

4. EVALUATION OF THE PERFORMANCE OF THE EMD-VLB AND ITS COMPARISON WITH THE VLB GENERATOR

4.1 Introduction

To evaluate the performance of the developed generator, it is necessary to determine its ability to replicate historic statistics. The effect of using EMD in block selection needs to be assessed by comparing the performance of the coupled EMD-VLB generator and the standard VLB generator. These tasks are carried out using the multi-site rainfall generation problem that uses the data set from 10 stations described previously in Section 3.2. In the assessment, the closeness of the historic statistics to the generated (stochastic) ones for both generators are determined and plotted. In this way, the assessment of performance and the comparison of the hybrid EMD-VLB with the standard VLB are done simultaneously. These statistics used at both the annual and the monthly time scale are; the mean, the median, the 25th percentile, the 75th percentile, the lowest and the highest rainfall, the standard deviation, and the skewness. In addition, the respective minimum run sums for annual lengths of 1-24 years are also computed and compared for both generators. The minimum run sums are necessary to assess the replication of long-term dependence characteristics (Ndiritu and Nyaga, 2014).

The two generators are used to generate 101 rainfall sequences of a length equal to that of the historic sequences (93 years) and the performance and comparison is based on these stochastic sequences. Whereas the original VLB set a minimum block length of 3 years, the EMD starts block generation from the second IMF generated from an original time series so as to eliminate mode mixing in which different hydrological occurrences could be quantified together if selection started from IMF1 or the original time series.

Since any difference between the hybrid EMD-VLB and the standard VLB would result from differences in the way the blocks are obtained, a comparison of the characteristics of the blocks from the two methods is done in Section 4.2. The two characteristics considered are the length of block and the scaled mean

annual precipitation (MAP) of the blocks. The performance and comparisons in replicating annual and monthly statistics then follows in Section 4.3 and Section 4.4 summarises the Chapter.

The summarized number of IMFs per station after decomposition is shown in Table 4.1 while the plots of the resulting IMFs are shown in appendix B. Table 4.1 also shows the percentage losses from the decomposition that have been used to verify the accuracy of decomposition. This verifies EMD as a lossless decomposition in which the completeness of the decomposition is demonstrated (Molla et al., 2005). All the 10 losses are negligible thus giving confidence that the decomposition was valid.

Table 4.1 Summary of resultant number of IMFs generated for each station

Station	0258894	0555857	0474295W	0320848	0240891	0142805	0052990	0020806	0149082	0678776
	W	W		W	W	W	W	W	W	W
Number of IMFs	5	4	5	4	4	5	4	5	6	6
% loss in decomposition	7.70E-06	4.01E-06	5.75E-06	1.01E-07	1.28E-05	8.77E-06	1.19E-05	2.85E-06	5.09E-06	1.00E-05

4.2 Length and MAP of blocks generated by the hybrid EMD-VLB and by the VLB model

The scaled mean annual precipitation (MAP) of each block is obtained as,

$$\text{Scaled MAP (mm)} = \frac{\text{MAP of block (mm)}}{\text{MAP of complete historic sequence (mm)}} \quad (4.1)$$

Figures 4.1 and 4.2 show the number of blocks obtained for various lengths of blocks that were applied in the stochastic generation. In addition, a plot in which the number of blocks is multiplied by the block length is included in the two Figures. This plot indicates the real significance of the specific length as a block

once selected takes up its length in the stochastic rainfall sequence. The plots show some distinct similarities and differences. The number of blocks used generally decreases as block length increases but the hybrid model has 2-year long blocks in much smaller numbers than 3-year long ones. The hybrid has the longest block length of 40 years while the VLB has the longest block length of 37 years which is only slightly shorter. The number of blocks also peak locally at 22 years for the hybrid and at a slightly shorter length of 20 years for the VLB. Because the VLB defines the condition for block termination stochastically within a set range while the hybrid uses specific lengths obtained from the IMFs of the EMD, the total numbers of blocks for the VLB are much greater than that those of the EMD-VLB as Figures 4.1 and 4.2 show.

Figures 4.3 and 4.4 show plots of the average scaled MAP for the hybrid and the VLB respectively. As expected, the scaled MAPs generally take values close to 1.0. The exception is mainly for the low lengths where the hybrid EMD-VLB obtained higher values while the VLB obtains lower values. The VLB terminates blocks at low rainfall periods and so the very short blocks will tend locate entirely multi-annual low rainfall periods and will therefore have low MAPs. The reason why the shortest lengths of the hybrid are distinctly higher is not known yet but is it considered possible that the short lengths are picking up the high peaks while missing on the lowest values that tend to locate in the middle of multi-annual low rainfall periods.

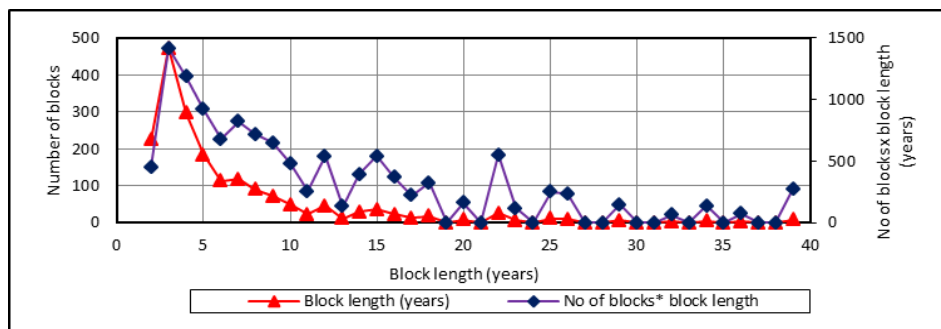


Figure 4.1 Number of blocks generated and the total length of blocks of specified length from the EMD-VLB generator.

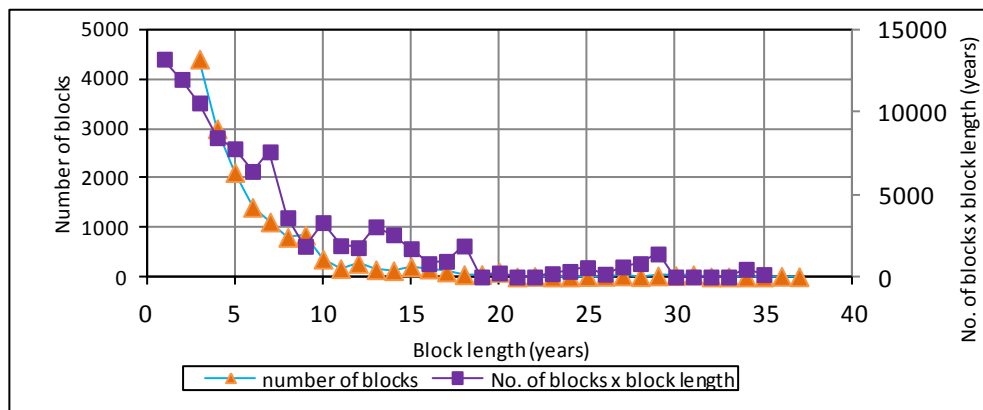


Figure 4.2 Number of blocks generated and the total length of blocks of specified length from the VLB generator.

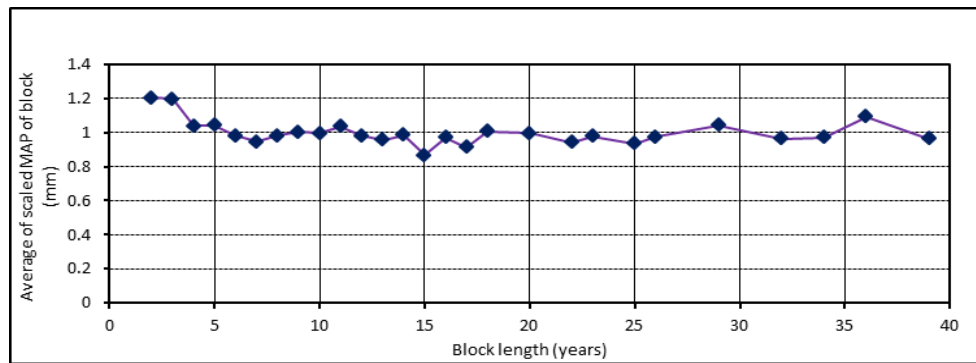


Figure 4.3 Scale average MAP of rainfall from the EMD-VLB generator.

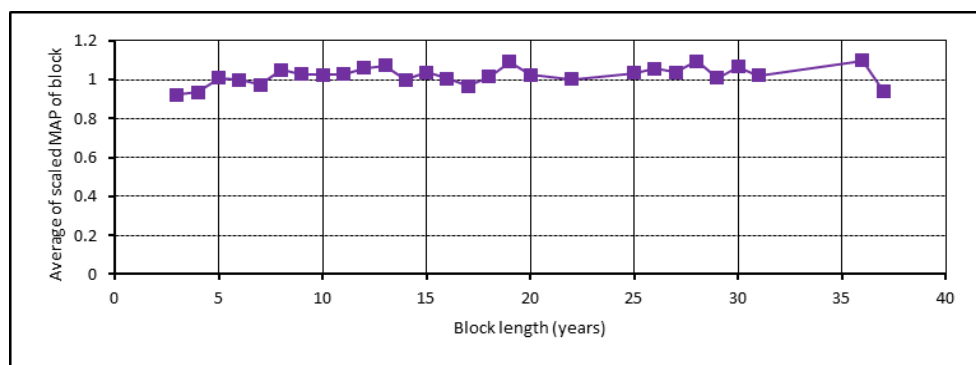


Figure 4.4 Scale average MAP of rainfall from the VLB generator.

4.3 Evaluation of the performance of the EMD-VLB generator and its comparison with the VLB

Evaluation of performance and comparison of the two generators is mainly done using box plots of the generated statistics. Figure 4.5 defines the percentiles that are used in the box plots. As a general guideline, the replication of a statistic is considered satisfactory if the historic value falls within the inter-quartile range of the box plot. The box plots also help to assess the variability and to identify the range of the generated statistics.

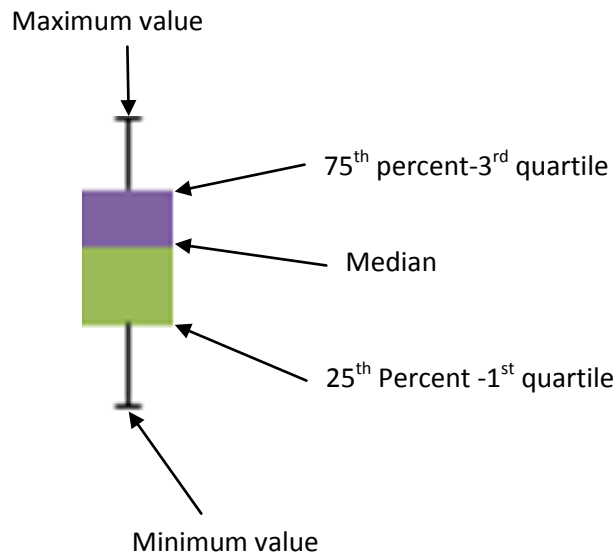


Figure 4.5 Box plot legend representing the various quartiles as used in the stochastic simulations.

4.3.1 Evaluation and comparison of EMD-VLB and VLB generators using minimum run sums

Figures 4.6 to 4.15 present the box plots of minimum run sums for the 10 rainfall sites from the EMD-VLB and the VLB model. The single lines across the box plots (with thin rectangular blocks) indicate the historic minimum run sums. For each site, the first figure is a box plot for the EMD-VLB and the second for the VLB.

