Hydrologic regionalization using wavelet-based multiscale entropy method

A. Agarwal a,b, R. Maheswaran a, V. Sehgal a,c, R. Khosa a, B. Sivakumar d,e,⇑, C. Bernhofer b

a Department of Civil Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110 016, India
b IHM, Technical University Dresden, 01737 Dresden, Germany
c Department of Biological Systems Engineering, Virginia Polytechnic and State University, VA 24061, USA
d School of Civil and Environmental Engineering, The University of New South Wales, Sydney, NSW 2052, Australia
e Department of Land, Air and Water Resources, University of California, Davis, CA 95616, USA

1. Introduction

Estimates of streamflow are a prerequisite for solving a number of engineering and environmental problems. These include design or dimensioning a water control structure, economic evaluation of flood protection projects, land use planning and management, water quality control, and stream habitat assessment, among others. When the availability of streamflow records is limited at the site of interest, it is a common practice to apply regionalization techniques to derive the streamflow quantile estimates at the sites where records are limited or in ungaged catchments (Kokkonen et al., 2003). Regionalization can be defined as the transfer of information from one catchment to another (Blöschl and Sivapalan, 1995). This information may comprise characteristics describing hydrologic data or models. To have greater confidence in extrapolating hydrologic behavior from catchments with flow records to an ungaged catchment, all these catchments should form a relatively homogeneous group (Pilgrim et al., 1988; Nathan and McMahon, 1990; Post and Jakeman, 1999). The homogeneity is not only in terms of geographic contiguity but also in terms of hydrologic similarity.

Some of the common approaches for regionalization in hydrology include: the method of residuals (MOR) (Choquette, 1988), the region of influence (ROI) approach (Zrinji and Burn, 1994, 1996), principal component analysis (PCA) (Singh et al., 1996), and cluster analysis and its extensions (Rao and Srinivas, 2006a,b; Isik and Singh, 2008; Srinivas et al., 2008; Satyanarayana and Srinivas, 2011); see also Razavi and Coulibaly (2013) for a review of regionalization methods for streamflow prediction in ungaged basins, and Sivakumar et al. (2015) for a comprehensive account of catchment classification more broadly. Nathan and McMahon (1990) used a combination of multiple regression, cluster analysis, principal component analysis, and graphical representation of eighteen physical catchment variables for predicting the low-flow characteristics of 184 catchments in south-eastern Australia. Notwithstanding their ability to provide reasonable outcomes, these approaches have an important disadvantage in that they mainly rely on the pre-conceived notion of the factors that are thought to influence the behavior of the streamflow from a catchment and that these factors are measurable (Zoppou et al., 2002). In reality, however, the streamflow is a resultant of integrated effects of
The concept of entropy, when applied in conjunction with wavelet analysis, can be used to determine the randomness (i.e. level of uncertainty) of a time series at different timescales. At a given scale, maximum entropy is possible when the information is evenly spread across time, and minimum entropy occurs when all the information is contained in a single location. The Wavelet-based Multiscale Entropy (WME), which is a measure of the degree of order/disorder of the signal and carries information associated with multi-frequency signal, can provide useful information about the underlying dynamic processes associated with the signal and can help in regionalization studies (Cazelles et al., 2008). This provides the motivation for the present study to develop a robust regionalization tool based on WME. In this study, the WME method is applied to monthly streamflow data observed at 530 stations in the contiguous United States. Continuous Wavelet Transform (CWT) is applied to each of the observed streamflow time series using the Morlet wavelet to capture the temporal multiscale variability of the streamflow in the form of wavelet coefficients. These wavelet coefficients for each scale are utilized to obtain the entropy for the respective scales. The spectral organization of this multi-spectral variability in terms of WME is identified using k-means clustering.

