.89, lr = 0.001	1.897	0.3170	46.51	0.119	-0.0448	1.0461	1.0064
	ECM1-2	1.894	0.3163	46.74	0.119	-0.0472	1.0466	1.0067
	ECME1-2 (k = 2), 3	1.894	0.3163	46.74	0.119	-0.0472	1.0466	1.0067

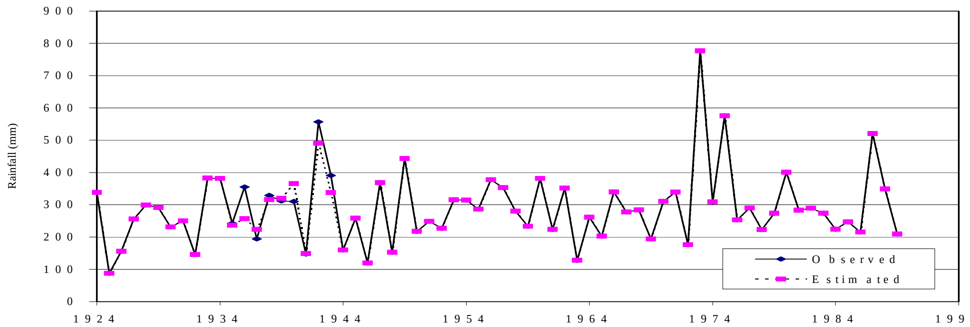


Figure 6.20a Annual rainfall at 0228458 (13.6 % missing from 1935) using standard BP with base gauge 0228495

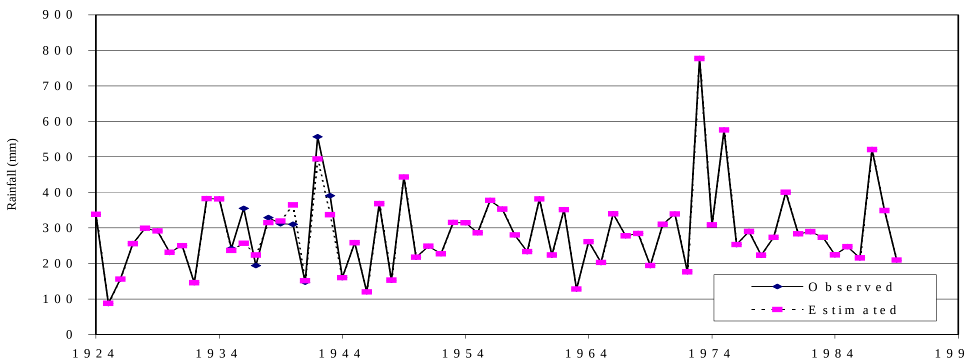


Figure 6.20b Annual rainfall at 0228458 (13.6 % missing from 1935) using GenerBP with base gauge 0228495

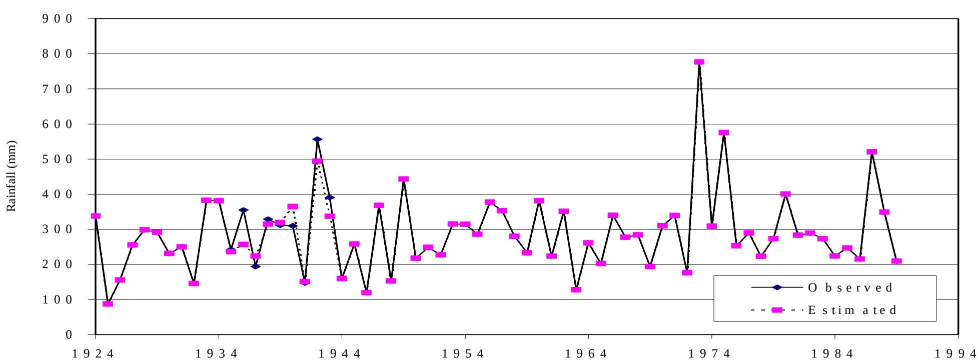


Figure 6.20c Annual rainfall at 0228458 (13.6 % missing from 1935) using GoldSBP with base gauge 0228495

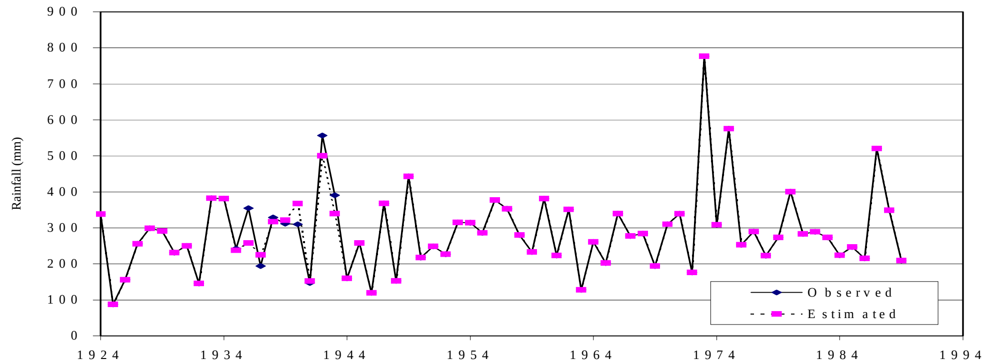


Figure 6.20d Annual rainfall at 0228458 (13.6 % missing from 1935) using VLRBP with base gauge 0228495

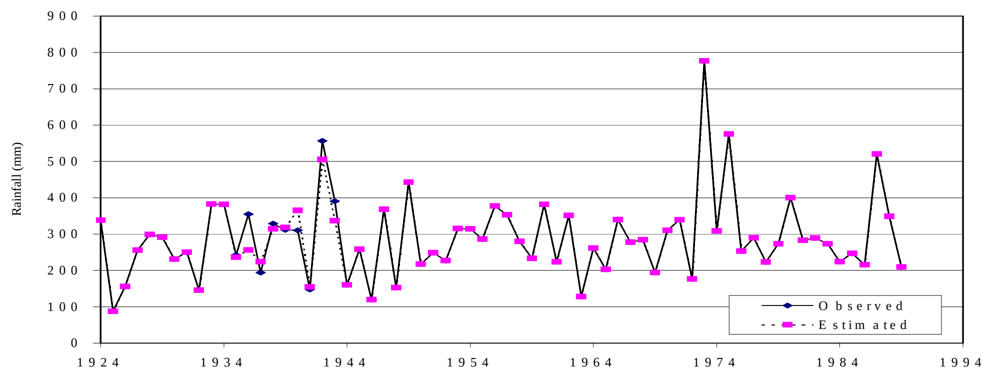


Figure 6.20e Annual rainfall at 0228458 (13.6 % missing from 1935) using standard EM with base gauge 0228495