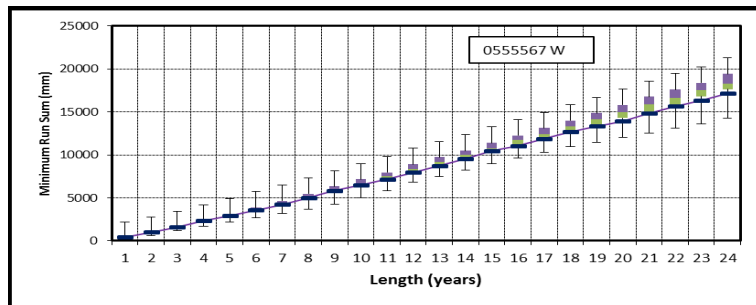


Figure 4.6a Box plots of minimum run sums for station 0555567 W from EMD-VLB generator

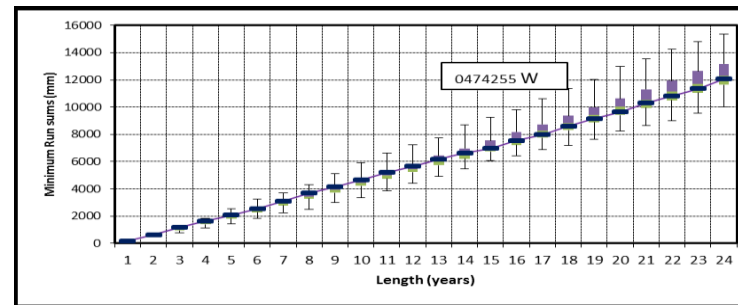


Figure 4.7a Box plots of minimum run sums for station 0474255 W from EMD-VLB generator

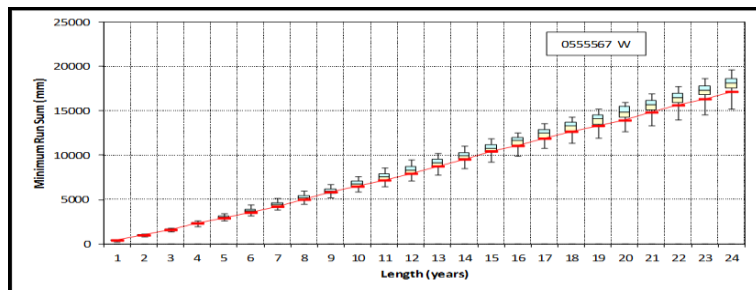


Figure 4.6b Box plots of minimum run sums for station 0555567 W from VLB generator

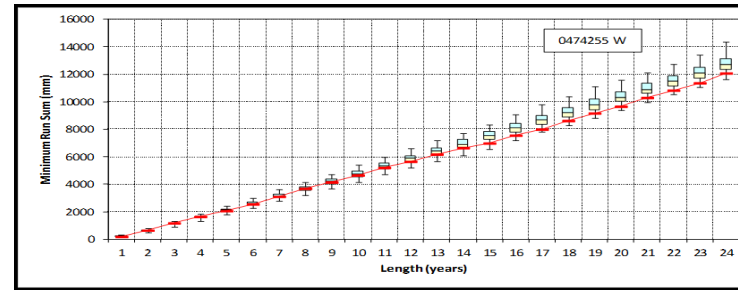


Figure 4.7b Box plots of minimum run sums for station 0474255 W from VLB generator

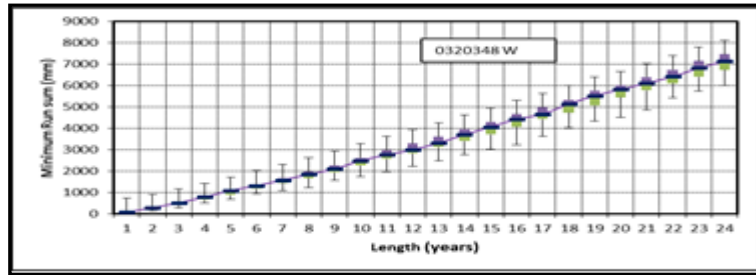


Figure 4.8a Box plots of minimum run sums for station 0320348 W from EMD-VLB generator

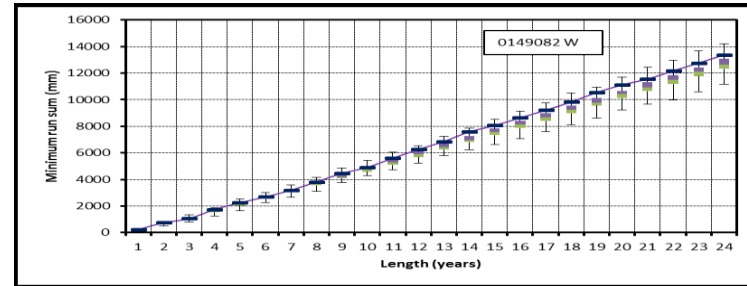


Figure 4.9a Box plots of minimum run sums for station 0149082 W from EMD-VLB generator

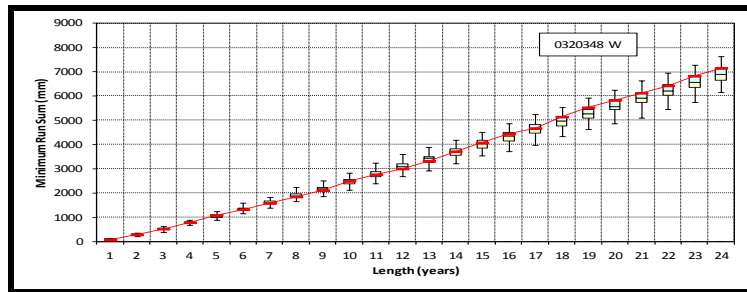


Figure 4.8b Box plots of minimum run sums for station 0320348 W from VLB generator

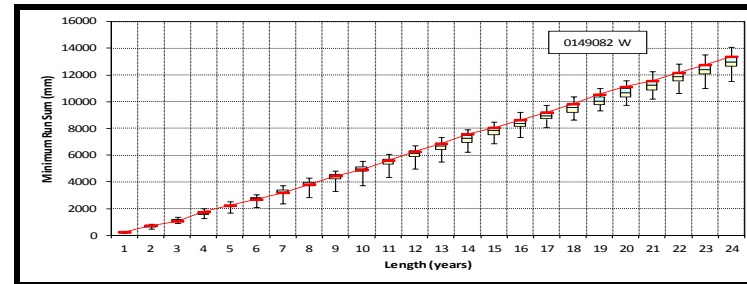


Figure 4.9b Box plots of minimum run sums for station 0149082 W from VLB generator

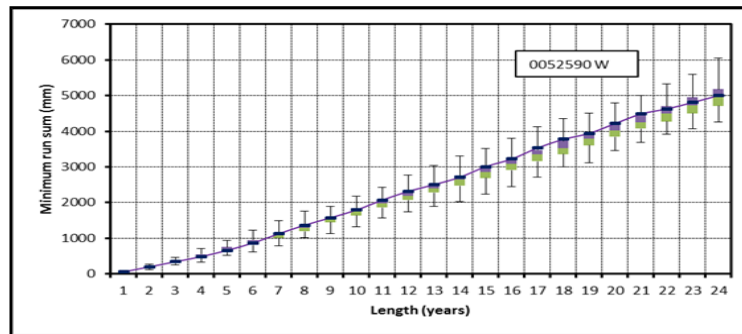


Figure 4.10a Box plots of minimum run sums for station 0052590 W from EMD-VLB generator

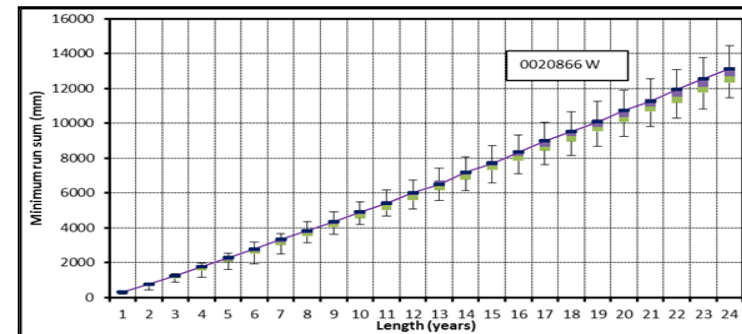


Figure 4.11a Box plots of minimum run sums for station 0020866 W from EMD-VLB generator

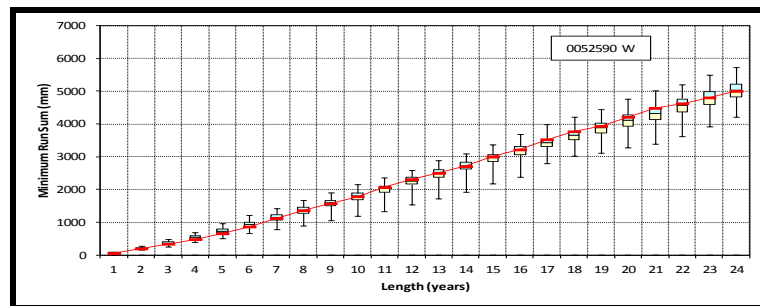


Figure 4.10b Box plots of minimum run sums for station 0052590 W from VLB generator

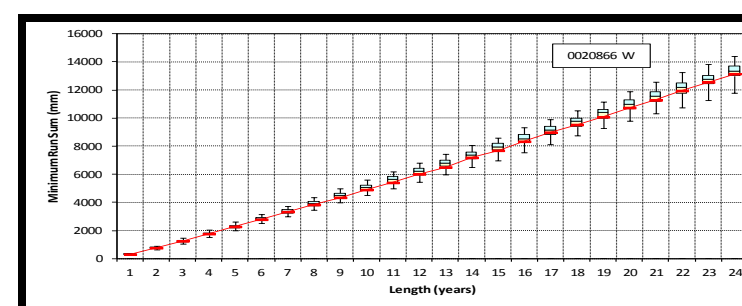


Figure 4.11b Box plots of minimum run sums for station 0020866 W from VLB generator

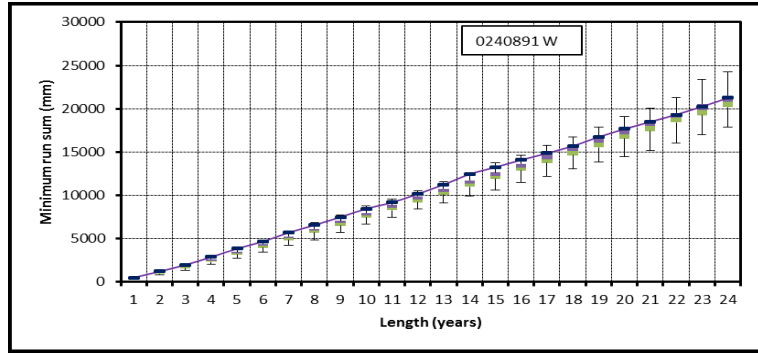


Figure 4.12a Box plots of minimum run sums for station 0240891 W from EMD-VLB generator

