

https://dspace.org/

The use of empirical mode decomposition (EMD) and variable length boostrap (VLB) for stochastic rainfall generation

Nyaga, Muriithi Job

2015-05-07
http://hdl.handle.net/10539/17668

Downloaded from a DSpace repository. https://dspace.org/

APPENDIX A

Use of cubic splines in EMD data generation

Discrete hydrological events plotted as data points on continuous time series graphs are interconnected by straight lines or curved splines to interpolate unknown data values in between. The possibility of equal rainfalls in a continuous time period is highly unlikely and therefore the missing data points in between do not often assume the same values as the adjacent ones in a rainfall time series. In this regard, smooth curved splines (as opposed to linear inter(extra)polations) are widely used for piecewise interpolation/extrapolations between (or extensions from) neighbouring data points. The use of cubic splines as interpolation functions is due to their stable and smooth characteristics(McKinley and Levine, 1998; Huang et al. ,1998; Kruger, 2002; Klasson, 2008), making them one of the most common piece-wise polynomial approximations that are fitted between each successive pairs of data points. Various modifications of the cubic splines include, rational splines (Pegram et al., 2008; Peel et al., 2009), constrained cubic splines (Greef, 2003) and envelope correction methods (Tianlu and Zengli, 2013), and the use other curved splines exist in literature(including Hermite splines, Beizer splines, Catmul-Rom Splines, B-splines and Nurbs).The simplicity and robustness of cubic splines in extrapolations beyond the time series and as interpolation functions between data points enhances their choice against various other mentioned splines whose applicability is limited by their computational complexity.

Cubic spline interpolation

Cubic spline construction

A piece-wise function, $f(x)$ that satisfies both smooth and continuous characteristics of a cubic spline's interpolation (and extrapolation) is given by;

$$f(x) = \begin{cases} f_1(x) & if\ x_1 \leq x \leq x_2 \\ f_1(x) & if\ x_2 \leq x \leq x_3 \\ & . \\ & . \\ f_{n-1}(x) & if\ x_{n-1} \leq x \leq x_n \end{cases} \qquad (A.1)$$

Where $f_i(x)$ is a third order polynomial given by;

$$f_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \qquad (A.2)$$

$for\ all\ i = 1,2, \dots, n-1,$

where, a, b and c are polynomial constants.

The piecewise function, $f(x)$,interpolates all the data points as well as the complete curve on the interval $[x_1, x_n]\ for\ all\ i = 2, \dots, n-1.$ The function is thereby joined by a pair of piecewise polynomials at all the shared data points. Apart from the extrema nodes at the ends of the time series in consideration, all the other nodes are defined by two polynomials on either side of each node; to achieve continued smoothness at a node, it is therefore notable that the joining polynomials at a particular node must be of the same value. This can be expressed as;

$$f_i(x_i) = f_i(x_{i-1}) \qquad (A.3)$$

Similarly, the first and second order derivatives of the splines polynomials need have the same values on each and every shared node for smooth continuity between the nodes$(x_i, x_{i+1}), (x_{i+1}, x_{i+2}), \dots \dots . (x_{n-1,} x_n)$.

The first and second order derivatives of the connecting splines are given by;

$$f'_i(x) = 3a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i \qquad (A.4)$$

$$f''_i(x) = 6a_i(x - x_i) + 2b_i \qquad (A.5)$$

$$for\ all\ i = 1,2, \dots, n-1,$$

and a, b, and c are polynomial constants.

As indicated, the functions $f(x), f'(x)\ and\ f''(x)$ need to be continuous over the interval $[x_1, x_n]$ of the time series both on each node as well as on consecutive intervals. The above simple theory explains the smoothness attributes of a cubic spline thus its adoption in this study.

A complete analysis of cubic splines can be obtained from various texts including Burden and Faires (1993), McKinley and Levine (1998), Kruger (2002), Greef (2003) and Klasson (2008).

Cubic spline extrapolation

Extrema end points are used to compute adjacent end and inner points in EMD and it has been found that their ineffective determination leads to wildly oscillations in subsequent siftings that could result to unrepresentative IMFs. Figure A1 demonstrates the relative effects two different extrapolation methods that eventually results to different oscillations from an original similar time series.

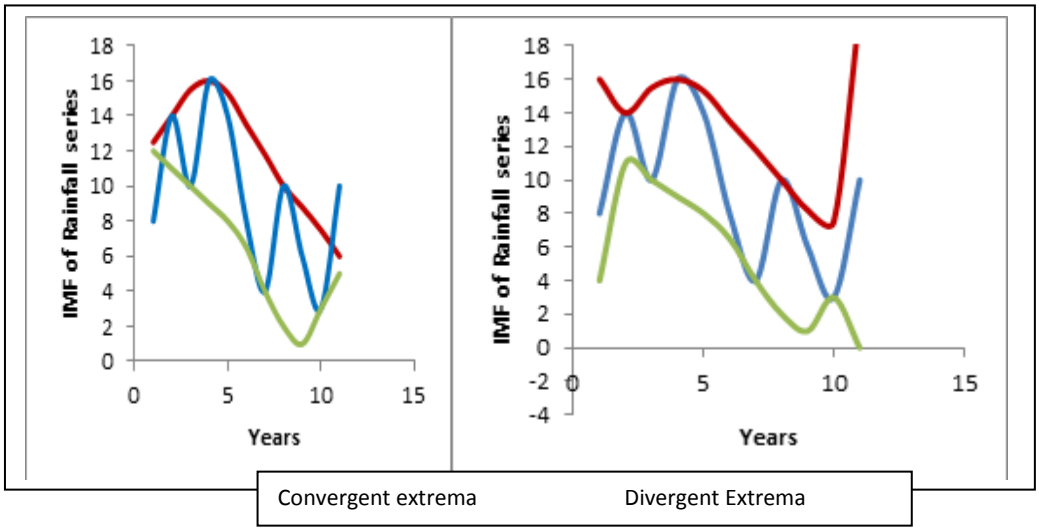


Figure A.1 Effects of different end conditions on EMD decompositions

From Figure A1, it can be seen that the choice of the extrapolation method greatly affects the shape of cubic splines from which inner data points are derived.