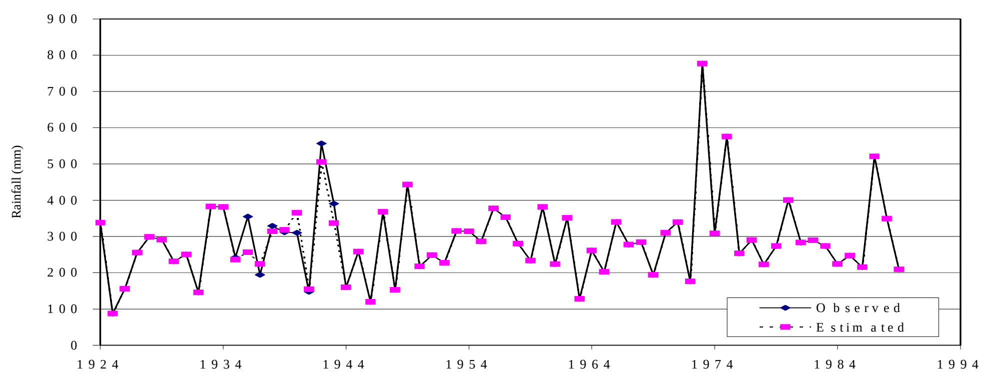


Figure 6.20f Annual rainfall at 0228458 (13.6 % missing from 1935) using MEM1-3 with base gauge 0228495

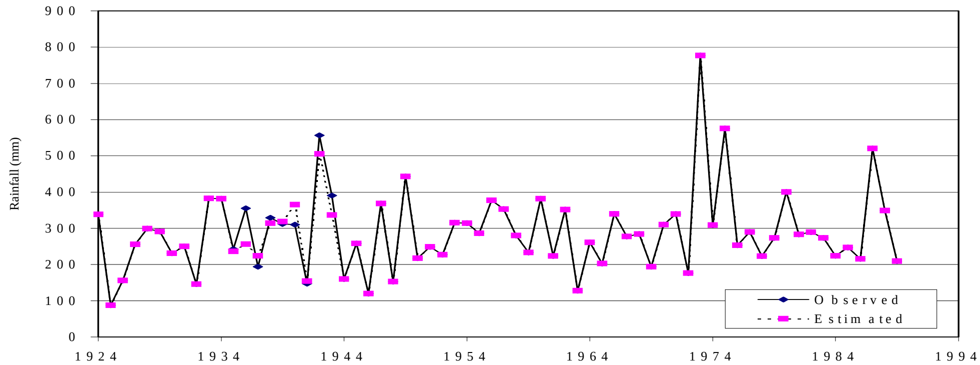


Figure 6.20g Annual rainfall at 0228458 (13.6 % missing from 1935) using MEM2 with base gauge 0228495

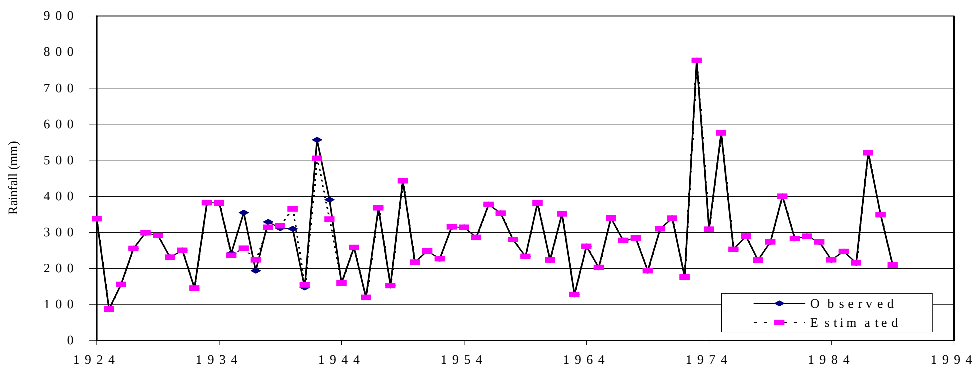


Figure 6.20h Annual rainfall at 0228458 (13.6 % missing from 1935) using ECM1-2 with base gauge 0228495

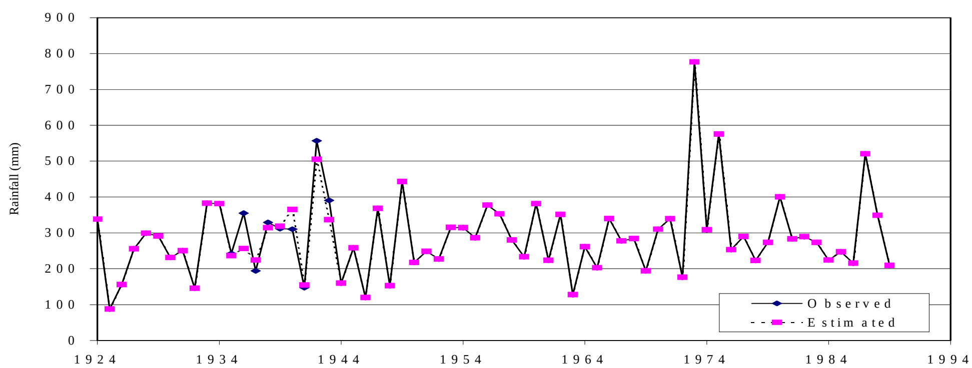


Figure 6.20i Annual rainfall at 0228458 (13.6 % missing from 1935) using ECME1-2-3 with base gauge 0228495

6.1.4.2.2 Comparison of performance among ANN techniques at gauge 0228458 using base gauge 0228495

Based on the entropic values for the selected techniques (Table 6.5), model performance assessment was then made for the estimation of missing annual total rainfall values at gauging station 0228458. For gap duration of 13.6 % of annual total rainfall at 0228458, the performance in a decreasing order of the different ANN techniques in terms of DIT (T) values is as follows:

-VLR

-GenerBP, s = 5

-GoldSBP

-Standard BP (hyperbolic tangent hidden layer)

McL1BP and QBP techniques were not selected among the best techniques, as they did not satisfy the entropy criterion. From Table 6.5 and Figures 6.21 (a-b), the values for DIT (T) range from 0.3138 (1.879 napiers) to 0.3170 (1.898 napiers). The first best technique (i.e. VLR) among the selected ANNs has relatively the highest entropic value, i.e. upper limit of the range for DIT (T). The lower limit value corresponds to the standard BP technique. However, these two limits are close. In other words, 31.70 % of uncertainty can be removed from the annual total rainfall at gauge 0228458 via the VLR technique. This can be also seen as the proportion of information transferred by the knowledge of the estimated series (through VRL) into the process to make the annual total rainfall data at 0228458 better defined. The VLR technique is thought to be relatively more capable (than the rest of ANNs) of estimating the missing annual total rainfall at 0228458. It can be noticed that 31.38 % of uncertainty can be removed from the subject station 0228458 via the standard BP. This technique is thought, in this study, to be a good estimator of missing annual total rainfall at 0228458.