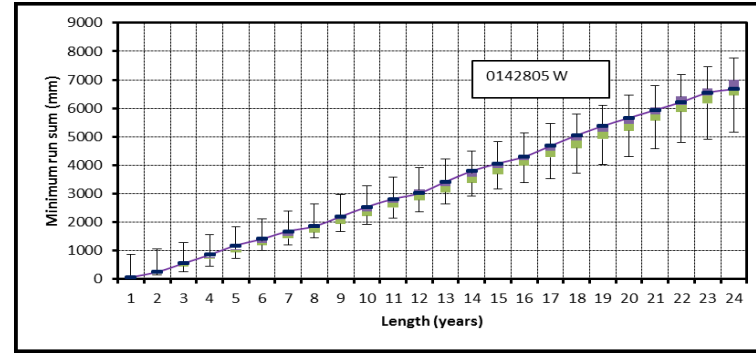


Figure 4.13a Box plots of minimum run sums for station 0142805 W from EMD-VLB generator

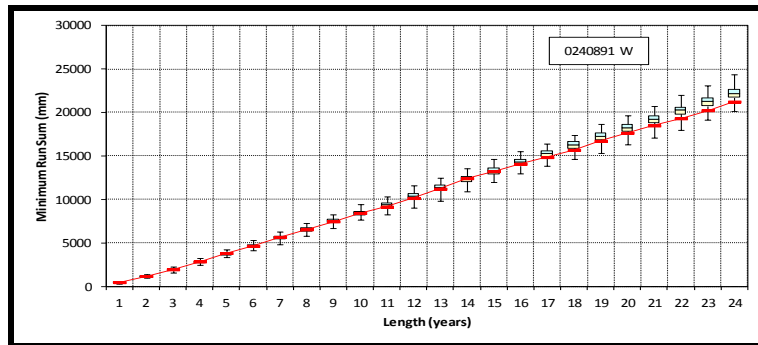


Figure 4.12 b Box plots of minimum run sums for station 0240891 W from VLB generator

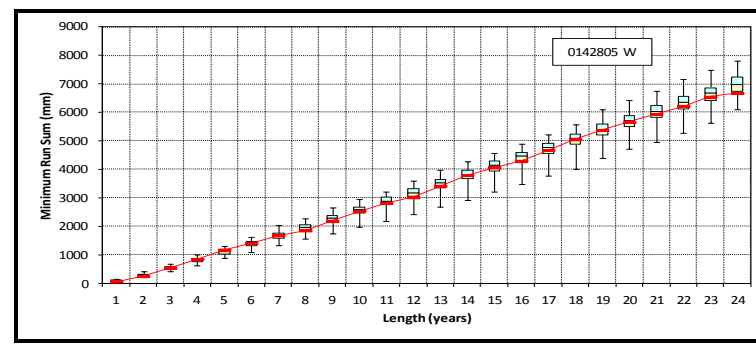


Figure 4.13b Box plots of minimum run sums for station 0142805 W from VLB generator

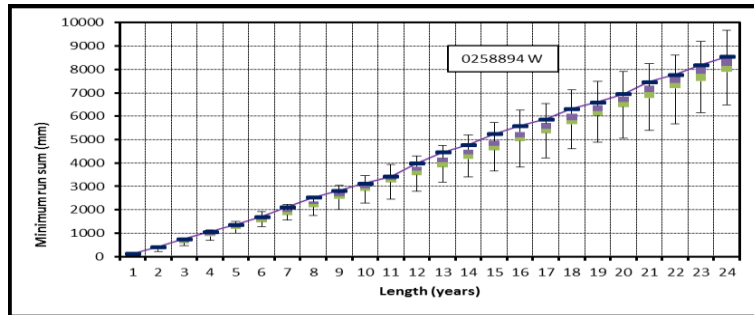


Figure 4.14a Box plots of minimum run sums for station 0258894 W from EMD-VLB generator

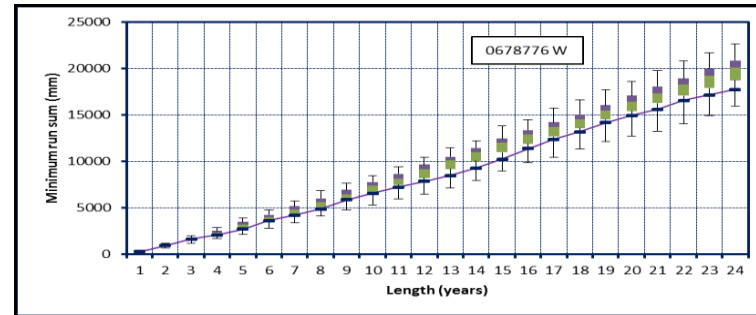


Figure 4.15a Box plots of minimum run sums for station 0678776 W from EMD-VLB generator

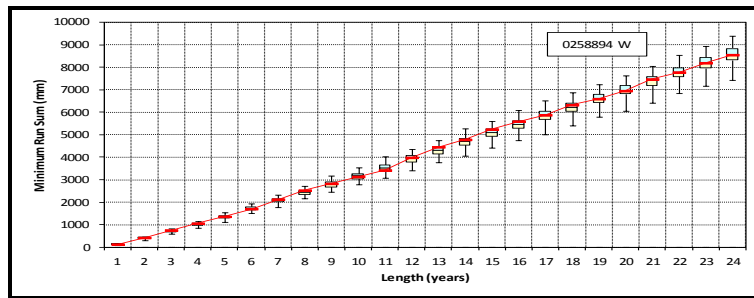


Figure 4.14b Box plots of minimum run sums for station 0258894 W from VLB generator

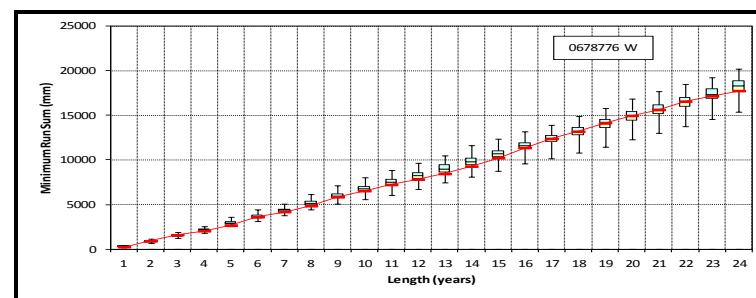


Figure 4.15b Box plots of minimum run sums for station 0678776 W from VLB generator

From the box plots on Figures 4.6 to 4.15, it is found that both generators generally replicate the historic minimum run sums reasonably well. The VLB however locates more of the historic run sums within the inter-quartile range than the EMD-VLB. The EMD-VLB generator has 70.8 % while the VLB has 95% of the historic run sums falling within the inter-quartile range of box plots of the simulated run sums. All the historic values lie within the box plot limits and the two generators are considered to have the ability to replicate long term dependence characteristics. It is also observed that the EMD-VLB model obtained box plots that are more widely spaced than the VLB. This difference is significant as the high exceedance statistics that are commonly used in water resources assessment) may be substantially different for the run sums from the two generators.

4.3.2 Evaluation and comparison of EMD-VLB and VLB generators using annual statistics

Figure 4.16 to 4.23 present the box plots of the 10 stations obtained from the hybrid EMD-VLB and the VLB generator.

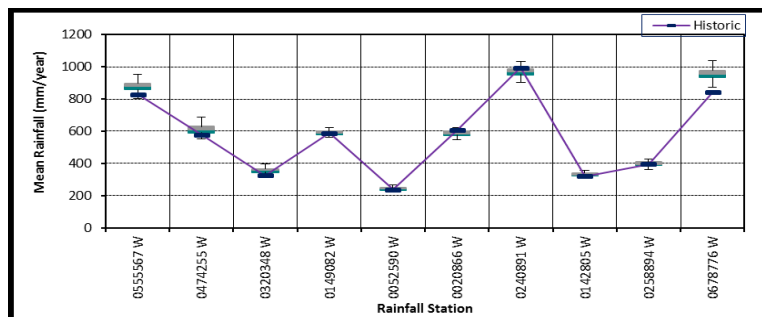


Figure 4.16a Box plots of annual mean rainfall from EMD-VLB generator

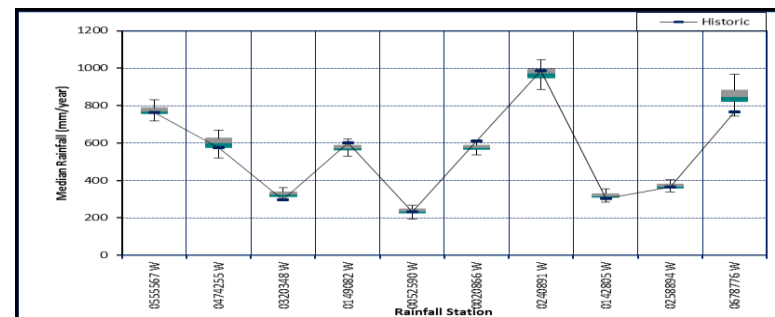


Figure 4.17a Box plots of annual median rainfall from EMD-VLB generator

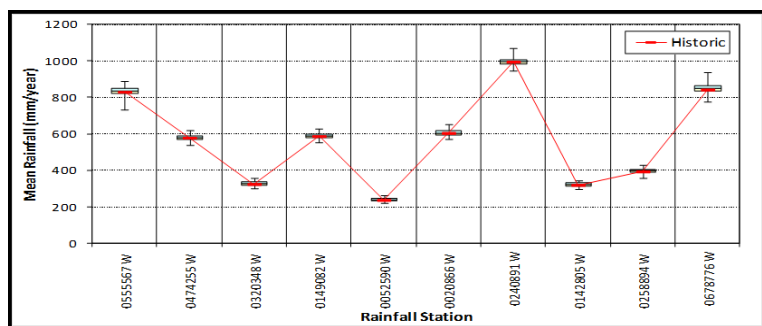


Figure 4.16b Box plots of annual mean rainfall from VLB generator

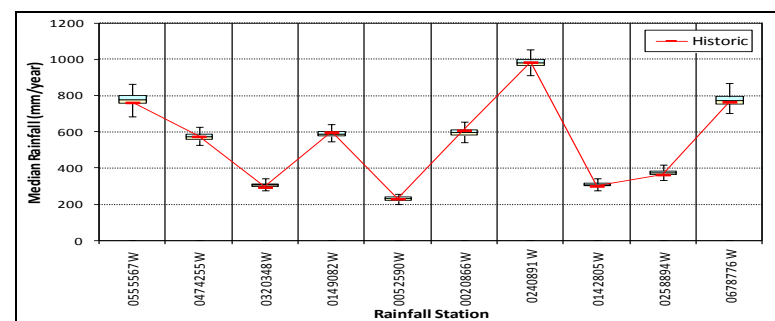


Figure 4.17b Box plots of annual median rainfall from VLB generator

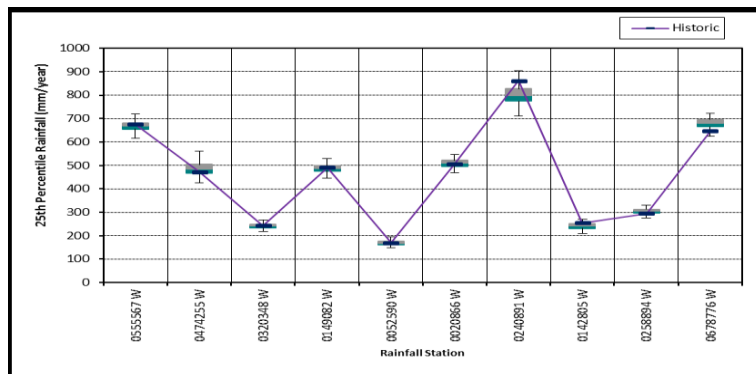


Figure 4.18a Box plots of annual 25th percentile rainfalls from EMD-VLB generator