Other effects of poor selection of end points are overshoots and undershoots (less or higher rainfall amounts than the series) at intermediate points in successive approximations(Pegram et al, et al, 2008, Kruger, 2002; Huang et al 1998 and Peel et al, 2005) as illustrated in Figure A2 below.

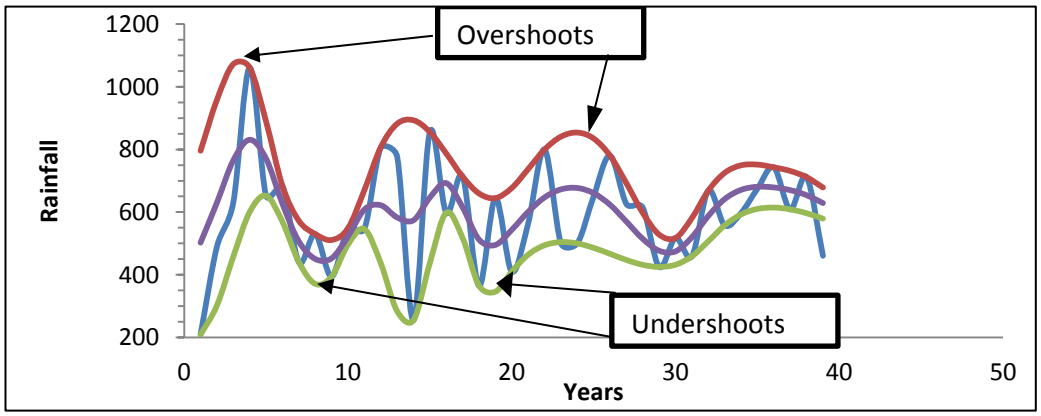


Figure A.2- The effects of envelope end effects corrections on oscillations in station0240891W

Since it takes several siftings (and hence computations) to realize an IMF, as end extrema progressively generate data in each subsequent sifting, the oscillations of the IMFs become inwardly infectious within each successive sifting. Because end points need to be correctly determined to avoid undesirable and unrepresentative results, it is thus conclusive to state that the effectiveness of EMD as a data processing tool is highly dependent on end extrema computations (Rato et al., 2008).

As illustrated in the interpolation, data extension (extrapolation) should be carried out similarly as in the interpolation. To limit unnecessary impact into the interiors of the IMFs during sifting processes, according to Zhao and Huang (2001), time series ends extrema are added such that the resultant cubic spline's derivatives (first and second order)as well as its function values result similarly to the interpolation process explained above so as to enhance smoothness.

The choice of extrapolation methods in Empirical Mode Decomposition

Various approaches that exist to correct end points effect include; Mirror Method (Zhao and Huang, 2001), Envelope Correction, (Tianlu and Zengli, 2013), Average Method (Dätig and Schlurmann (2004)), Slope Zero method (Rilling et al., 2003; Chiew et al., 2005), The Improved Slope Based Method (Wu and Qu, 2008) and the Characteristic Waves approach by Huang et al. (1998). The presence of these several methods highlights the absence of concrete theoretical principles that can be used for data extensions. Some approaches could lead to indefiniteness because of the different characteristic waves produced from different data (Zhao and Huang, 2001). This leaves researchers with options of choosing end extension methods as long as the selected methods produce plausible results.

The Mirror Method in Empirical Mode Decomposition

The mirror approach reproduces extension end data values from the original data to construct end points as opposed to simulating new values and since this demonstrates a true reflection of available data; that simplicity warrants its adoption for this study. The spline interpolation and extrapolation coefficients of the Mirror Method at the ends are entirely produced by the available data and not by the extended data set, which enhances the reliability of the method. The mirror method is demonstrated in a simple analogy in Figure A3. It consists of data placed on top of a mirror being reflected below and a cyclic pattern is eventually formed in a continuous circle. In the method, end points are constructed by reflecting the ends of the original time series by the same data amounts and time differences as the extrema closest to the edges (Wu and Qu, 2008).

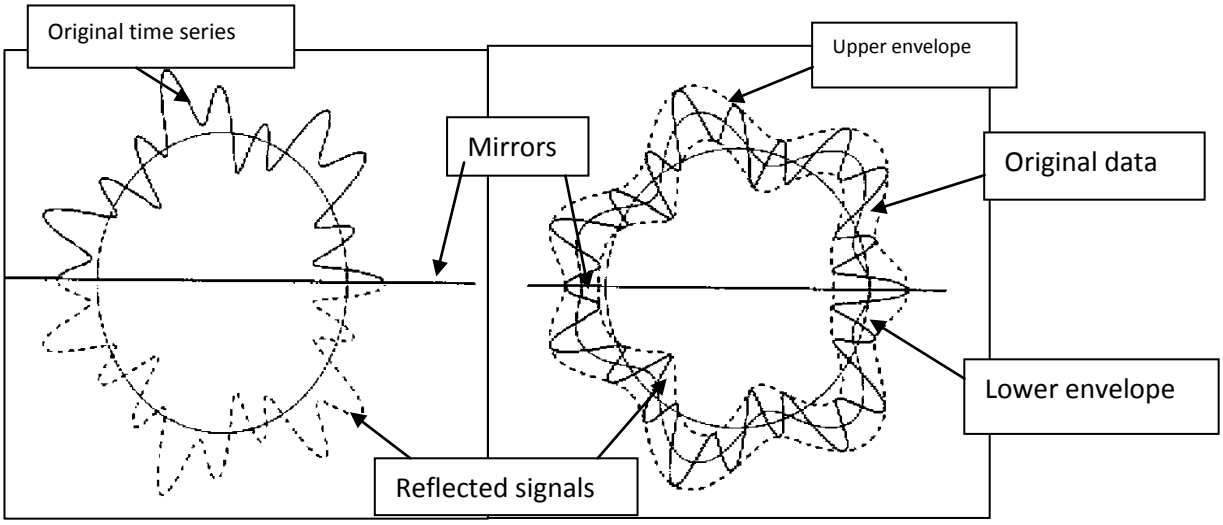


Figure A.3 Illustration of the mirror method construction (Zhao and Huang, 2001).

It is however noted that the upper mirror's output and not the lower, is the one used eventually to compute the IMFs. Zhao and Huang (2001) introduce and demonstrate the theoretical criterion of the method's reliability as well as its strengths in modelling variability in any kind of data. A complete graphical illustration of the mirror method for station 0240891W is shown in Figure A4.

