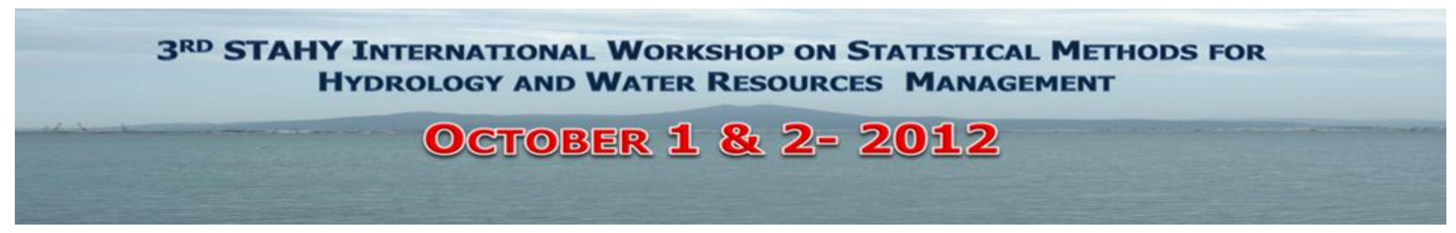
Empirical Mode Decomposition using skeletonization

Prashanth R. Marpu, Taha B.M.J. Ouarda, and Hosni Ghedira
Water and Environmental and Engineering Program, Masdar Institute of Science and Technology, PO Box 54224, Abu Dhabi, UAE
Email addresses: pmarpu@masdar.ac.ae, touarda@masdar.ac.ae, hghedira@masdar.ac.ae



1. Introduction

- Empirical Mode Decomposition (EMD) or the Hilbert-Huang transform (Huang et al., 1998) captures the non-stationary oscillations in hydro-climatic variables by decomposing the time-series signal into different components showing various frequencies in the signal as intrinsic mode functions (IMFs).
- Traditionally, EMD uses mathematical functions such as cubic splines or rational functions to interpolate the local extrema and thereby extract the IMFs by sifting through the centroid curve of the interpolated functions (Wu et al., 2007). Such approaches are not robust to outliers and can result in wrong IMFs.
- In this work, we employ a cognitive approach to estimate the IMFs. Skeletons are considered as the best solutions for the generalization of outlier induced edges of the polygons based on the observations in cognitive science (Aichholzer et al., 1995).
- An application to the prediction of annual precipitation data is presented where the precipitation IMF oscillatory pattern is predicted based on dominant low-frequency climate oscillation indices and the proposed approach is compared to the traditional EMD.

2. Data

The precipitation records for 105 years from 1901-2005 of 117 stations in eastern Canada presented by Vincent et al. (2002) are used in this study. The data of each station is normalized by subtracting the mean and dividing by the standard deviation to eliminate the geographical differences. The mean of 117 scaled data sets for each year is taken and the data set, called the Eastern Canada Scaled Mean Winter Precipitation (ESWP), is employed in the following analysis as shown in Figure 1.

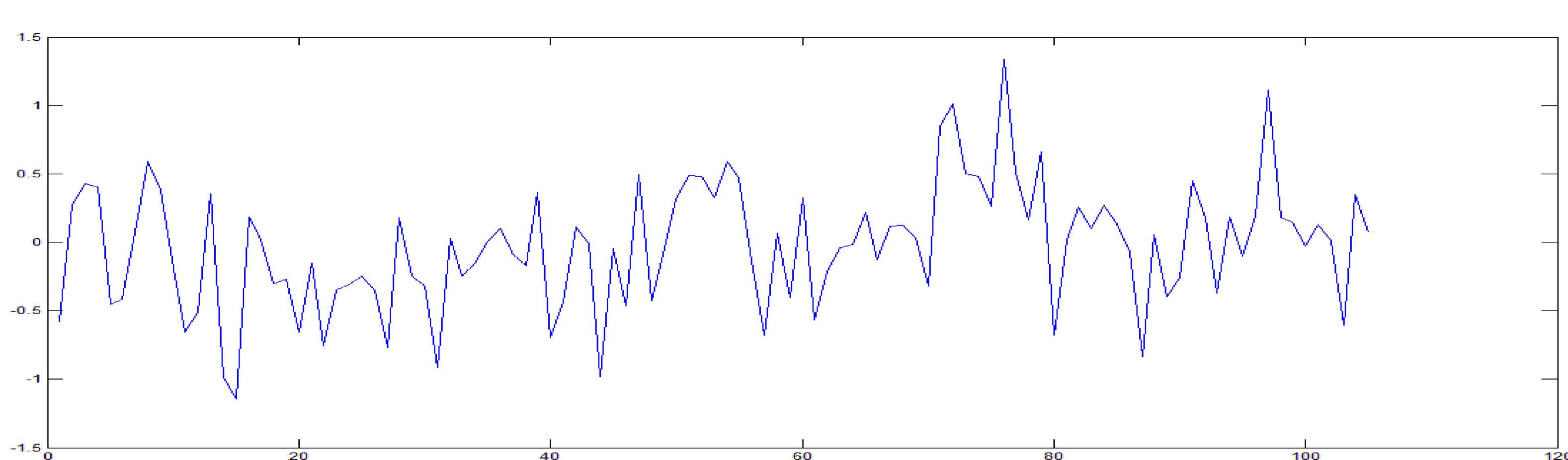


Figure 1: Time series showing scaled annual mean winter (DJFM) precipitation from 1901 to 2005.