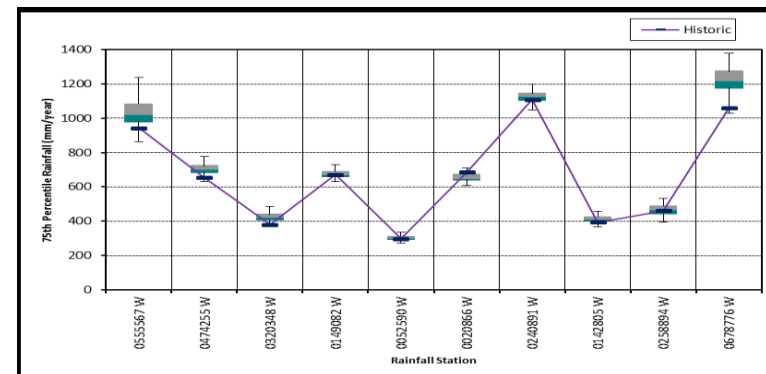


Figure 4.19a Box plots of annual 75th percentile rainfalls from EMD-VLB generator

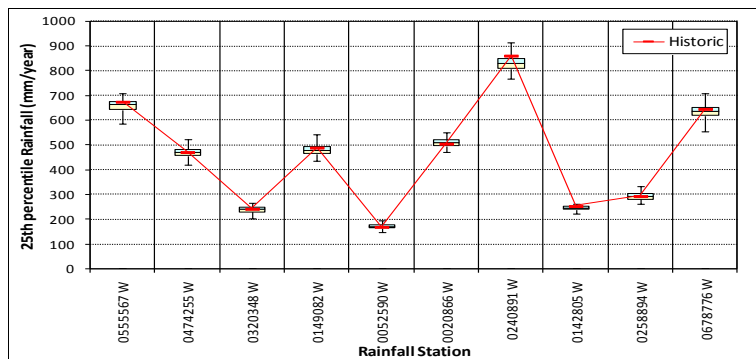


Figure 4.18b Box plots of annual 25th percentile rainfalls from VLB generator

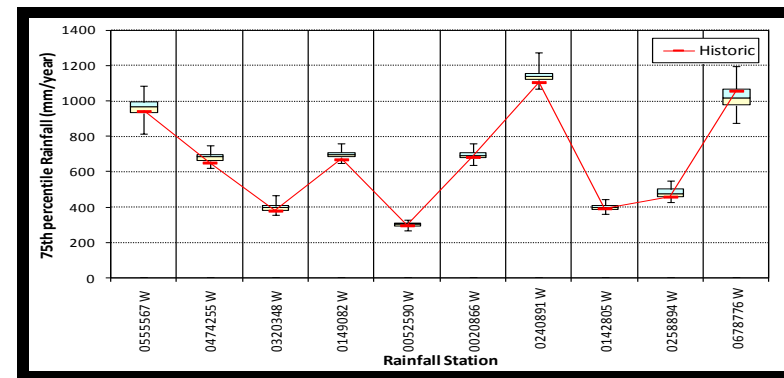


Figure 4.19b Box plots of annual 75th percentile rainfalls from VLB generator

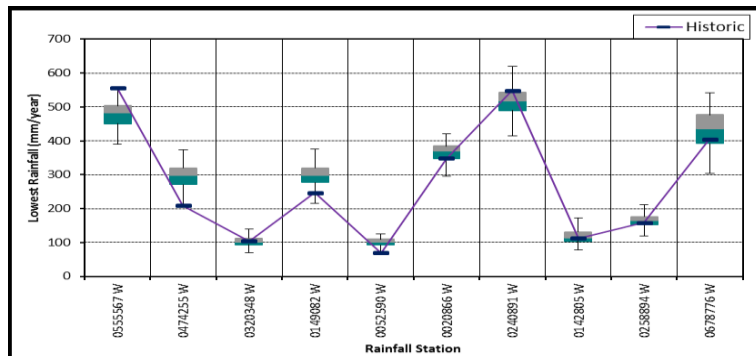


Figure 4.20a Box plots of lowest annual rainfalls from EMD-VLB generator

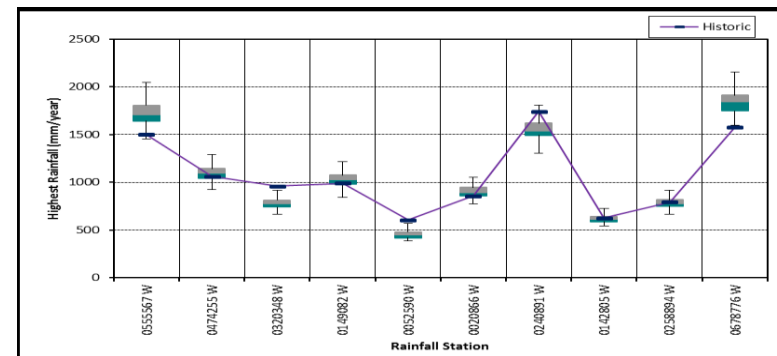


Figure 4.21a Box plots of highest annual rainfalls from EMD-VLB generator

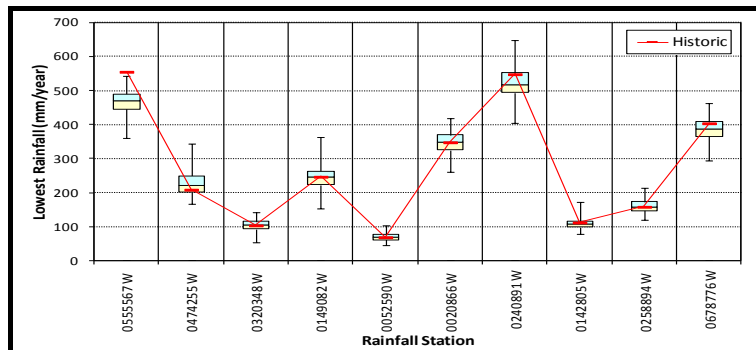


Figure 4.20b Box plots of lowest annual rainfalls from VLB generator

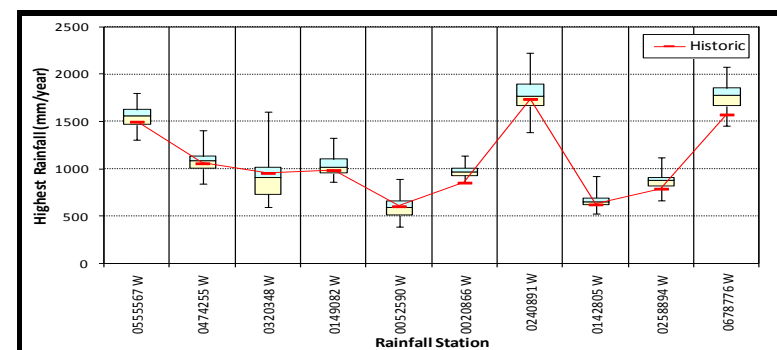


Figure 4.21b Box plots of highest annual rainfalls from VLB generator

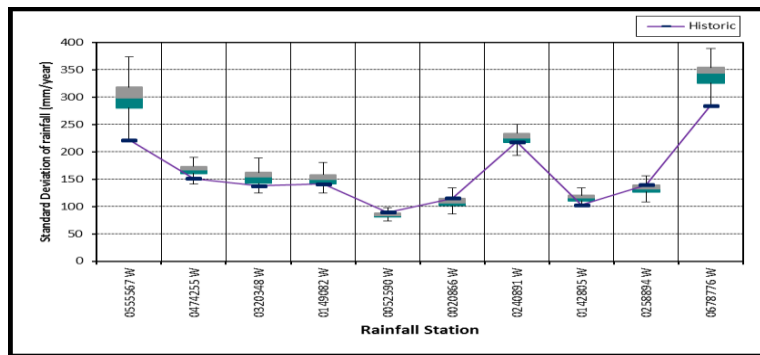


Figure 4.22a Box plots of standard deviation of annual rainfall from EMD-VLB generator

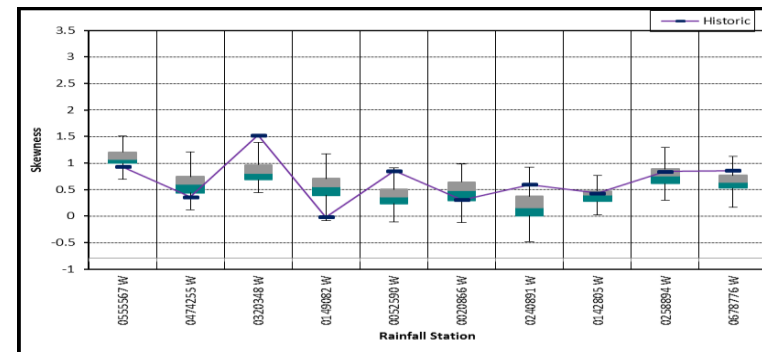


Figure 4.23a Box plots of skewness of annual rainfalls from EMD-VLB generator

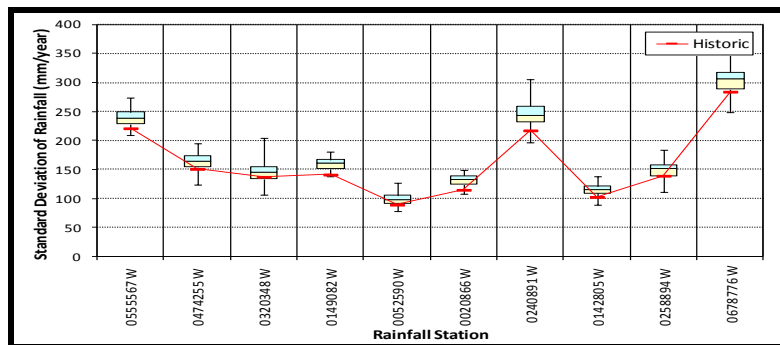


Figure 4.22b Box plots of standard deviation of annual rainfall from VLB generator

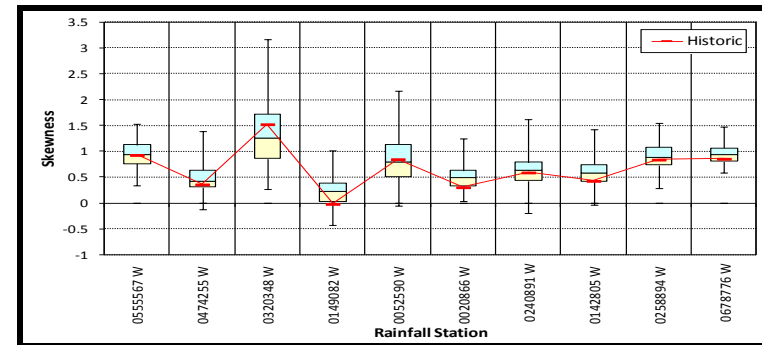


Figure 4.23b Box plots of skewness of annual rainfalls from VLB generator

The box plots in Figures 4.16 – 4.23 reveal that both generators are able to adequately replicate the annual statistics with 67 % and 84% of the historic statistics falling within the inter quartile range for the EMD-VLB and VLB respectively. It is however seen that a small percentage (3%) of historic values are beyond the minimum and the maximum values of the whisker plots for the EMD-VLB model. Table 4.2 compares average mean stochastic values and the historic ones for the 10 stations while Table 4.3 compares the median. Figure 4.24 shows the number of times the historic annual statistics are beyond the interquartile ranges for the EMD-VLB and the VLB generators for 8 statistics analysed.