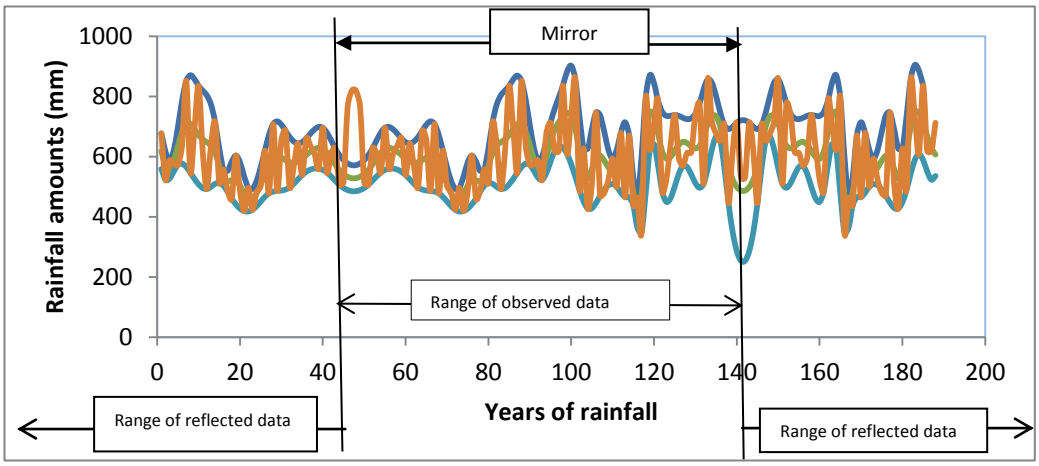


Figure A 4 Graphical illustration of the mirror method from station 0240891W.

Appendix B

Empirical Mode Decomposition of individual stations

Using the methodology described in chapters 2.2 and 3.3 for the EMD, each station is decomposed separately to realize its time series constituent IMFs and a residual. The 10 stations described in chapter 3.2 are used in order to achieve the objective of generating multi-site rainfall. In order to realize the variabilities exhibited by decompositions from all the stations, the Figures B1 to B72 below appear in the following order, the historic time series, the first IMFs to the last IMFs (residual trends) as a result of subsequent siftings. This is vital in observing the time-frequency differences in subsequent IMFs from the first to the last stages of decomposition. For easier and clearer visibility, each of the stations' decomposition appears on each separate page.