Figures 6.21 (a-b) show the comparison of selected ANNs in terms their entropic values.

On the other hand, the values of RMSEp range from 46.861 mm to 48.316 mm of rainfall for the ANNs. The values of RMEp range from 0.1247 to 0.1168. The VLRBP technique has relatively the lowest values of RMSEp and RMEp while the standard BP has the highest value for the same statistical indicators. However, these values are close. These results enhance the conclusions drawn from the entropy calculations.

Figure 6.22 summarizes the hydrographs for different ANNs. This figure shows that the results are very close. Once again, this enhances the conclusions from entropy calculations.

In general, the techniques presented above appear to be good estimators of the missing annual total rainfall values (13.6 %) at gauge 0228458. For this gap duration, the above selected ANNs preserved the mean and the variance of the estimated series compared to the observed one as shown in Table 6.5.

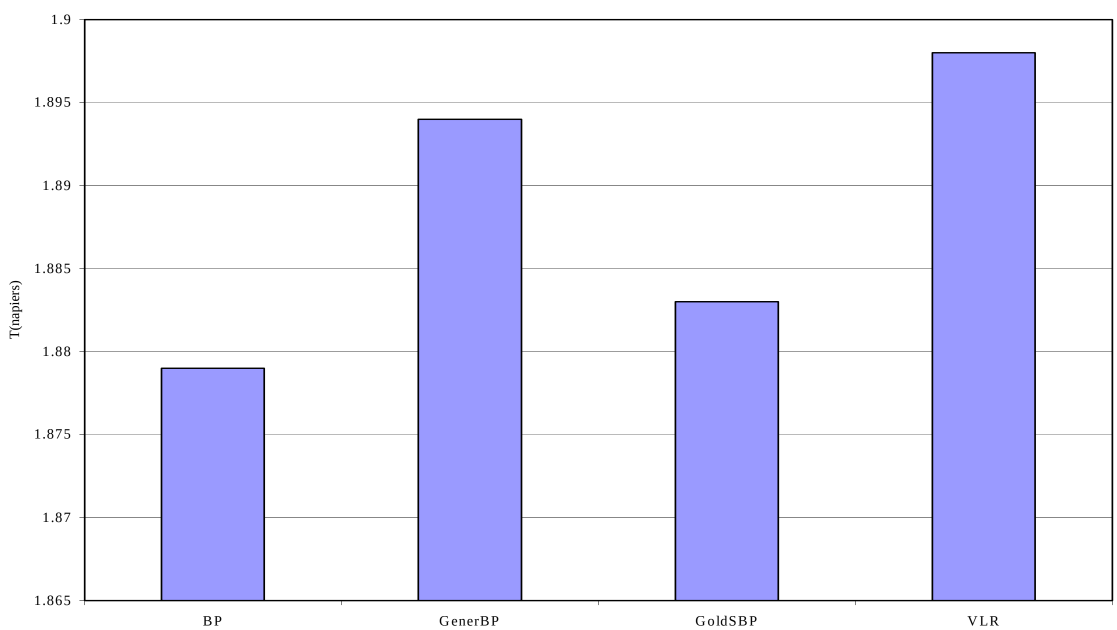


Figure 6.21a Comparison of ANNs in terms of T (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

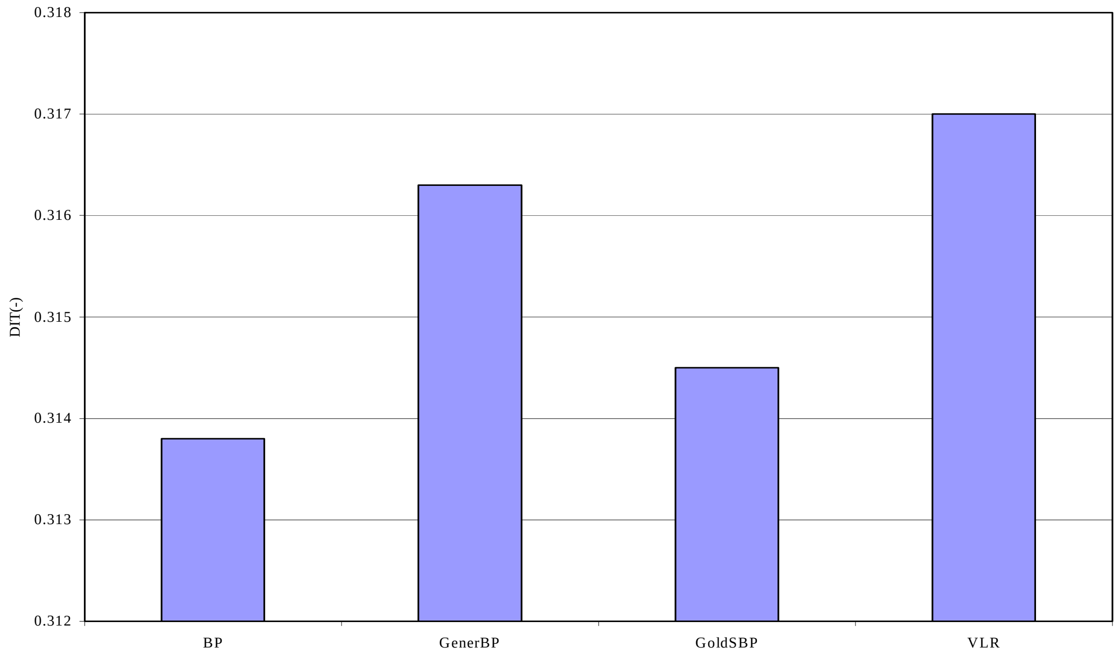


Figure 6.21b Comparison of ANNs in terms of DIT (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