Table 4.2 Mean values of simulated annual rainfall from the EMD-VLB generator compared with the historic means.

Station	0555567_W	0474255_W	0320348_W	0149082_W	0052590_W	0020866_W	0240891_W	0142805_W	0258894_W	0678776_W
Simulated (mm)	872.41	603.96	352.95	589.75	239.61	590.75	971.09	329.04	399.38	937.22
Historic (mm)	830.11	578.78	324.92	587.62	237.97	604.87	995.25	319.96	394.20	842.65
Difference (mm)	42.30	25.17	28.03	2.13	1.65	-14.12	-24.17	9.08	5.18	94.56

Table 4.3 Median values of simulated annual rainfall from the EMD-VLB generator compared with the historic medians.

Station	0555567_W	0474255_W	0320348_W	0149082_W	0052590_W	0020866_W	0240891_W	0142805_W	0258894_W	0678776_W
Simulated	770.46	590.05	324.49	577.71	231.87	578.69	972.17	316.18	377.29	832.76
Historic (mm)	763.35	576.25	296.05	599.65	232.35	609.85	986.05	302.6	365.35	766.6
Difference (mm)	7.11	13.80	28.44	-21.94	-0.48	-31.16	-13.88	13.58	11.94	66.16

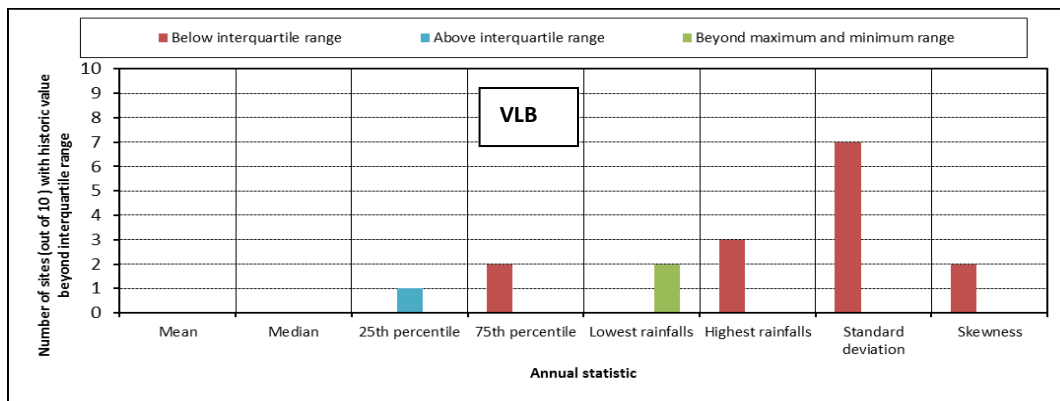
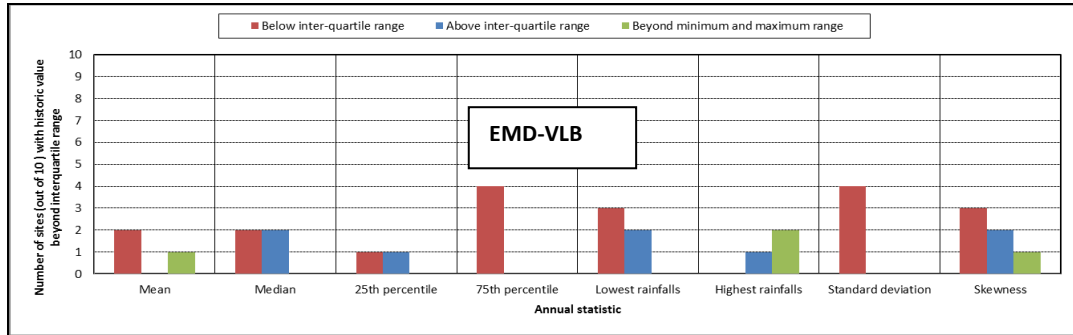


Figure 4.24 The number of times historic statistics fall beyond interquartile ranges of the box plots of stochastic sequences of the EMD-VLB and VLB generator.

Apart from station 0678776, Tables 4.1 and 4.2 reveal that means of the simulated rainfall values are close to the historic statistics. Figure 4.24 shows that for the EMD-VLB the historic values for 5 statistics have fallen above the interquartile range for at least one rainfall station. For the EMD-VLB, historic values for 7 statistics have fallen below the interquartile range for at least one rainfall station. The EMD-VLB has 3 out of 10 statistics falling beyond the maximum and minimum ranges for at least one rainfall station. The VLB has 4 out of 10 of the historic statistics falling above the interquartile range of the stochastic and 1 statistic (lowest rainfalls) falls beyond the whisker plots ranges for at least one rainfall station. The 25th percentile statistic for one station falls above interquartile range of the stochastic for the VLB.

4.3.3 Evaluation and comparison of EMD-VLB and VLB generators using monthly statistics

Figures 4.25 to 4.40 compare the monthly values of the same statistics applied for the yearly analysis in the previous section. The box and whisker plots are presented in two parts of five stations each and on each page for clarity.

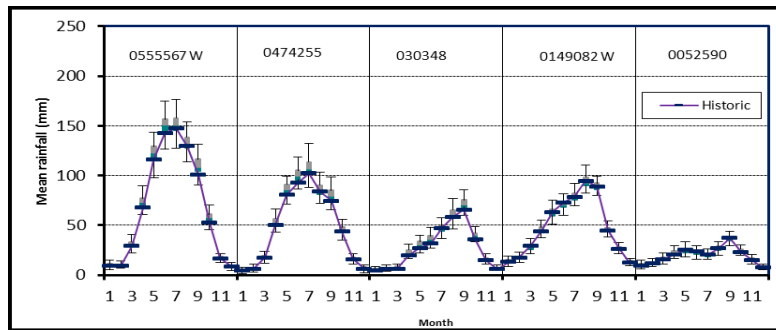


Figure 4.25a Box plots of monthly mean rainfalls for stations 0555567 W, 0474255 W, 0320348 W, 0149082 W and 0052590 W from EMD-VLB generator

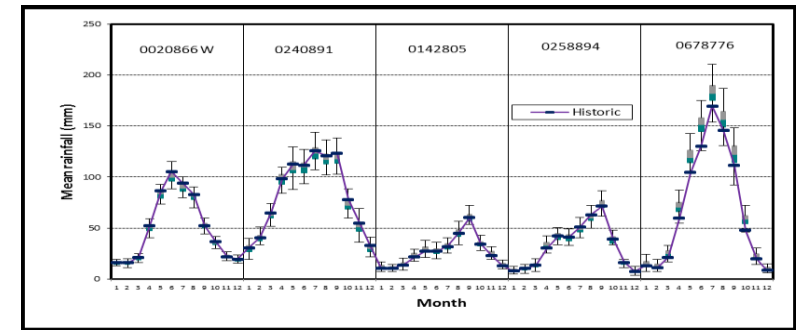


Figure 4.26a Box plots of monthly mean rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

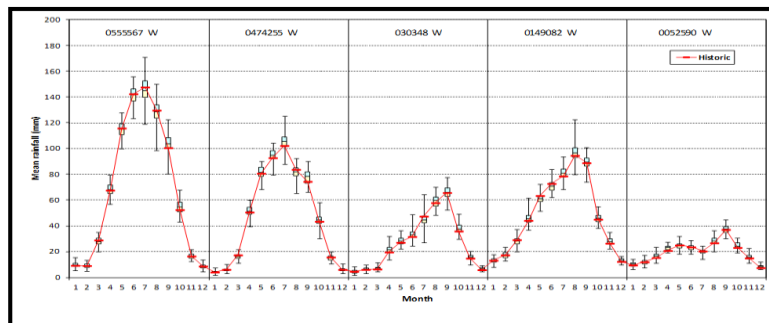


Figure 4.25b Box plots of monthly mean rainfalls for stations 0555567 W, 0474255 W, 0320348 W, 0149082 W and 0052590 W from VLB generator

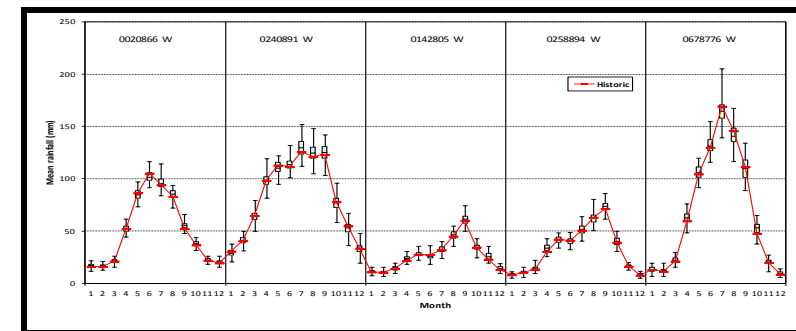


Figure 4.26b Box plots of monthly mean rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

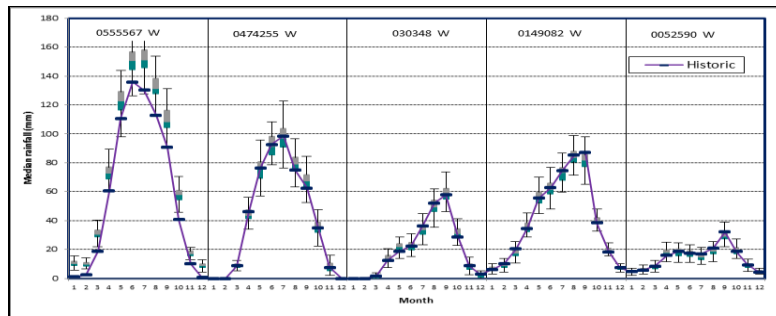


Figure 4.27a Box plots of monthly median rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from EMD-VLB generator

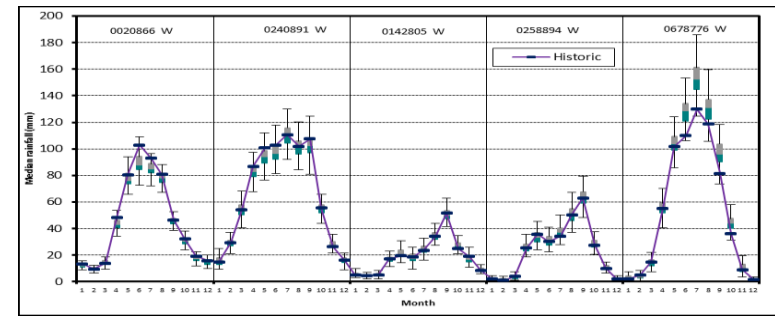


Figure 4.28a Box plots of monthly median rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

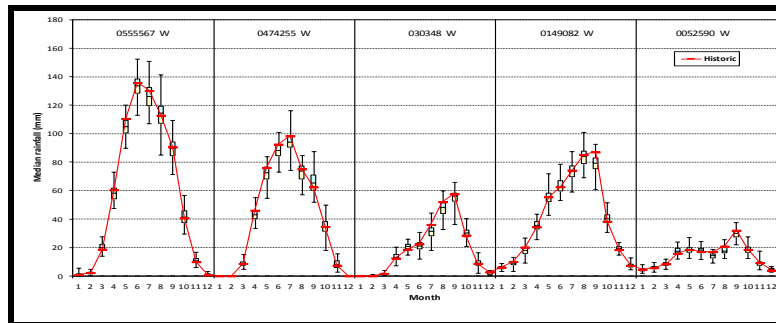


Figure 4.27b Box plots of monthly median rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from VLB generator