Analysis of the decomposition derived from station 0020866 W

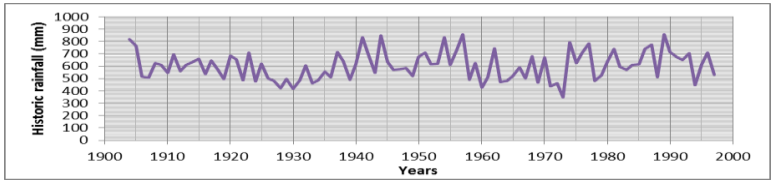


Figure B1- The original time series of station 0020866 W

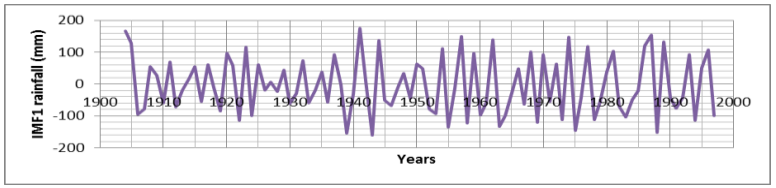


Figure B2- IMF1 of the decomposed time series for station 0020866 W

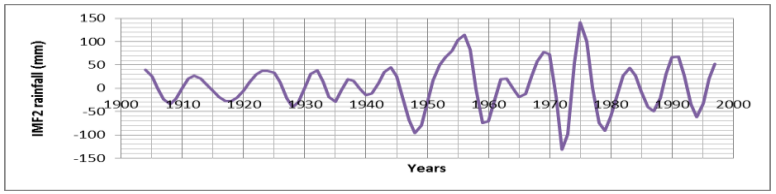


Figure B3- IMF2 of the decomposed times series for station 0020866 W

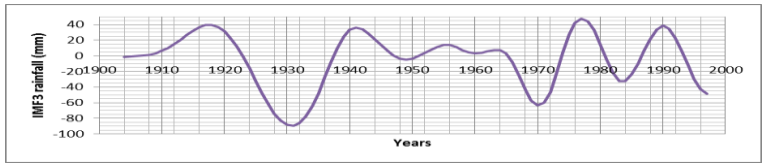


Figure B4.-IMF3 of the decomposed time series for station 0020866 W

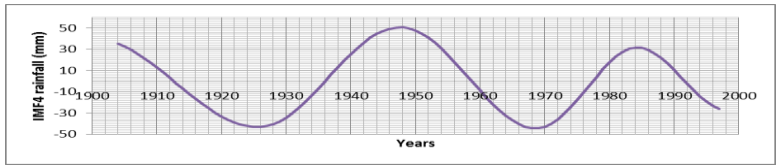


Figure B5- IMF4 of the decomposed time series for station 0020866 W

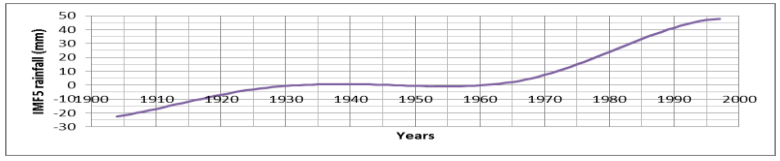


Figure B6- IMF5of the decomposed time series of station 0020866 W

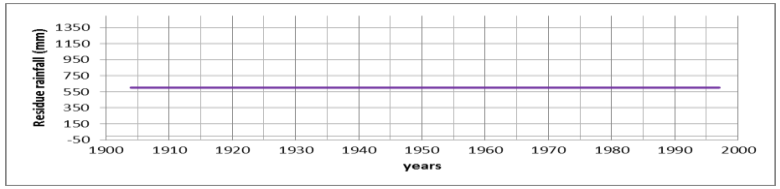


Figure B7-The residual trend of the times series from station 0020866 W

Analysis of decomposition derived from station 0142805 W

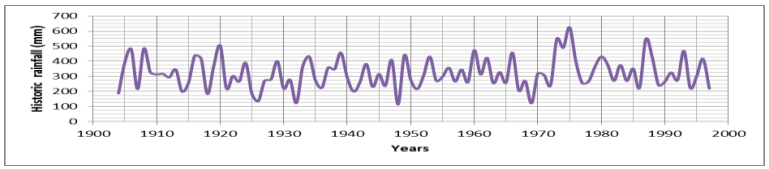


Figure B8-The original time series of station 0142805 W

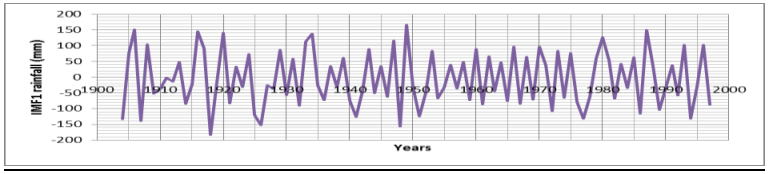


Figure B9-IMF1 of the decomposed time series of station 0142805 W

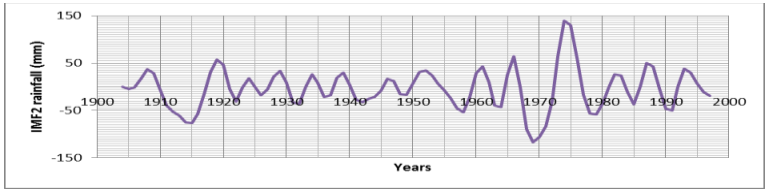


Figure B10-IMF2 of the decomposed time series of station 0142805 W

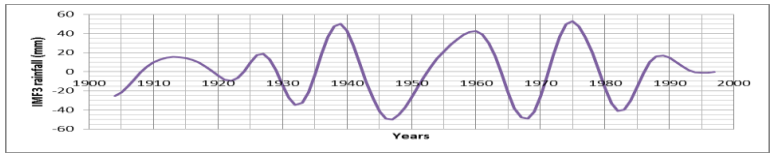


Figure B11-IMF3 of the decomposed time series of station 0142805 W

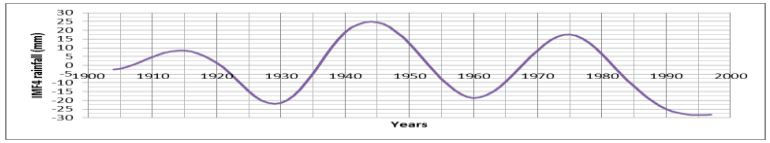


Figure B12-IMF4 of the decomposed time series of station 0142805 W

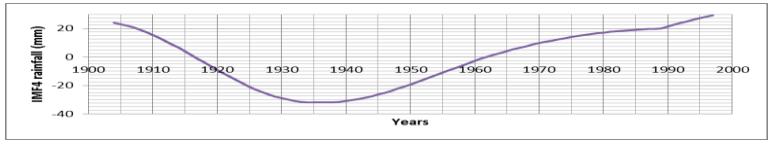


Figure B13-IMF4 of the decomposed time series of station 0142805 W