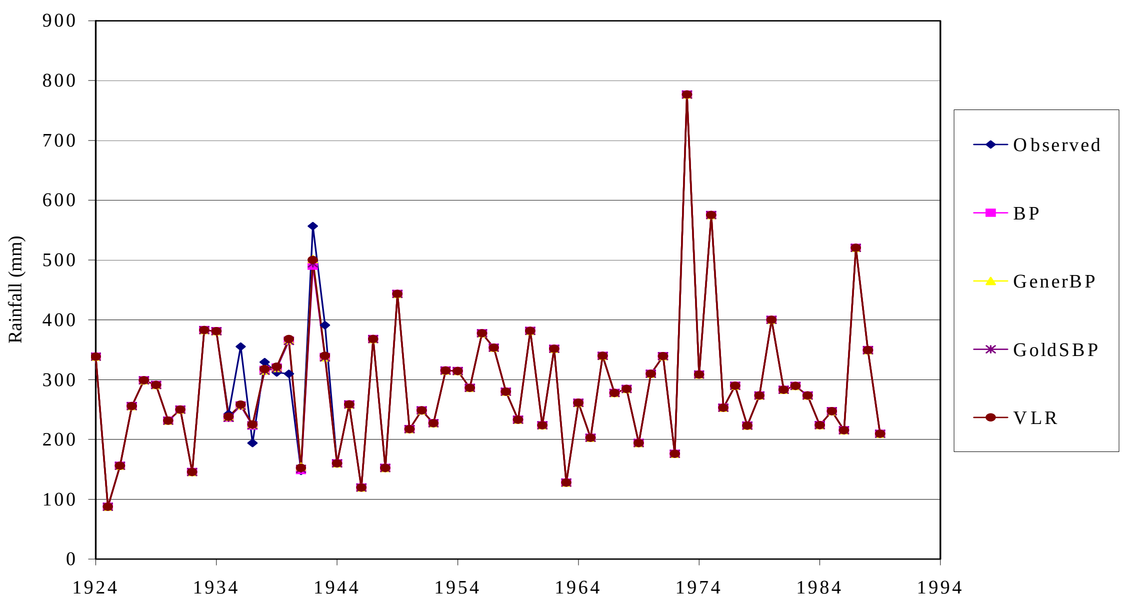


Figure 6.22 Comparison of ANNs in terms of hydrographs (13.6 % missing annual total rainfall from 1935 at 0228458) with base gauge 0228495

6.1.4.2.3 Comparison of performance of EM techniques at gauge 0228458 using base gauge 0228495

The selected EM techniques are thought to be promising estimators of the missing annual total rainfall values at 0228458 for gap duration of 13.6 % as long as they satisfy the entropy criterion. From the results shown in Table 6.5 and Figures 6.23 (a-b), it can be seen that the standard EM technique and its variants have almost identical entropic values although a very slight (negligible) improvement was noticed for MEM2 compared to the standard EM technique. It is believed that the standard EM leads almost 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 total rainfall (e.g. 13.6 %) at gauge 0228458.

The statistical indicators were also almost identical for all the EM techniques. These techniques preserve the mean and variance of the annual total rainfall data at 0228458.

Figure 6.24 summarizes the hydrographs of the different EM techniques. Hence, from a visual perspective, all the estimated hydrographs were almost identical.

Both the graphical comparison and the statistical indicators between the different EM techniques confirm the results obtained from the entropy calculations.

In an overall assessment, the EM techniques and its variants as presented here lead almost to the same solution and are thought to be promising estimators of the missing annual total rainfall at station 0228458 for gap duration of 13.6 %.

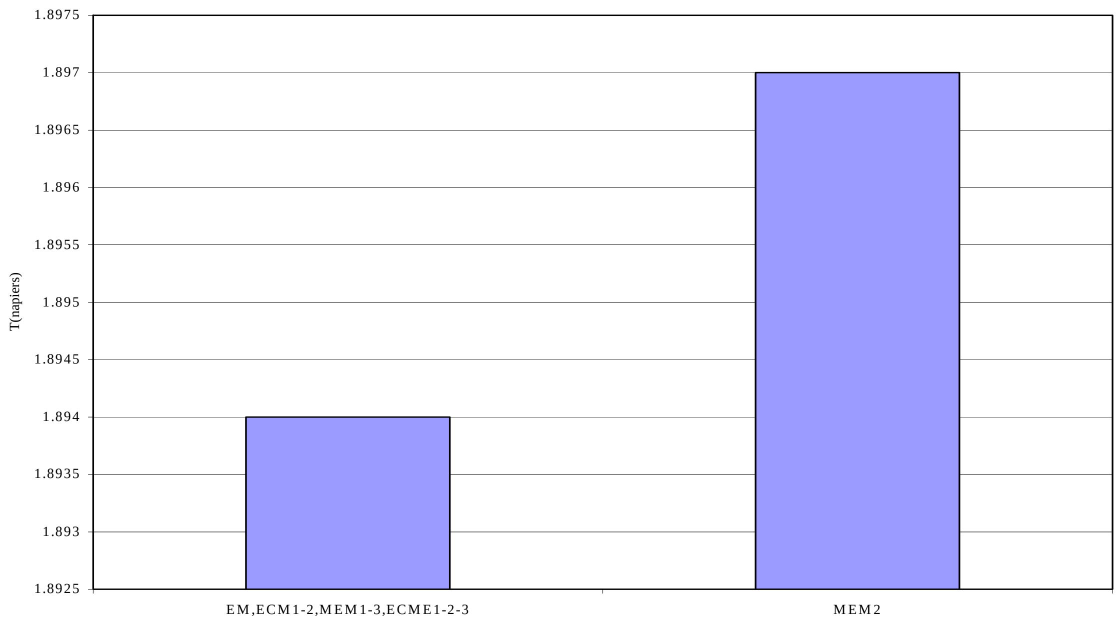


Figure 6.23a Comparison of EM techniques in terms of T (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

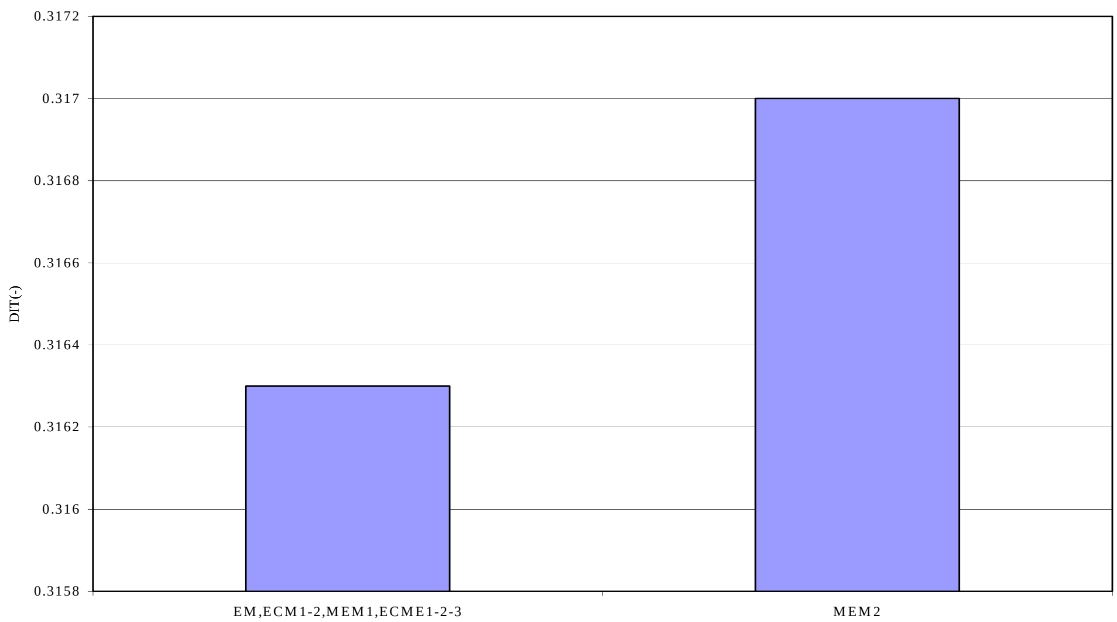


Figure 6.23b Comparison of EM techniques in terms of DIT (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