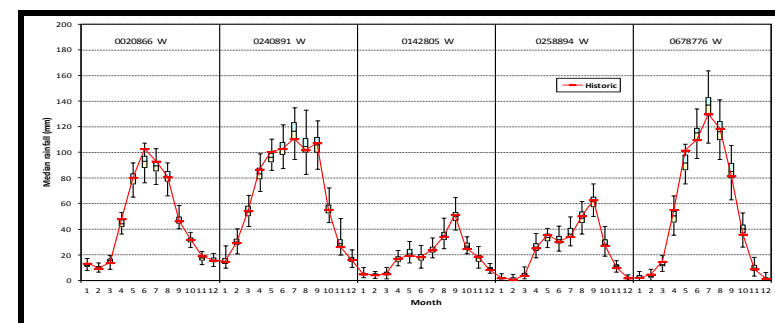


Figure 4.28b Box plots of monthly median rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

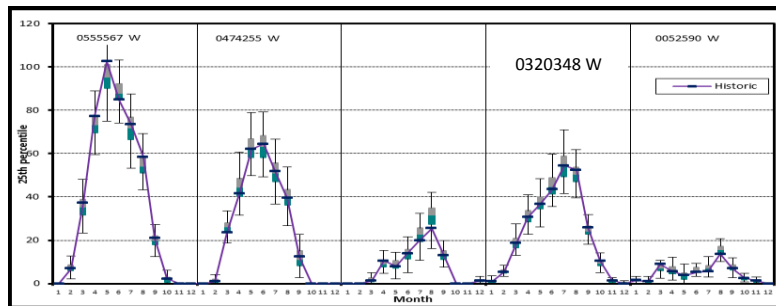


Figure 4.29a Box plots of monthly 25th percentile rainfalls for stations 0555567 W, 0474255 W, 0320348 W, 0149082 W and 0052590 W from EMD-VLB generator

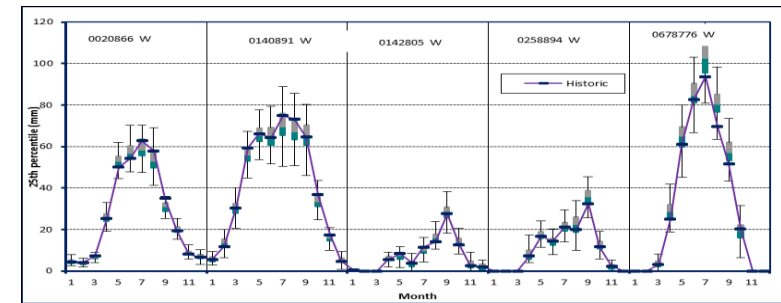


Figure 4.30a Box plots of monthly 25th percentile rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

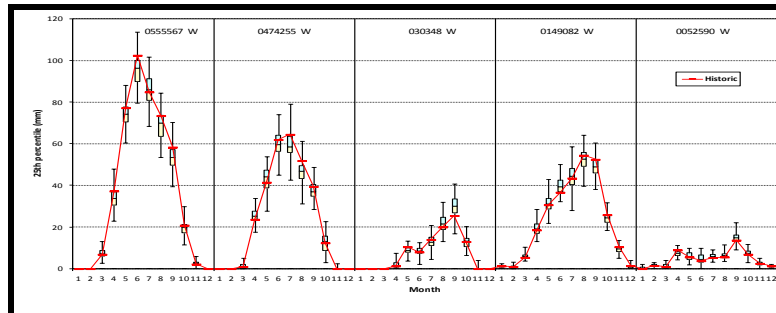


Figure 4.29b Box plots of monthly 25th percentile rainfalls for stations 0555567 W, 0474255 W, 0320348 W, 0149082 W and 0052590 W from VLB generator

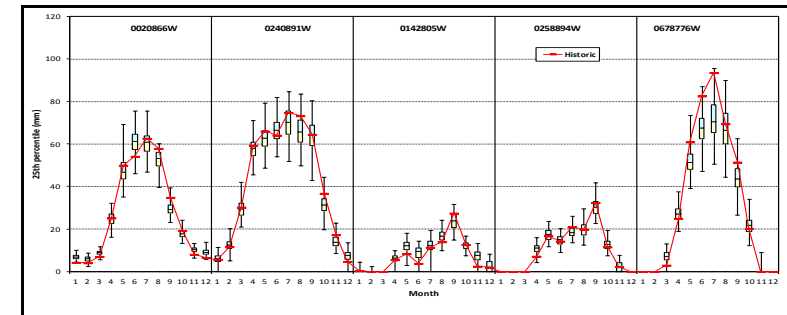


Figure 4.30b Box plots of monthly 25th percentile rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

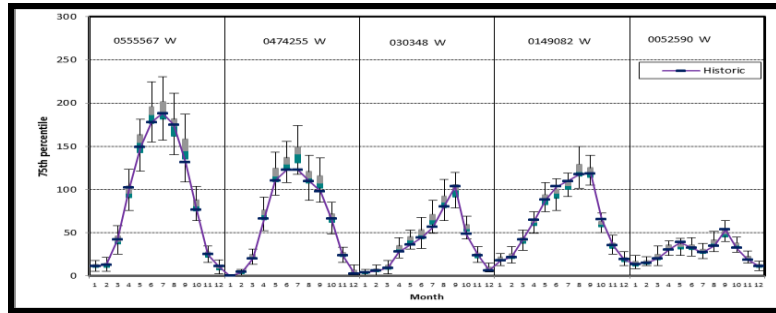


Figure 4.31a Box plots of monthly 75th percentile rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from EMD-VLB generator

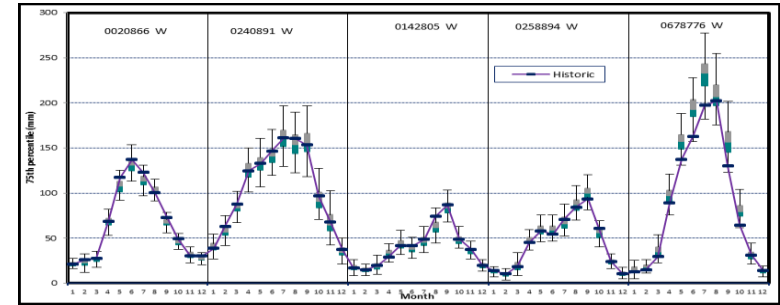


Figure 4.32a Box plots of monthly 75th percentile rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

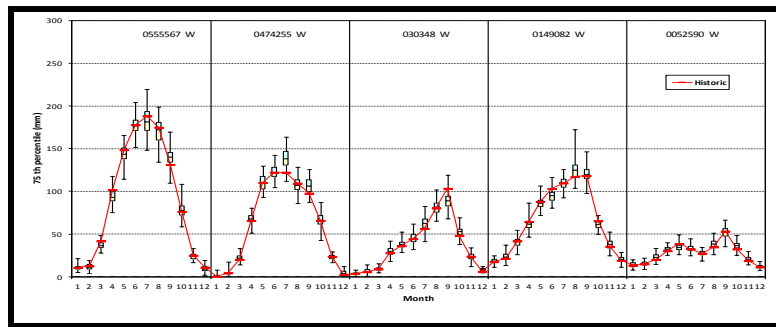


Figure 4.31b Box plots of monthly 75th percentile rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from VLB generator

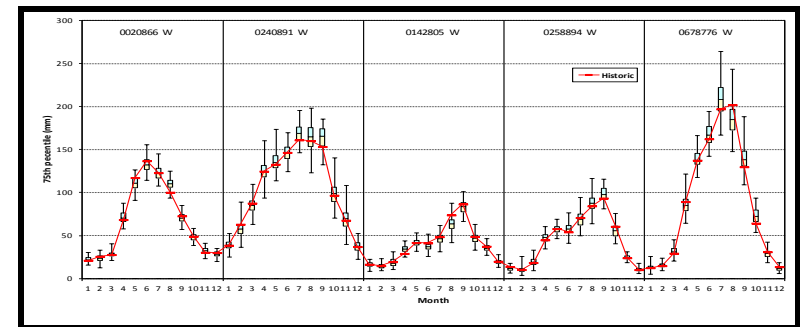


Figure 4.32b Box plots of monthly 75th percentile rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

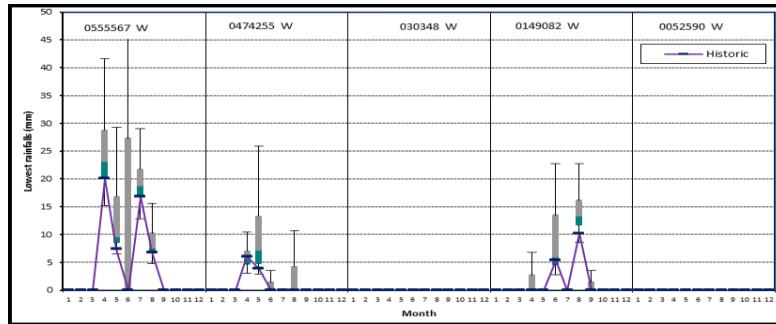


Figure 4.33a Box plots of monthly lowest rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from EMD-VLB generator

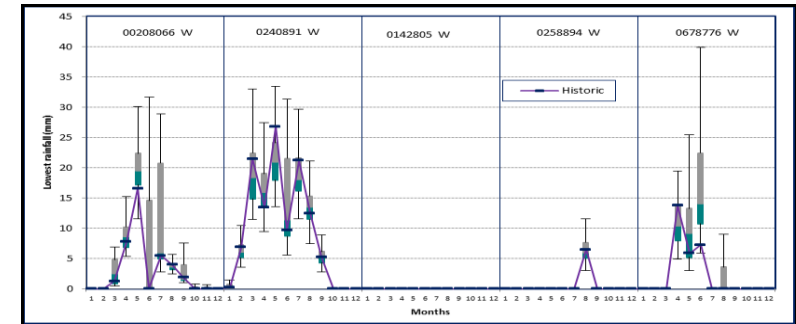


Figure 4.34a Box plots of monthly lowest rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

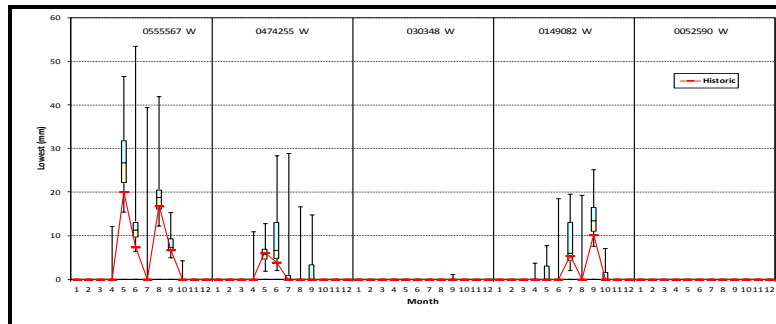


Figure 4.33b Box plots of monthly lowest rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from VLB generator

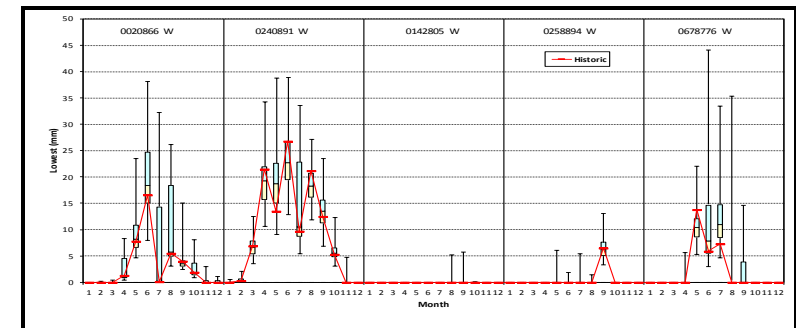


Figure 4.34b Box plots of monthly lowest rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