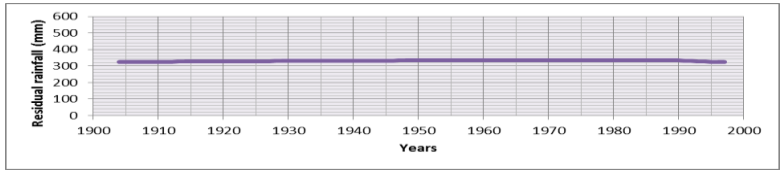


Figure B14-The trend of the rainfall time series for station 0142805 W

Analysis of decomposition derived from station 0149082W

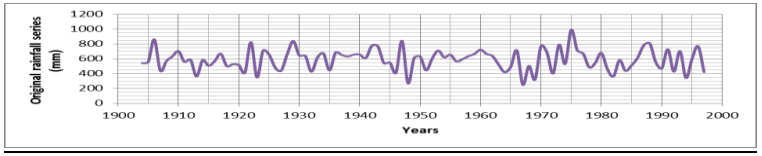


Figure B15- Original time series for station 0149082W

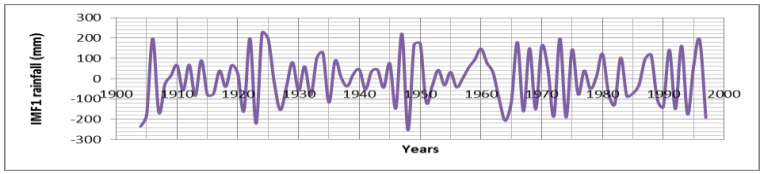


Figure B16- IMF1 of the decomposed time series for station 0149082 W

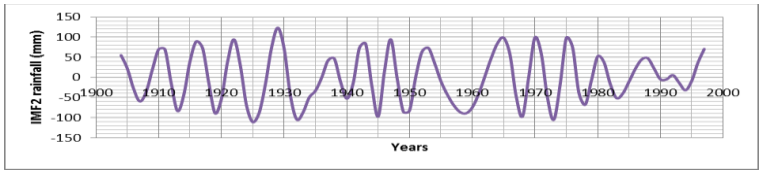


Figure B17- IMF2 of the decomposed time series for station 0149082 W

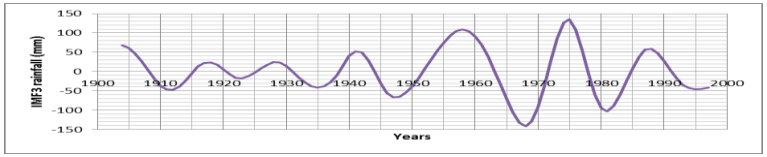


Figure B18- IMF3 of the decomposed time series for station 0149082 W

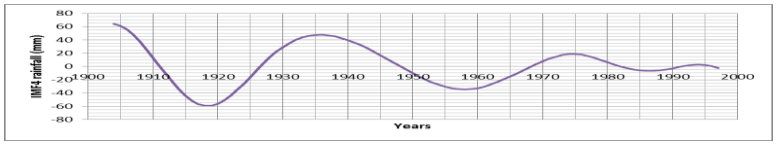


Figure B19- IMF4 of the decomposed time series of station 0149082 W

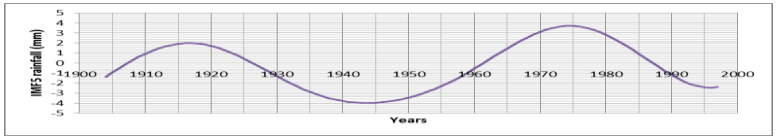


Figure B20- IMF5 of the decomposed time series for station 0149082 W

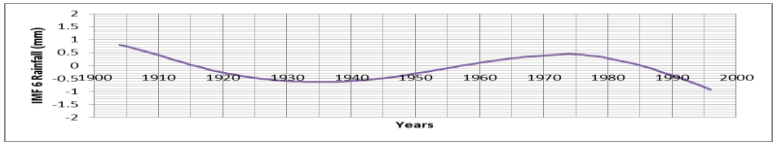


Figure B21- IMF6 of the decomposed time series for station 0149082 W

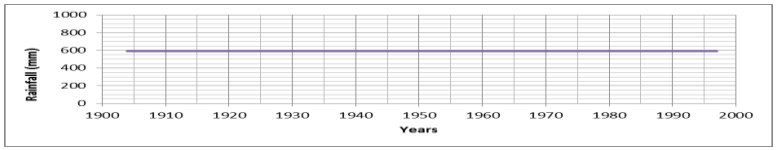


Figure B22- Residue of the decomposed time series for station 0149082 W

Analysis of decomposition derived from station 0320348 W

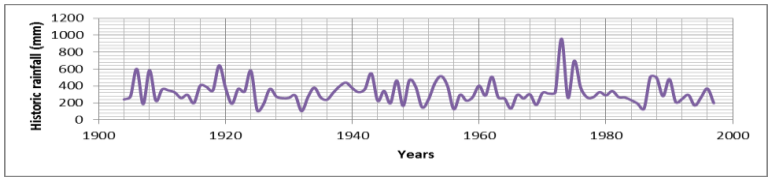


Figure B23- Original time series of station 0320348 W

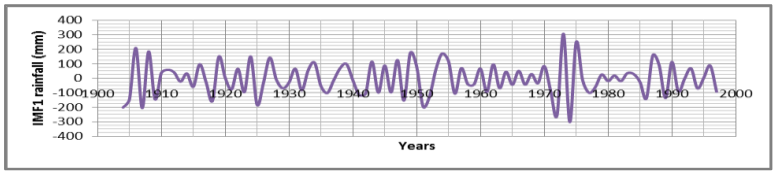


Figure B24-IMF1 of the decomposed time series for station 0320348 W

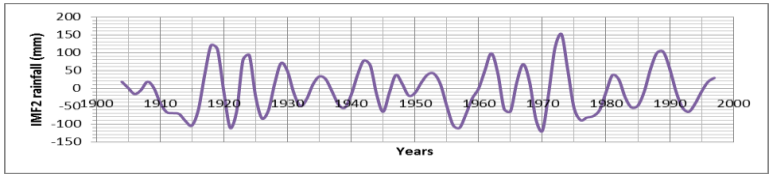


Figure B25- IMF2 of the decomposed time series for station 0320348 W

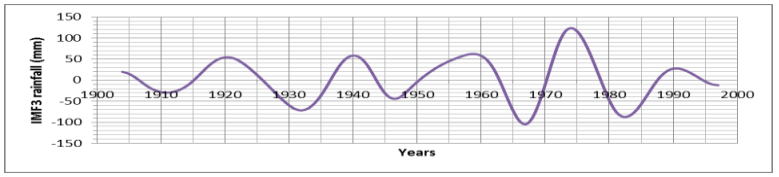


Figure B26- IMF3 of the decomposed time series for station 0320348 W

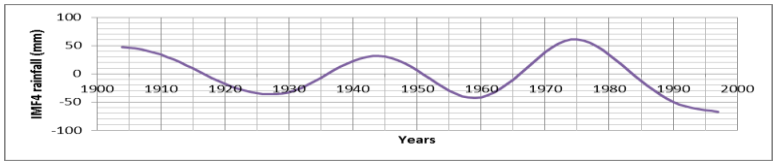