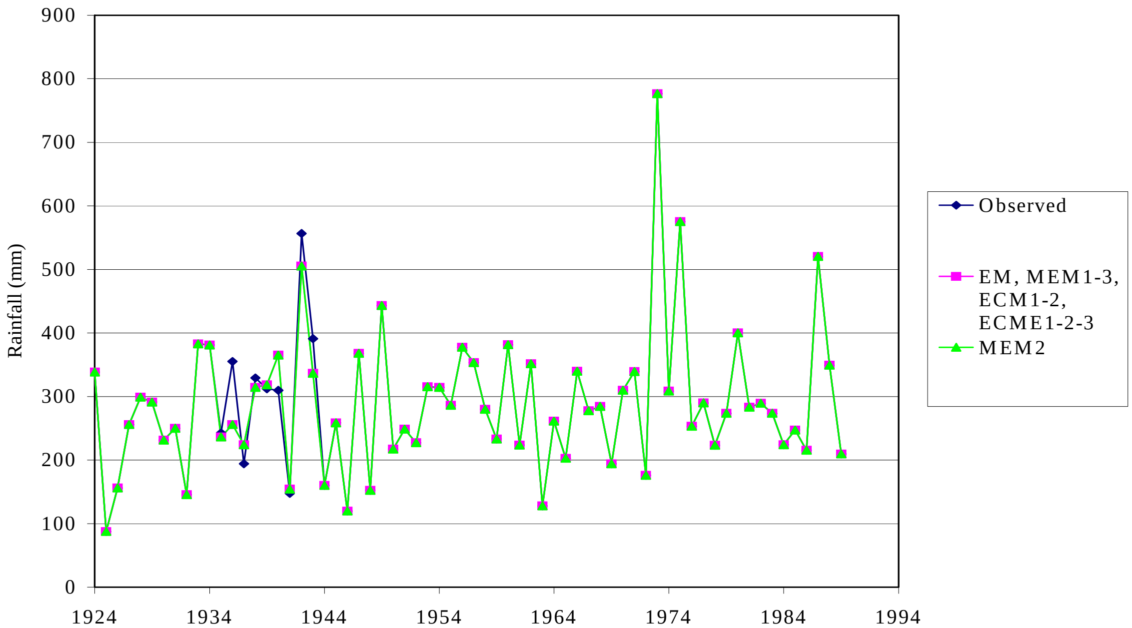


Figure 6.24 Comparison of EM techniques in terms of hydrographs (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

6.1.4.2.4 Comparison of performance of ANN and EM techniques at gauge 0228458 using base gauge 0228495

Comparison between these techniques is made in terms of DIT (T) and is exhibited in Figures 6.25 (a-b). For visual convenience, only the first best two ANN techniques and all the EM techniques are presented here.

The EM techniques come in the second after VLR. Thus, the EM techniques show good estimation capabilities for the missing annual total rainfall at 0228458. This may be due to the presence of a strong linear relationship in the rainfall data between gauges 0228458 and 0228495 since the EM techniques used in this study were developed within the linear regression context. Only VLR technique showed slightly a higher capability (than the EM techniques) of mapping the non-linear characteristics of the missing annual total rainfall.

The difference in entropic values between these ANNs and the EM techniques were small.

The graphical comparison (hydrographs) between the EM and the first best two ANNs techniques for the gauge 0228458 is exhibited in Figure 6.26. The visual observation reveals that the hydrographs are very close.

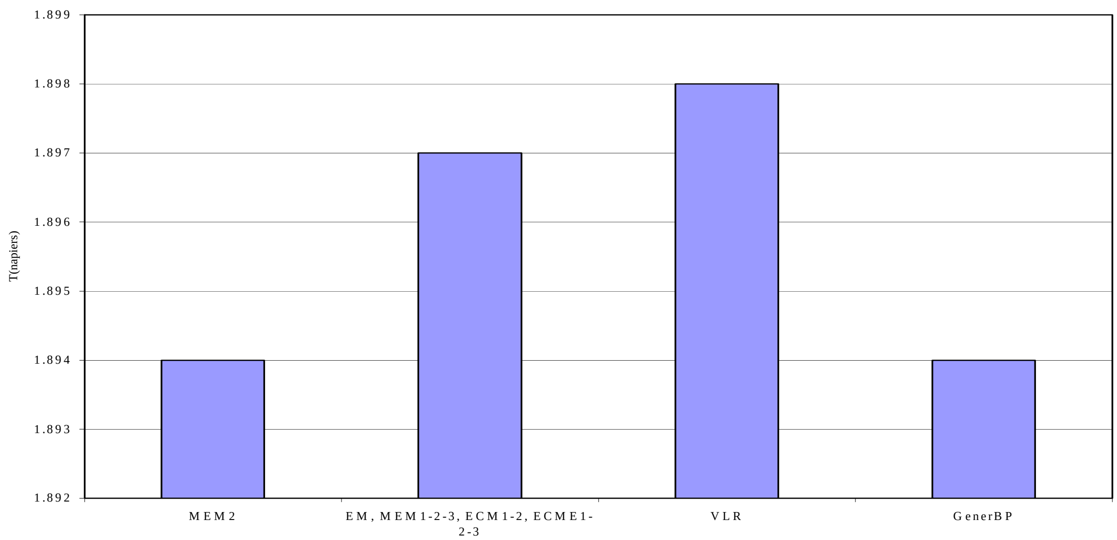


Figure 6.25a Comparison of ANN and EM techniques in terms of T (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