3. Methodology

Empirical Mode Decomposition

A time series can be decomposed into a finite number of oscillatory modes whose frequencies are significantly different from each other.

The procedure to obtain the IMFs [Wu et al., 2007] is:

1. Find all the local maxima and minima of the time series .
2. Connect all these local maxima and minima with a cubic spline to form the upper and lower envelopes.
3. The medial line is the average of the upper and lower envelopes (see Fig. 2).
4. Obtain the first component, ***g***, from the difference between the data and the medial line.
5. Repeat steps 1- 4 using ***g*** as the data until the residual between consecutive residuals is smaller than a critical value.
6. Designate the final ***g*** as the first IMF component.
7. Repeat steps 1-6 for the following IMFs by replacing the original data with the difference between the original data and the preceding IMF.

In this work, steps 2 and 3 are replaced by estimation of medial line using skeleton of the polygon.

Skeletonization

- The time series is represented as a closed polygon by connecting the extrema and the corresponding skeleton is estimated (see Fig. 3).
- Straight skeleton representation can be used to represent the medial line (Aichholzer et al., 1995).
- However, a more simple method is used in this work. The skeleton is estimated by combining the center of the consecutive incenters of the Delaunay triangles of the polygon.
- Using the incenters reduces the effect of outliers which are represented as acute angle triangles. A cubic spline interpolation is applied to the obtained skeleton to obtain a smoother line.

4. Results

The proposed skeleton-based EMD is applied to the precipitation data shown in Fig. 1. To demonstrate the effectiveness of the proposed method in the presence of outliers, the method is tested by introducing an outlier at t=31 with a value of -3. The results of EMD using traditional approach and the proposed skeleton based approach are shown in Fig. 4. It can be seen in the results of the traditional EMD are not stable when the outlier is introduced. However, the effect of the outlier is restricted to the top components in the skeleton-based approach as shown in Fig. 4(d). It can also be noted that the IMFs of traditional and skeleton-based approach are different.