Figure B27- IMF4 of the decomposed time series for station 0320348 W

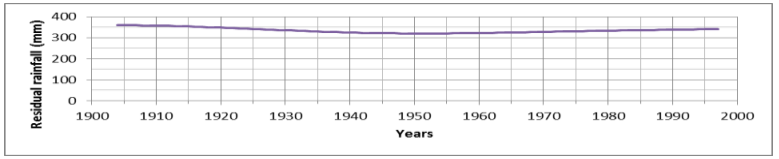


Figure B28-Residue of the decomposed time series of station 0320348 W

Analysis of decomposition derived from station 0258894 W

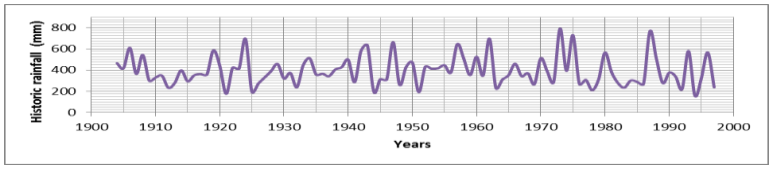


Figure B29-Historic time series for station 0258894 W

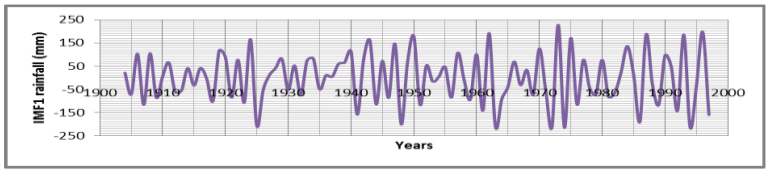


Figure B30-IMF1 of the decomposed time series for station 0258894 W

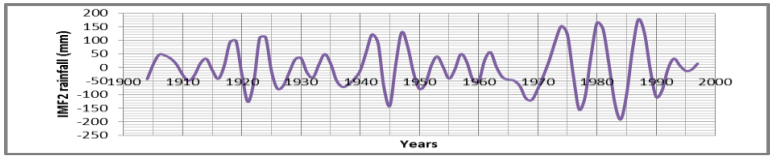


Figure B31-IMF2 of the decomposed time series for station 0258894 W

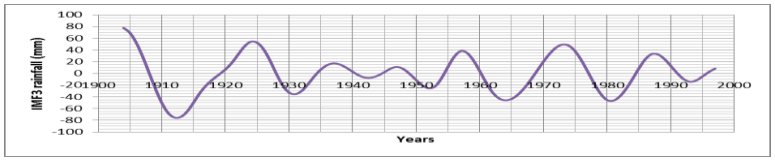


Figure B32-IMF3 of the decomposed time series for station 0258894 W

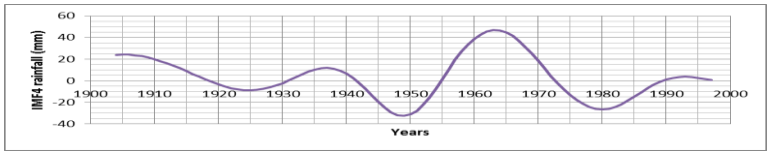


Figure B33-IMF4 of the decomposed time series for station 0258894 W

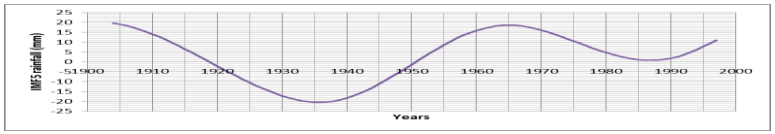


Figure B34-IMF5 of the decomposed time series for station 0258894 W

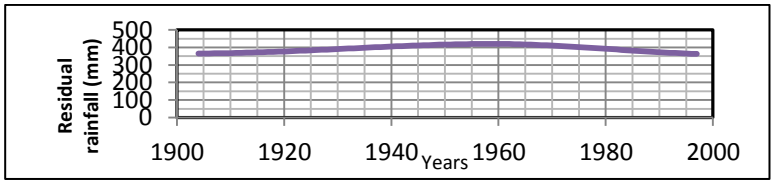


Figure B35-Residue of the decomposed time series for station 0258894 W

Analysis of decomposition derived from station 0474255 W

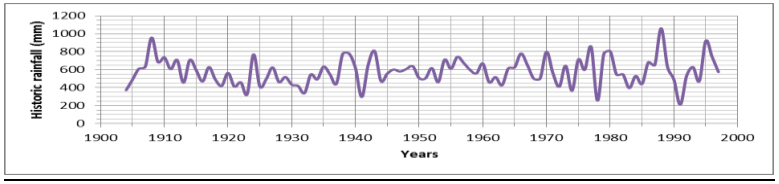


Figure B36-Historic times series for station 0474255 W

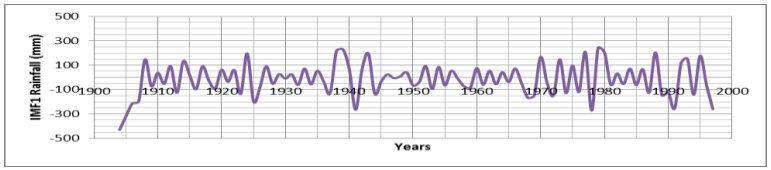


Figure B37-IMF1 of the decomposed time series for station 0474255W

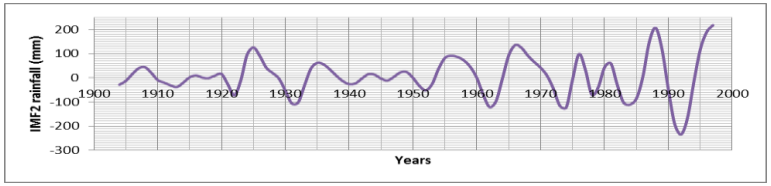


Figure B38-IMF2 of the decomposed time series for station 0474255 W

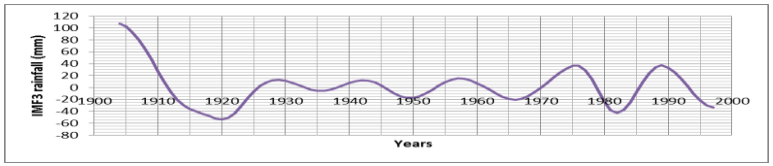


Figure B39-IMF3 of the decomposed time series for station 0474255 W

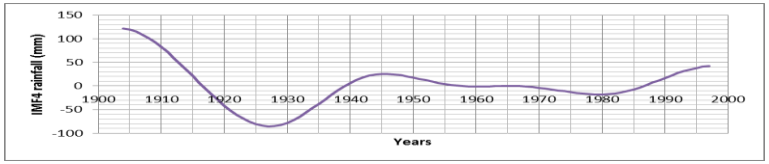