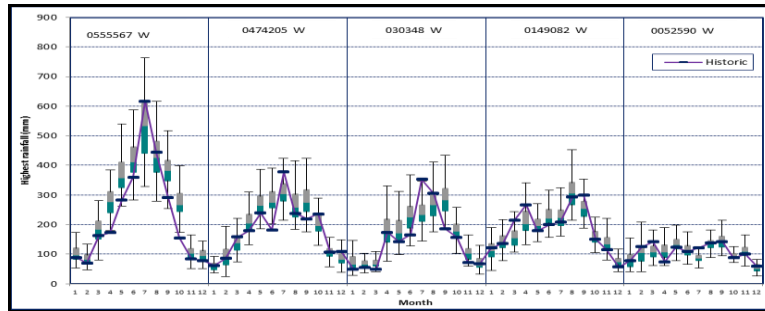


Figure 4.35a Box plots of monthly highest rainfalls for stations 0555567 W, 0474205 W, 030348 W, 0149082 W and 0052590 W from EMD-VLB generator

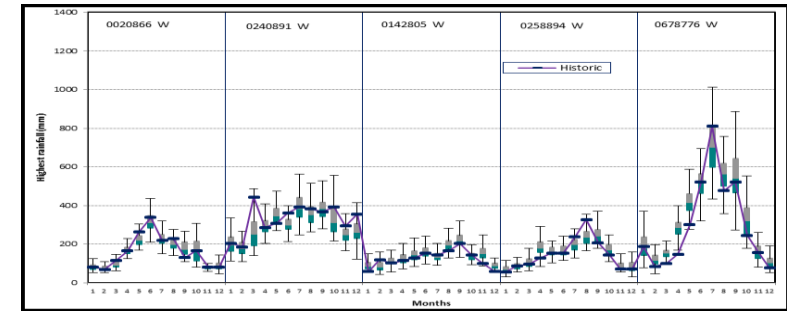


Figure 4.36a Box plots of monthly highest rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

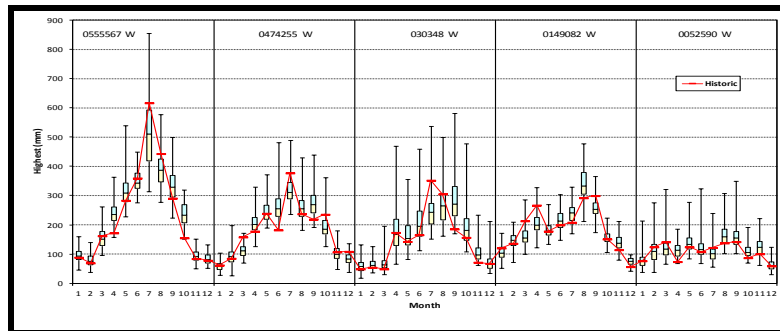


Figure 4.35b Box plots of monthly highest rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from VLB generator

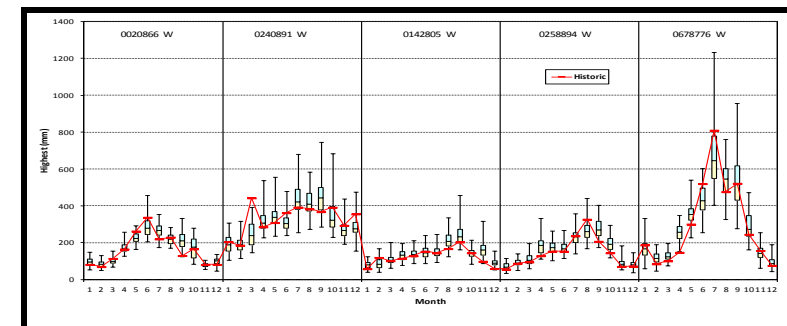


Figure 4.36b Box plots of monthly highest rainfalls for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

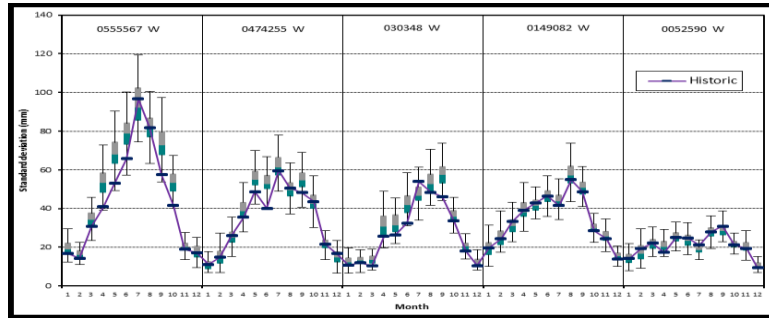


Figure 4.37a Box plots of monthly standard deviations for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from EMD-VLB generator

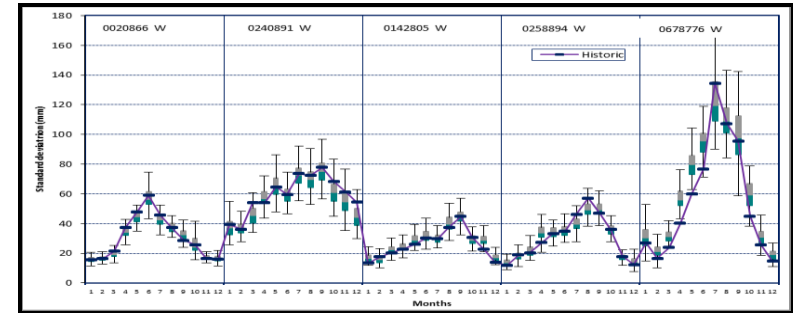


Figure 4.38a Box plots of monthly standard deviations for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

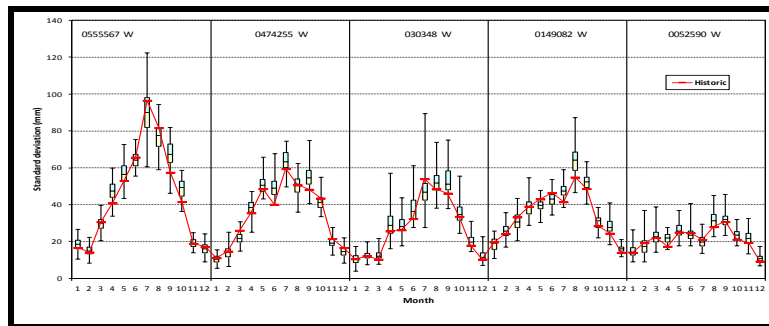


Figure 4.37b Box plots of monthly standard deviations for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from VLB generator

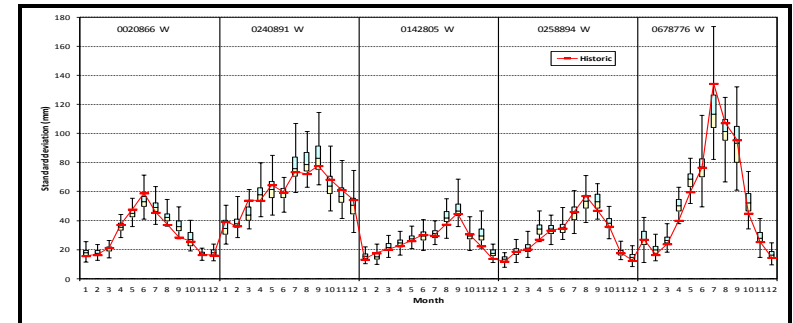


Figure 4.38b Box plots of monthly standard deviations for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

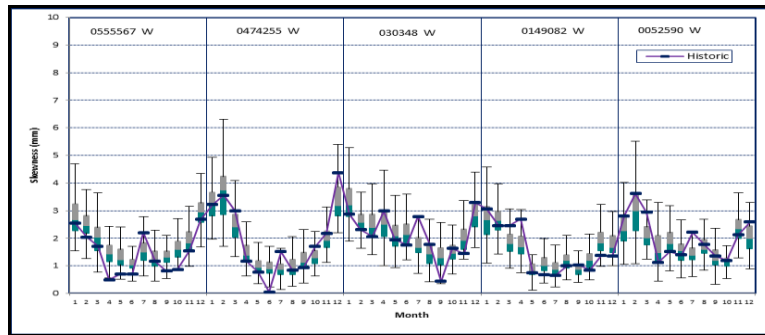


Figure 4.39a Box plots of monthly skewness of rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from EMD-VLB generator

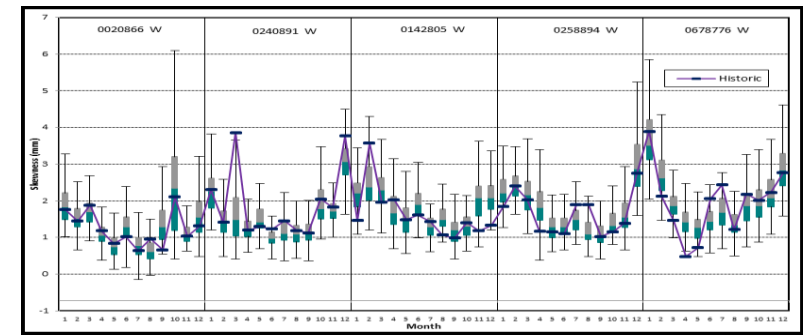


Figure 4.40a Box plots of monthly skewness for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from EMD-VLB generator

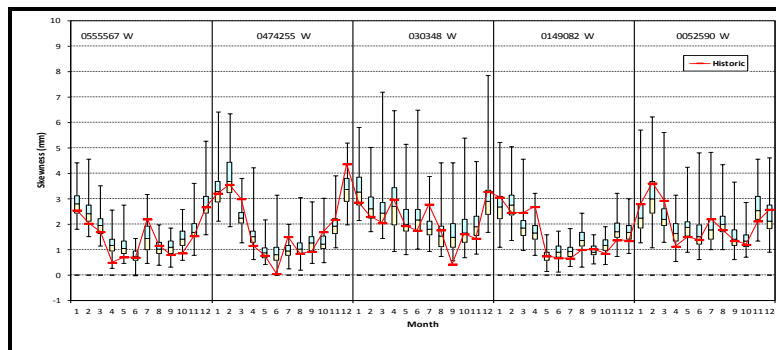


Figure 4.39b Box plots of monthly skewness of rainfalls for stations 0555567 W, 0474255 W, 030348 W, 0149082 W and 0052590 W from VLB generator

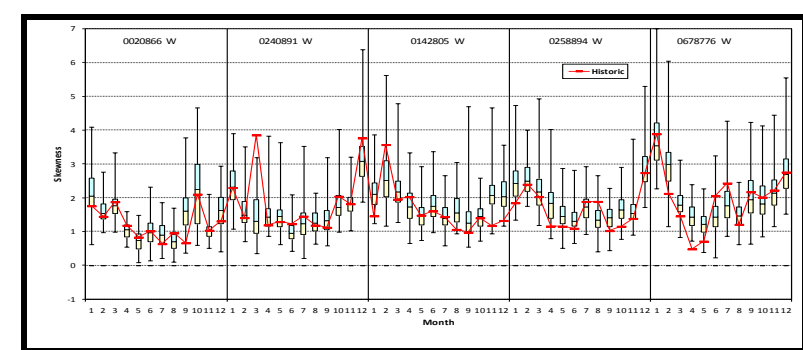


Figure 4.40b Box plots of monthly skewness for stations 0020866 W, 0240891 W, 0142805 W, 0258894 W and 0678776 W from VLB generator

Table 4.4 shows the average mean of the stochastic simulated rainfall compared to the historic while Table 4.5 compares the median values.