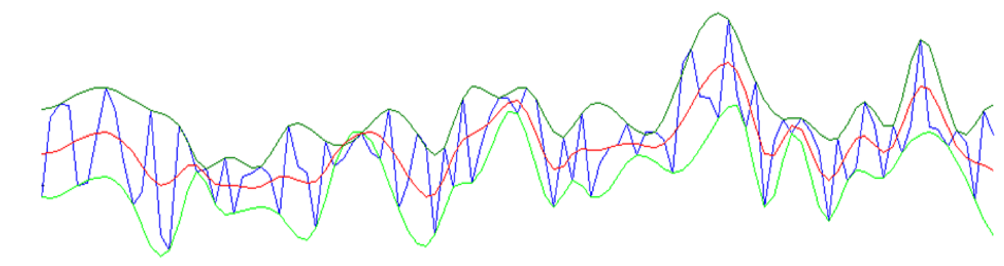


Figure 2: Medial line as mean of envelopes

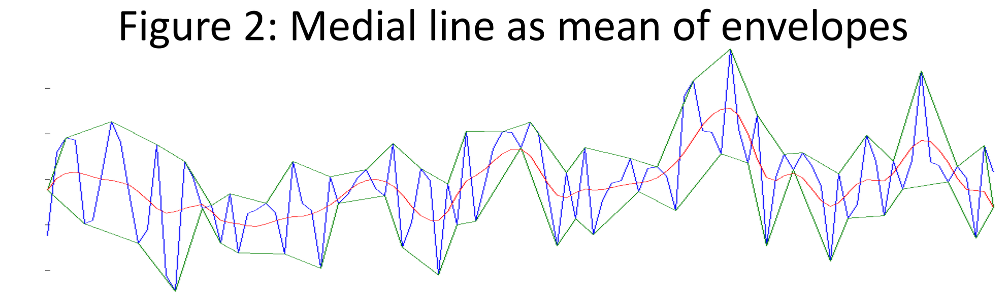


Figure 3: Medial line using skeleton

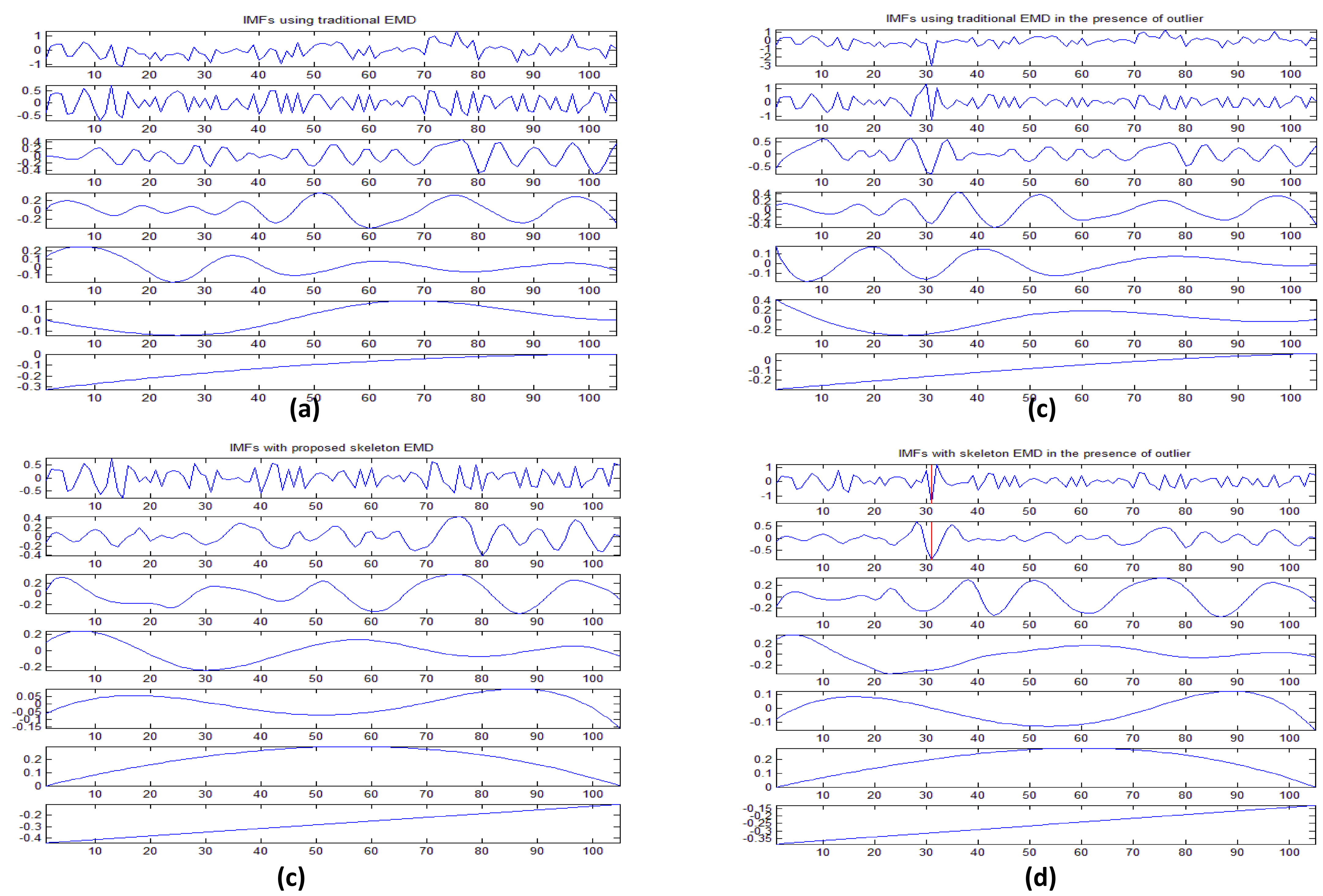


Figure 4: Comparison of EMD with traditional approach (a) and skeleton approach (c). The corresponding effect of outlier is shown in (b) and (d)

5. Conclusion and Future Work

- We propose to use skeleton of the polygon to represent the medial line in the estimation of the intrinsic mode functions in empirical mode decomposition.
- The usage of skeleton is found to be more adaptive to outliers in the data.
- In the future, the proposed method will be used in the ensemble EMD framework proposed by Wu and Huang (2009) and experiments will be carried out on a wide range of time series data.

References

[1] Aichholzer, Oswin; Aurenhammer, Franz; Alberts, David; Gärtner, Bernd (1995). "A novel type of skeleton for polygons". *Journal of Universal Computer Science* **1** (12): 752–761.
[2] Huang, N. E., Z. Shen, S. R. Long, M. L. C. Wu, H. H. Shih, Q. N. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu (1998), The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. R. Soc. London Ser. A*, 454(1971), 903–995.
[3] Wu, Z. H., N. E. Huang, S. R. Long, and C. K. Peng (2007), On the trend, detrending, and variability of nonlinear and nonstationary time series, *Proc. Nat. Acad. Sci. U. S. A.*, 104(38), 14889–14894.
[4] Wu, Z. H., and N. E. Huang (2009), Ensemble empirical mode decomposition: A noise-assisted data analysis method, *Adv. Adaptive Data Anal.*, 1(1), 41.
[5] Vincent, L. A., X. Zhang, B. R. Bonsal, and W. D. Hogg (2002), Homogenization of daily temperatures over Canada, *J. Clim.*, 15(11), 1322– 1334.