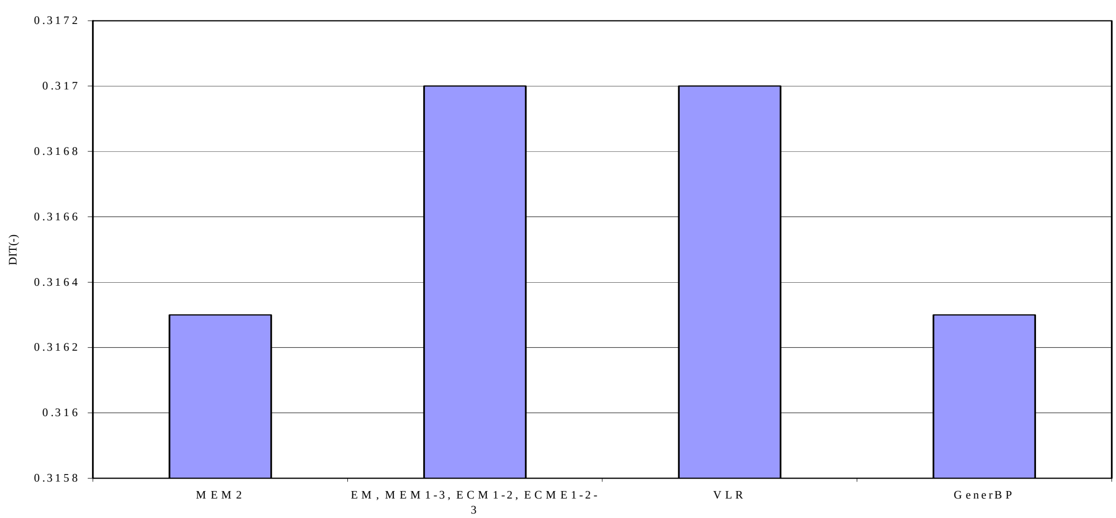


Figure 6.25b Comparison of ANN and EM techniques in terms of DIT (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

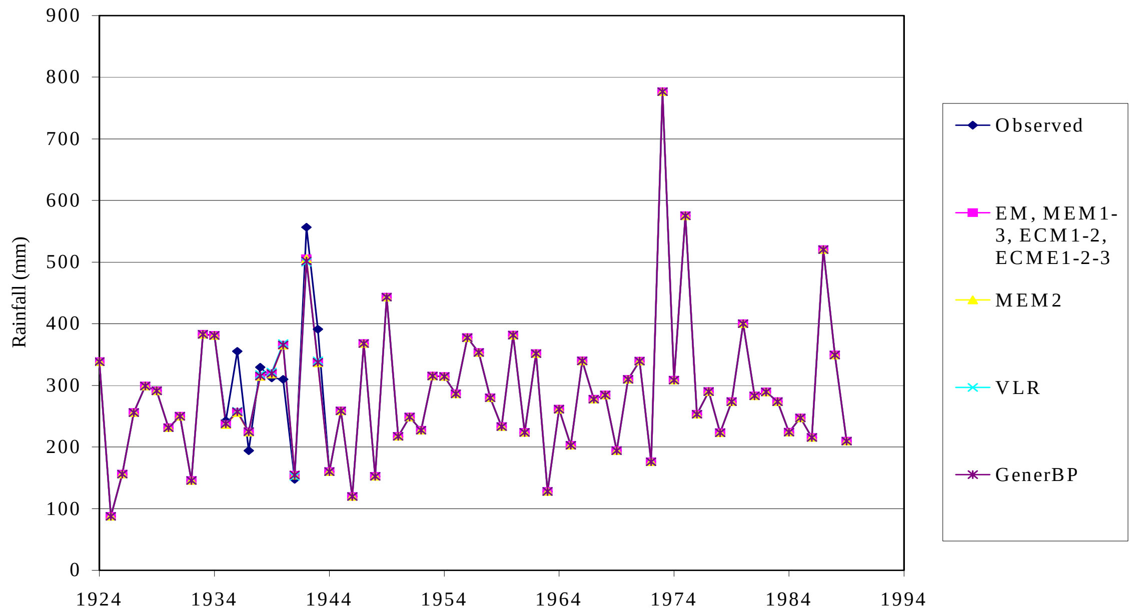


Figure 6.26 Comparison of ANN and EM techniques in terms of hydrographs (13.6 % missing annual total rainfall from 1935 at 0228458) using base gauge 0228495

Considering the different values of DIT, the rainfall data infilling techniques could be also compared for the different catchment areas. In general, for a given gap size (duration), the overall value of DIT for the station pair 0228495-022858 was not quite different from the one for the station pair 0228458-0228495. Hence, station 0228458 (0228495) was a good predictor for the estimation of missing annual total rainfall at 0228495 (0228458).

6.2 SUMMARY

The entropy concept was shown to be a versatile tool. Firstly, it enables to know the information content of the rainfall gauges. Secondly, it enables the selection of base/target station. Lastly, it is also used to assess technique performance when applied to annual total rainfall data infilling. Recall that these data are used in design-oriented studies of hydraulic structures such as reservoirs, bridges, etc.

It was shown that the gauges within each respective station-pair (i.e. 0228170-0228495 and 0228495-0228458) could infer mutually the information (contained in the annual total rainfall series) about one another. For each station pair, when one gauge could be considered as predicted station, the other one was predictor station and vice-versa.

The results from the rainfall station pairs 0228495-0228458 and 0228458-0228495 were the only ones to be discussed. The results related to the station-pairs 0228170-0228495 and 0228170-0228495 were just depicted in the Appendix C since they were similar to those obtained from the former two station-pairs. For the rainfall station pair 0228495-0228458 and 0228458-0228495, the gap duration of 7.6 % and 13.6 % (from 1935) in annual total rainfall at either station were discussed. This was done so, as the entropy criterion was satisfied with regard to the performance of the different rainfall data infilling techniques. The results from the rest of the gap duration (e.g. 20 % and 30 %) were not satisfactory results and were not therefore discussed. Nonetheless, the results from these gap sizes (duration) were just used to investigate the relationship between gap duration and accuracy of estimated values for the different techniques.

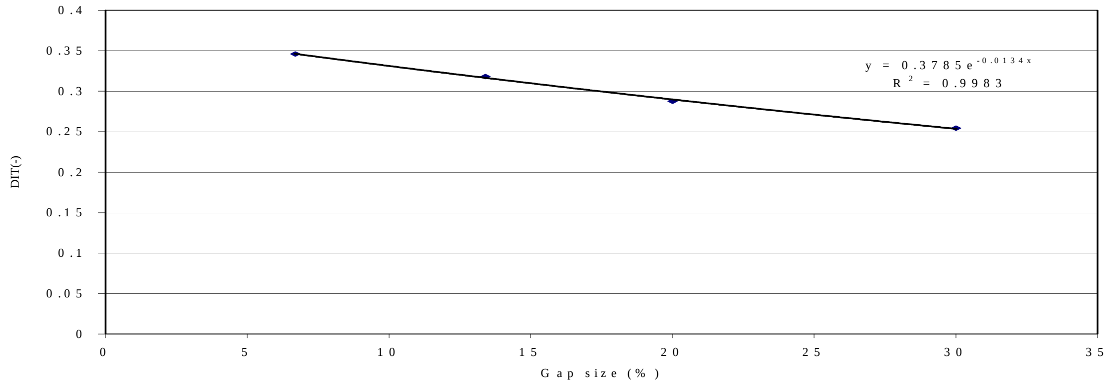
Considering the different values of DIT, rainfall data inflling techniques could be also compared for different catchment areas. In general, for a given gap size, it could be shown that station 0228458 (0228495) was a good predictor for the estimation of missing annual total rainfall at 0228495 (0228458).