Figure B40-IMF4 of the decomposed time series for station 0474255 W



Figure B41- IMF5 of the decomposed time series for station 0474255 W

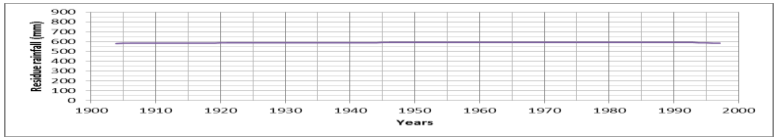


Figure B42-Residue of the decomposed time series for station 0474255 W

Analysis of decomposition derived from station 0240891W

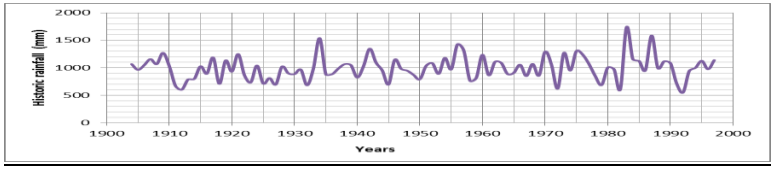


Figure B43-Historic time series for station 0240891 W

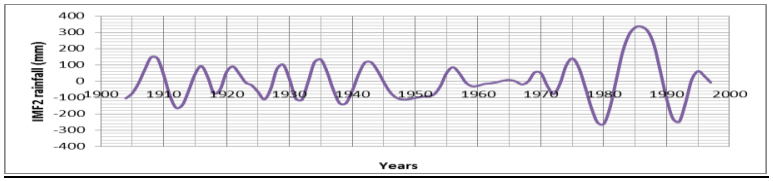


Figure B44- IMF2 of the decomposed time series for station 0240891 W

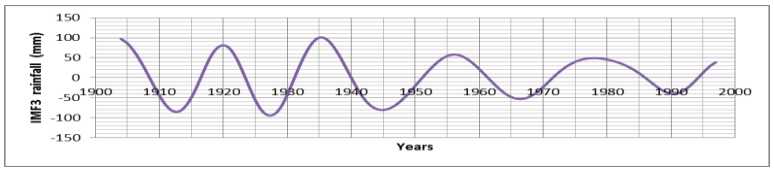


Figure B45- IMF3 of the decomposed time series for station 0240891 W

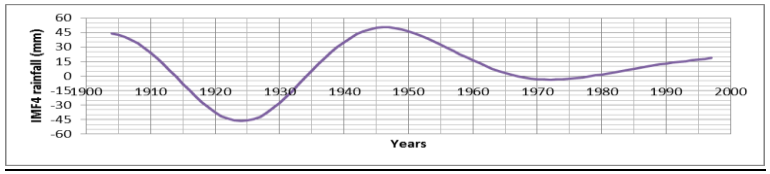


Figure B46- IMF4 of the decomposed time series for station 0240891 W

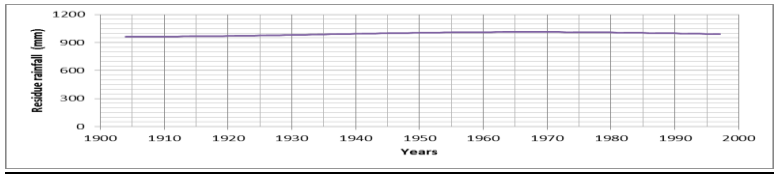


Figure B47- Residue of the decomposed time series for station 0240891 W

Analysis of decomposition derived from station 0555567 W

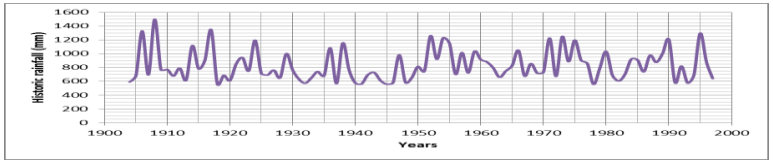


Figure B48- Historic rainfall series for station 0555567 W

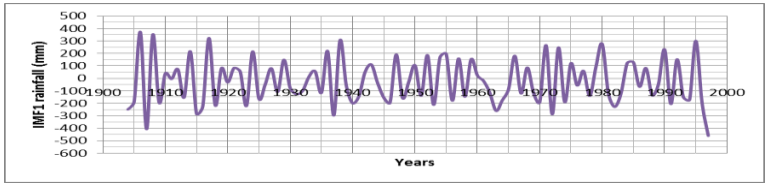


Figure B49-IMF1 of the decomposed time series for station 0555567 W

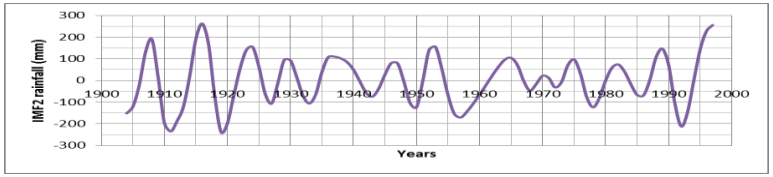


Figure B50-IMF2 of the decomposed time series for station 0555567 W

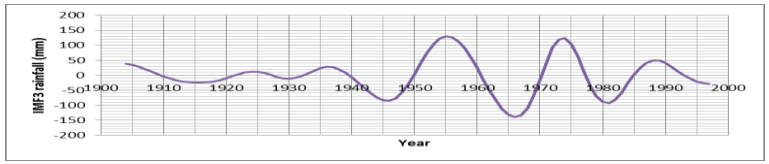


Figure B51- IMF3 of the decomposed time series for station 0555567 W

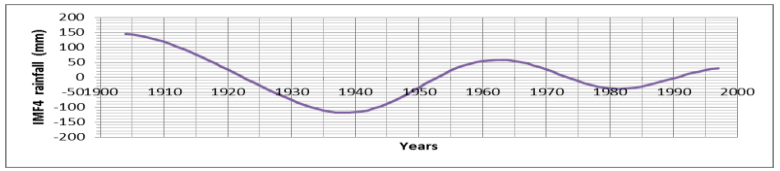


Figure B52 –IMF4 of the decomposed time series for station 0555567W

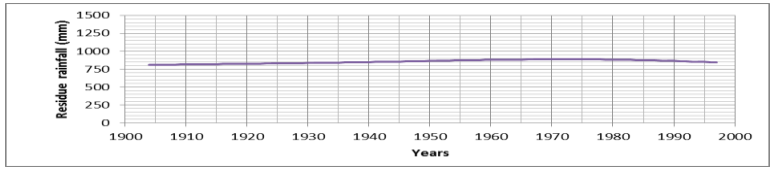


Figure B53-Residue of the decomposed time series for station 0555567 W