Table 4.4 Average stochastic monthly mean rainfall from the EMD-VLB generator compared with historic mean rainfall.

Station	0555567_W	0474255_W	0320348_W	0149082_W	0052590_W	0020866_W	0240891_W	0142805_W	0258894_W	0678776_W
Simulated (mm)	72.40	50.17	29.48	48.98	19.92	49.22	80.83	27.45	33.16	77.93
Historic (mm)	69.18	48.23	27.08	48.97	19.83	50.41	82.94	26.66	32.85	70.22
Difference (mm)	3.23	1.93	2.41	0.01	0.09	-1.18	-2.11	0.79	0.31	7.71

Table 4.5 Average stochastic monthly median rainfall from the EMD-VLB generator compared with the historic median rainfall.

Station	0555567_W	0474255_W	0320348_W	0149082_W	0052590_W	0020866_W	0240891_W	0142805_W	0258894_W	0678776_W
Simulated (mm)	60.05	41.75	20.28	40.95	13.78	43.43	66.09	19.41	24.42	61.01
Historic (mm)	59.53	41.85	20.14	41.69	14.37	46.32	67.29	19.30	23.77	55.49
Difference(mm)	0.52	-0.11	0.14	-0.74	-0.59	-2.89	-1.20	0.12	0.66	5.52

Tables 4.4 and 4.5 highlight the closeness of the mean and median averages of the simulated rainfall to the historic thus demonstrating the effectiveness of the EMD-VLB generator in preserving both the mean and median statistics. Figure 4.41 shows the percentages of the historic statistics that fall within the interquartile ranges of the stochastic ones.

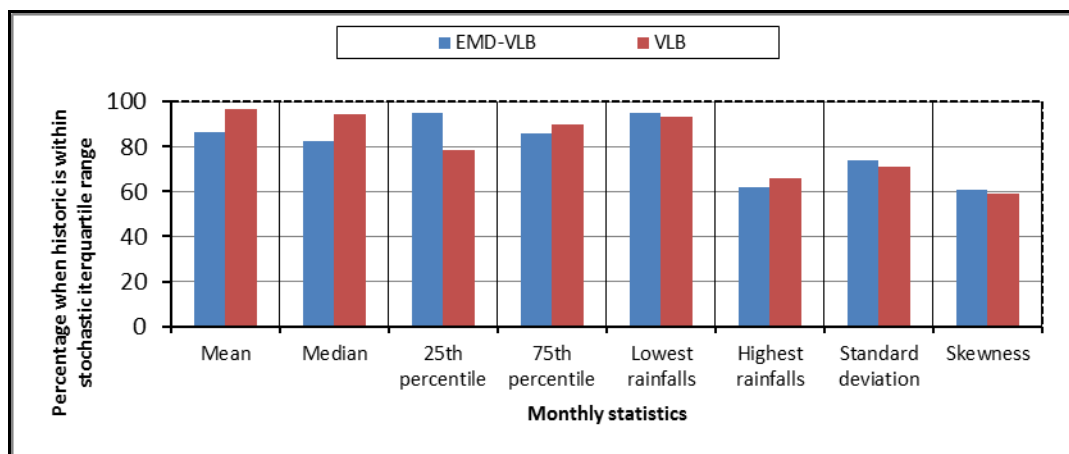


Figure 4.41 Percentages when the historic monthly values are within the interquartile ranges of the stochastic monthly statistics from the EMD-VLB and the VLB generators

The EMD-VLB obtains 80.2% while the EMD obtains 81.0% of the historic statistics falling within the interquartile range of the stochastic statistics. Both generators therefore replicate the monthly statistics adequately. Out of the 8 statistics, the EMD-VLB performs marginally better than the VLB in capturing the 25th percentile, the lowest rainfalls, the standard deviation and the skewness, while the VLB performs marginally better in capturing the mean, the median, the 75th percentile and the highest rainfalls.

Although the generators adequately replicate monthly statistics well, some statistics are either above or below the interquartile ranges or beyond the minimum and maximum ranges as shown in Figure 4.42.

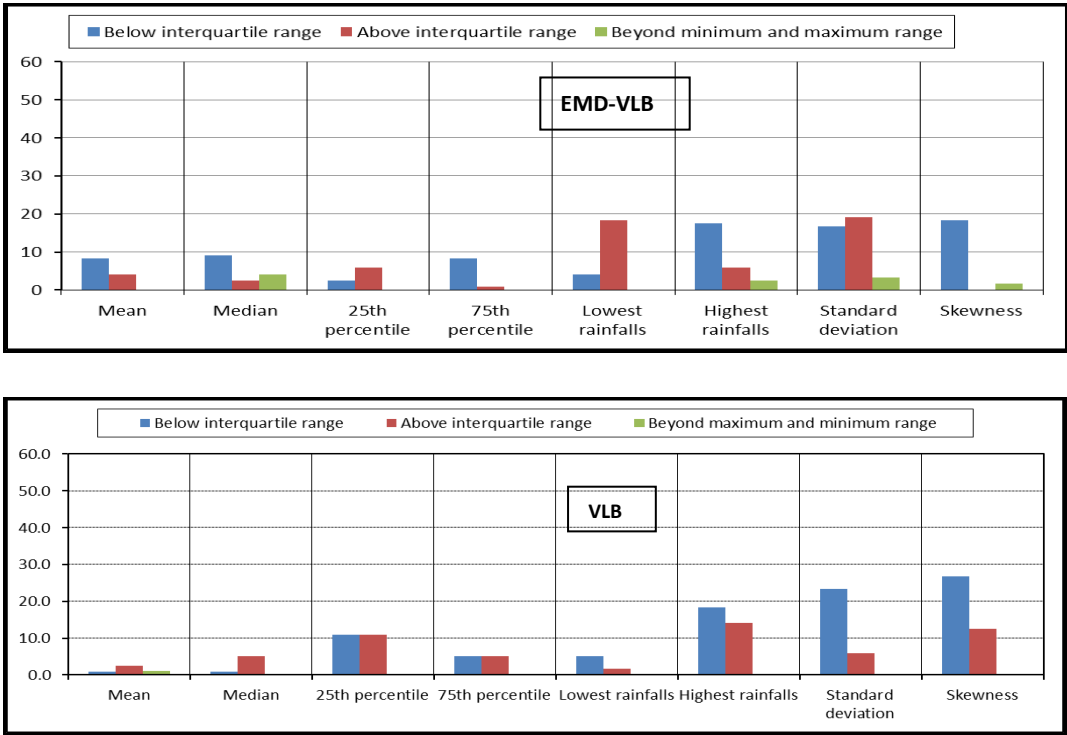


Figure 4.42 The percentages that monthly historic statistics fall above and below the interquartile range and beyond the range of generated statistics for the EMD-VLB and the VLB generator.

4.4 Discussion on evaluation of performance and comparison of generators

Section 4.3 in this chapter illustrates that both EMD-VLB and the VLB are potentially able to replicate historic statistics reasonably well with the VLB performing better in replication of the annual statistics. It is however noted that the poorer performance of the EMD-VLB mainly arises from the poor replication of statistics for one (0678776 W) out of the 10 stations. The same station was found to produce undesirable oscillations in the generation of IMFs that the other 9 stations were free from. It is therefore possible that the block lengths used to develop synthetic rainfall for this station lead to some bias as a result of its unique behaviour. A simple comparison of the historic statistics and correlation with the other stations however did not reveal any exceptional behaviour. The poor replication of the statistics for this station prompts the need for a more in-depth analysis.

The difference between the EMD-VLB and the VLB that is considered most significant is the higher variability of the box plots of the minimum runs sums by the EMD-VLB generator (Section 4.3.1). This difference in variability means that the EMD-VLB could predict drier conditions than the VLB generator for low probabilities of occurrence. Since the extreme exceedance statistics from stochastics are the ones often used to make practical decisions, this difference in variability of the run sums calls for additional analysis and comparison of the methods. This could lead to a refinement of the EMD-VLB generator.

5. CONCLUSIONS AND RECOMMENDATIONS

This research report aimed to find out if the performance of the Variable Length Block (VLB) monthly stochastic rainfall generator can be improved by using empirical mode decomposition (EMD). From a review of the literature, it was perceived that EMD has the ability to identify and quantify the effects of both short-term (intra-decadal) and long-term (inter-decadal) processes that affect rainfall. Block identification by EMD could therefore be more rational than the subjective approach used in the standard VLB generator. The hybrid EMD-VLB generator was developed and assessed using a multi-site stochastic rainfall generation problem that used 93 year-long rainfall data from 10 stations distributed across South Africa. To evaluate the performance of the generator, 8 statistical measures namely; the mean, the median, the 25th percentile, the 75th percentile, the lowest rainfalls, the highest rainfalls, the standard deviation and the skewness were used at the annual and the monthly time scale. In addition, the minimum run sums test ranging from 1-24 years was applied to test the preservation of the long-term dependence characteristics. The hybrid model was used to generate 101 stochastic sequences of the same length as the historic data (93 years) and box plots of the statistics were plotted. The historic values were plotted on the boxplots in order to assess how well the historic statistics were replicated. In addition the hybrid EMD-VLB model was compared with the standard VLB generator using similar box plots.

The EMD-VLB generator was found to replicate the historic statistics and the minimum run sums reasonably well. When compared with the standard VLB, it was found out that the VLB performed better than the hybrid EMD-VLB model in replicating yearly statistics. The two models were found to be almost at par in replicating monthly statistics with the VLB being marginally better. In the overall, both the EMD and the hybrid EMD-VLB are considered to perform adequately well for practical applications. The minimum run sums test

however revealed that the EMD-VLB consistently obtained more variable statistics than the VLB generator. This implies that practical decisions based on extreme statistics (low probabilities of exceedance) could at times be significantly different depending on the stochastic generator used.

Following are the recommendations from this study;

1. More tests and some aspects of the EMD need to be carried out for further improvement of the EMD-VLB model in areas of; use and comparison of alternative end condition rules for extrapolations and use of other types of splines for interpolations. This is because according to Pegram et al (2008) and Wu and Qu (2008), different end conditions will produce different number of IMFs in a station thus the block lengths produced as a result will vary differently.
2. A test should be done to establish the effect of the results if Ensemble Empirical Mode Decomposition (EEMD) is applied to the original historic time series. This entails addition of white noise before decomposition that assists in removal of different fluctuations occurring within the same IMF as was observed in the presence of very short fluctuations within long ones.
3. Consider other methods of resampling the blocks obtained from the different IMFs as the resampling adopted here implicitly assumes that the processes that affect rainfall are mutually exclusive while EMD reveals a lot of concurrence of these processes.
4. Test out the hybrid generator on more typical basin or catchment scale problem instead of the sub-continental (South Africa wide) problem that the model was tested on.
5. Carry out additional comparison on the EMD-VLB and the VLB generators to find out why the EMD-VLB obtains a higher variability of the minimum run sums than the VLB and use these findings to further improve the rainfall generation.