It was also noticed that the directional information transfer index between observed and estimated values generally decreases when the proportion of missing annual rainfall increases. In other words, the proportion of information transferred by the knowledge of the estimated series into the process to make the observed annual total rainfall data better defined, will decrease as the gap duration increases at the subject station.

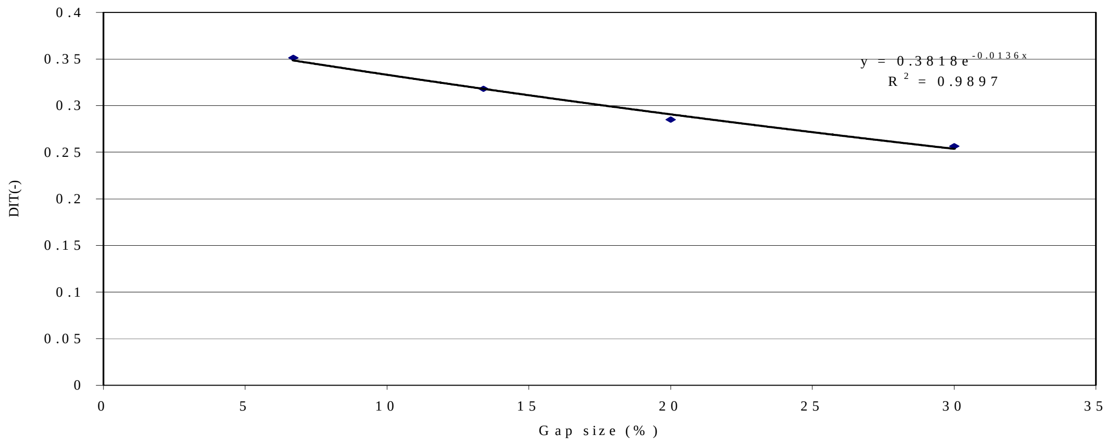
Generally, a decay power equation could describe better (than exponential or linear) the relationship between the gap duration (ranging from 7.6 % to 30 %) and the technique accuracy in terms of the directional informational transfer index. Figures 6.27 (a-d) and Figures 6.28 (a-d) give just an illustration of that relationship. Thus, for a given technique, it is possible to find approximately the expected accuracy of the estimated values at the subject station when the gap size is known. Increasing the number of data points (e.g. up to seven values of gap size at the target gauge) did not sensitively affect the above-mentioned relationship. It was also noticed that an earlier start (e.g. at 1930) or later start (e.g. at 1970) for the gaps created on the records of the subject station 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 selected station pair involved in the estimation process of missing annual total rainfall.

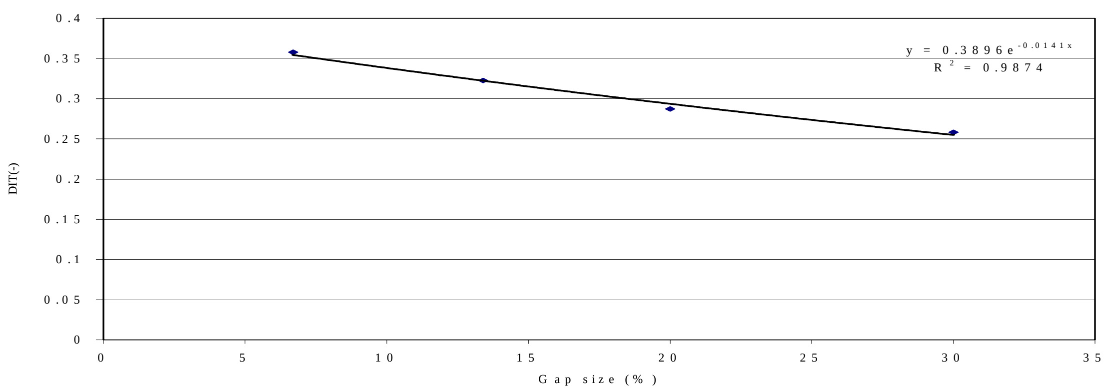
An overall assessment of the different rainfall data infilling techniques revealed that the standard BP and standard EM as well as their respective variants showed good capabilities of estimating the annual total rainfall at the target gauge 0228495, when considering the base gauge 0228458 (and vice-versa). 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 in 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).



(a)

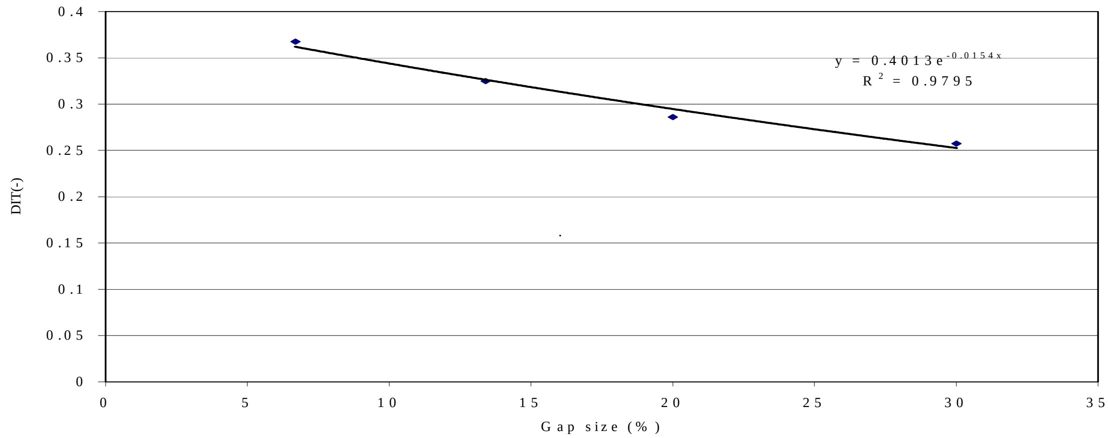


(b)