Analysis of decomposition derived from station 0052590 W

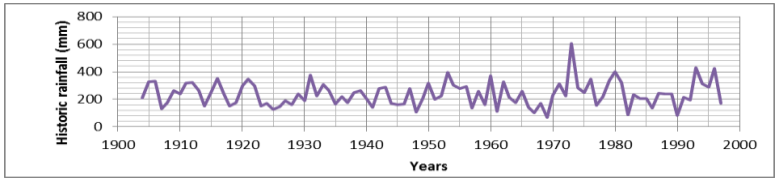


Figure B54-Historic rainfall series for station 0052590 W

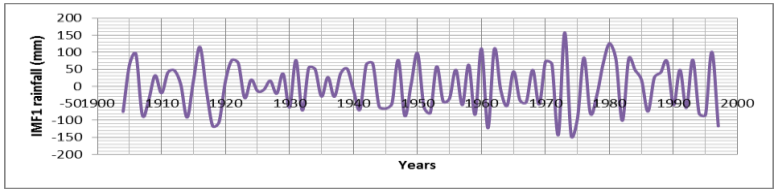


Figure B55- IMF1 of the decomposed time series for station 0052590 W

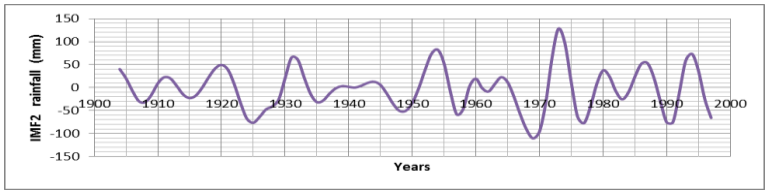


Figure B56- IMF2 of the decomposed time series for station 0052590W

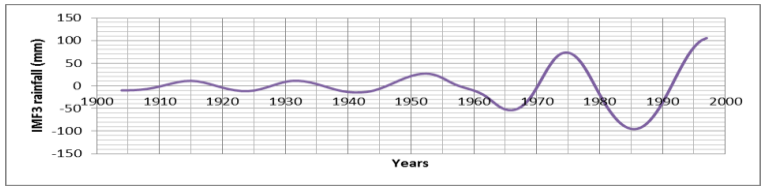


Figure B57- IMF3 of the decomposed time series for station 0052590 W

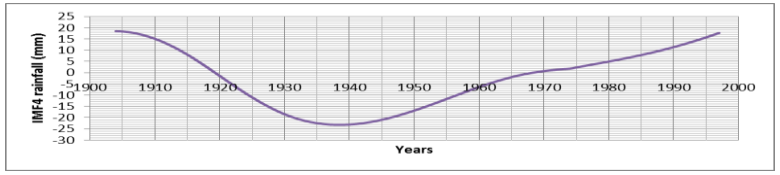


Figure B58- IMF4 of the decomposed time series for station 0052590 W

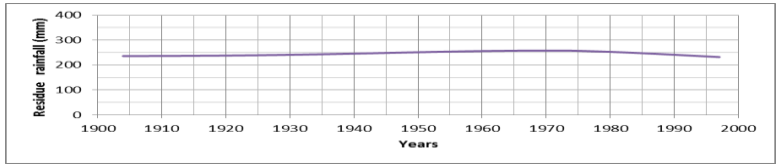


Figure B59-Residue of the decomposed time series for station 0052590 W

Analysis of decomposition derived from station 0678776 W

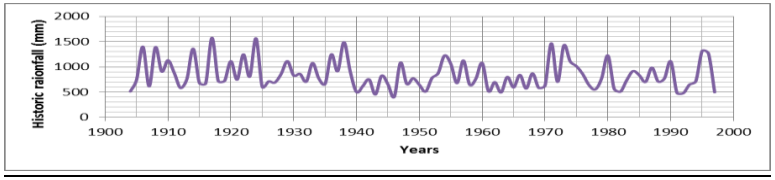


Figure B60-Historic rainfall series for station 0678776 W

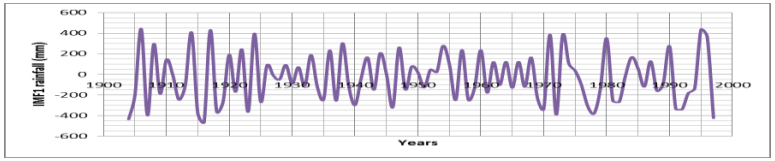


Figure B61- IMF1 of the decomposed time series for station 0678776 W

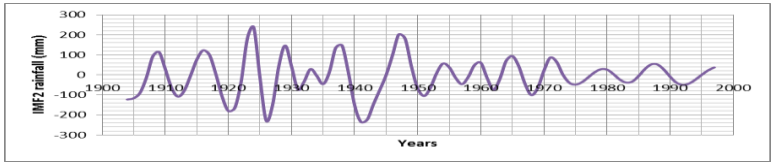


Figure B62- IMF2 of the decomposed time series for station 0678776 W

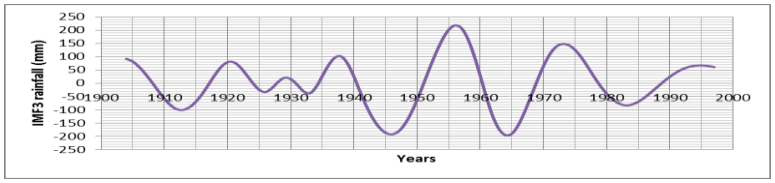


Figure B63- IMF3 of the decomposed time series for station 0678776 W

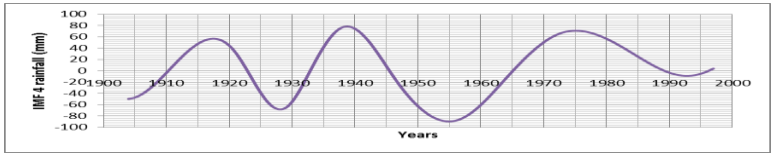


Figure B64- IMF4 of the decomposed time series for station 0678776 W

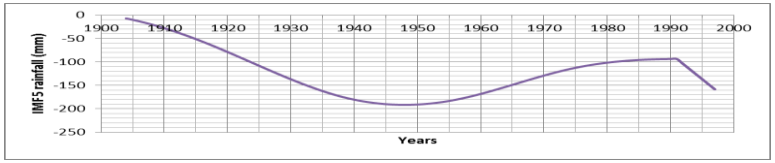


Figure B65- IMF5 of the decomposed time series for station 0678776 W

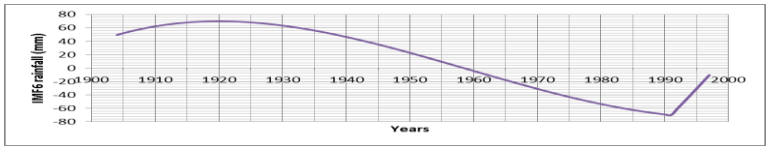


Figure B66- IMF6 of the decomposed time series for station 0678776 W