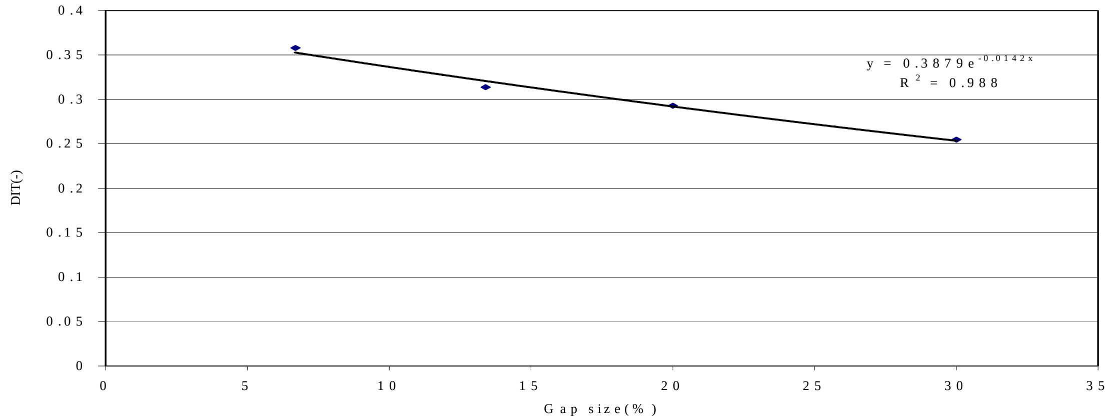
(c)

Figures 6. 27 (a-c) DIT versus gap size for annual total rainfall at 0228495 (base gauge 0228458): (a) BP, (b) GoldSBP, (c) GenerBP

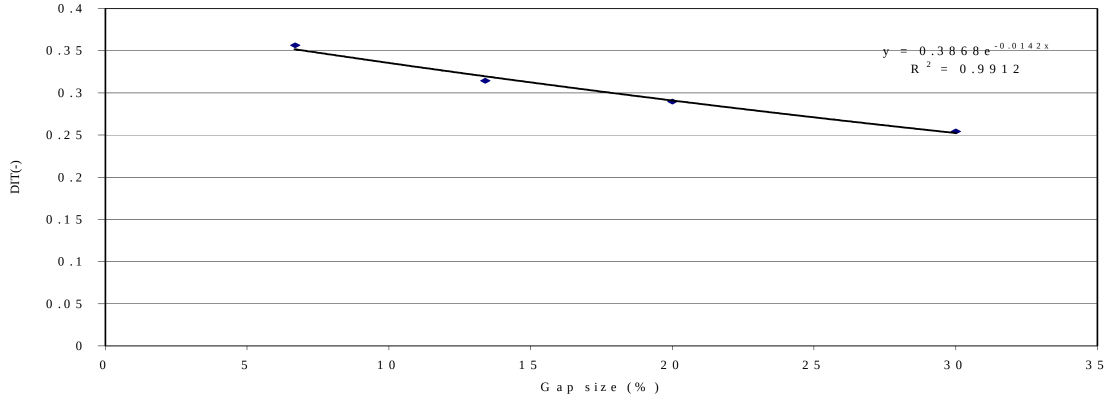


(d)

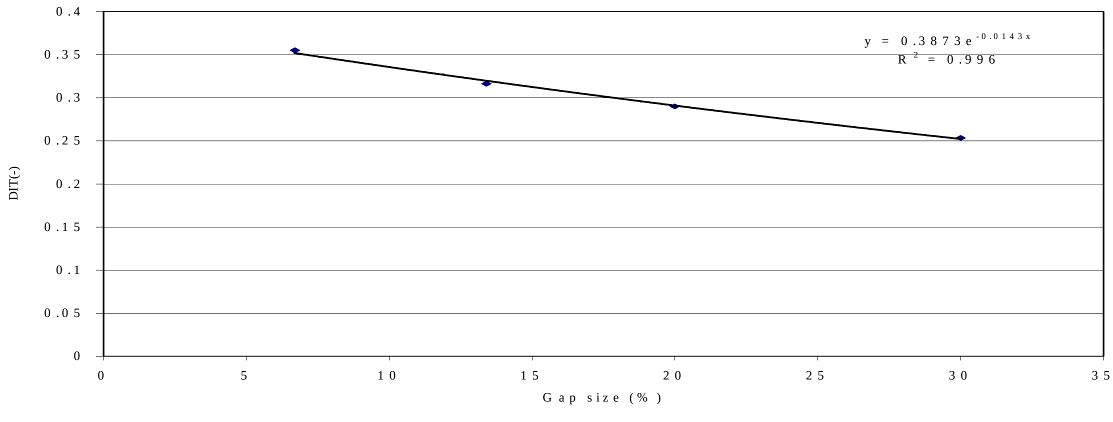
Figure 6.27d DIT versus gap size for annual total rainfall at target gauge 0228495 (base gauge 0228458): (d) EM



(a)

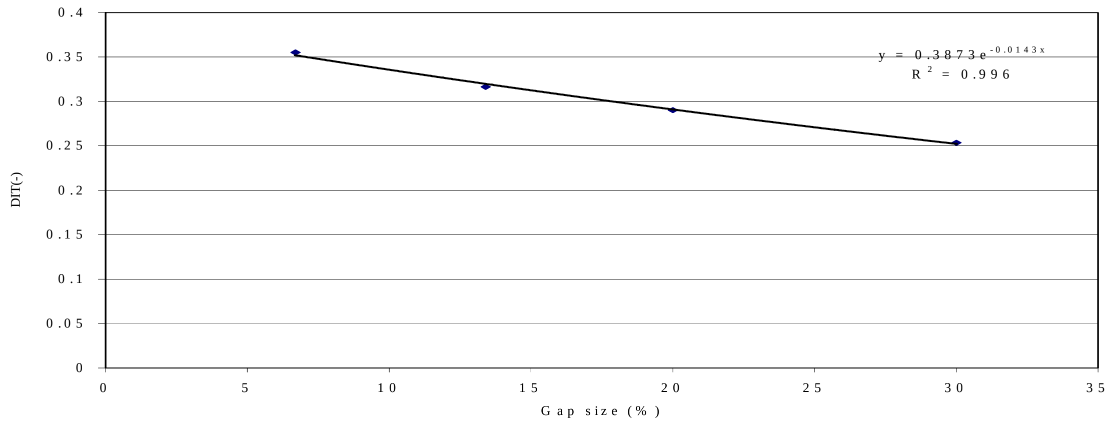


(b)



(c)

Figures 6.28 (a-c) DIT versus gap size for annual total rainfall at target gauge 0228458 (base gauge 0228495): (a) BP, (b) GoldSBP, (c) GenerBP



(d)

Figure 6.28d DIT versus gap size for annual total rainfall target gauge 0228458 (base gauge 0228495): (d) EM