The rest of the paper is organized as follows. Section 2 describes the proposed methodology with description of the wavelet transform, WME, and k-means clustering technique. Details of the study area and dataset are presented in Section 3. Section 4 presents the application of the proposed methodology to streamflow data, followed by a discussion of the results. Finally, Section 5 presents some of the important conclusions and scope for further research.
2.2. Multiscale wavelet entropy

In order to gauge the complexity of a time series (such as the streamflow time series in this study), the wavelet coefficients produced from the CWT analysis of the time series (Section 2.1) can be utilized to obtain the multiscale wavelet entropy coefficient using the Shannon entropy measure (Shannon, 1948), which is defined as:

\[ S_{\text{wt}}(x) = - \sum_{i=1}^{n} P(x_i) \ln(P(x_i)) \]  

(2)

where \( p(x_i) \) is the probability distribution function (pdf) used to describe the random behavior of variable \( x \) with the length of \( n \). Entropy is a measure of the statistical variability of the random variable \( x \) as described by the pdf. The base of the logarithm is arbitrary, but if base \( \log \), is used, entropy is measured in bits. \( S_{\text{wt}}(x) \) is a measure of information content in the signal; more information represents a higher entropy value and less information represents a lower entropy value and vice versa. Therefore, a high value of entropy represents a high degree of unpredictability and, hence, a highly complicated and disordered hydrologic system.

In order to measure the pdf, \( P(x_i) \), in Eq. (2), Cek et al. (2009) proposed an entropy based on the wavelet energy distribution of a time series. Because the value of the entropy \( S_{\text{wt}}(x) \) calculated is based on the wavelet results, using this approach, Sang et al. (2011) were able to propose four entropy measures, namely continuous wavelet entropy, discrete wavelet entropy, continuous relative wavelet entropy, and discrete relative wavelet entropy. The present study uses the approach by Sang et al. (2011) to propose a new entropy measure, named wavelet-based multiscale entropy (WME).

The CWT-based pdf, \( P(x_i) \), is estimated according to the wavelet energy (i.e., variance):

\[ P(x_i) = \frac{E(i,j)}{E(j)} = \frac{|W(i,j)|^2}{\sum |W(i,j)|^2} \]  

(3)

where \( E(i,j) \) represents the wavelet energy under time position \( i \) and time scale \( j \) and \( E(j) \) represents the total wavelet energy of the time series under timescale \( j \) (Cek et al., 2009; Sang et al., 2011).

To illustrate the concept of multi-scale wavelet entropy, a synthetic time series \( S \) is analyzed. \( S \) is obtained through linear combination of a stationary time series \( S_1 \), a linear component \( S_2 \), a non-linear signal \( S_3 \), and random noise \( S_4 \) of range 0–10. These signals are mathematically described as follows:

\[ S_1 = \sin\left(\frac{2\pi t}{50}\right) + \cos\left(\frac{2\pi t}{60}\right) \]  

(4)

\[ S_2 = \frac{t}{200} \]  

(5)

\[ S_3 = 10 \cdot S_1^2 + S_1 + \frac{t^2}{100} \]  

(6)

\[ S_4 = \text{Random noise between } [0,10] \]  

(7)

\[ S = S_1 + S_2 + S_3 + S_4 \]  

(8)

Fig. 1 presents the resultant synthetic time series (top), plot of the wavelet coefficients (bottom left) and multiscale entropy (bottom right). The wavelet coefficients plot show that the given time series has features having periods 16–32 units. The plots of the wavelet coefficients show that there is a consistent feature around a scale of the order of 32–64 units. Also, it can be seen that there is a strong trend which is captured at the scale of the order of 128 units. It is evident that multiscale entropy is sensitive to these features of the time series. High value of entropy is observed at the periods 16–32 and 128 units, indicating low degree of orderliness and inconsistent features around these periods. Further, a dip in the multiscale entropy plot around the period 32–64 units indicates that an ordered feature exists in the signal around this period. The high value of entropy around the scale of 4 units are the resultant of the variability in the high-frequency components of the signal at that scale. The dip around the scale 8–16 corresponds to a consistently low strength but consistent feature around the scales. Lower values of entropy indicate orderliness, and higher values of entropy indicate variability.

2.3. k-means clustering

One of the most popular clustering algorithms is the k-means method, in which the data is partitioned into k clusters, with each cluster represented by its centroid, which is the mean (weighted or otherwise) of feature vectors within the cluster. If \( N_k \) represents the number of feature vectors in cluster \( k \) and \( C_k \) is the mean of cluster \( k \), then the centroid of each cluster is calculated using Eq. (9)

\[ C_k = \frac{1}{N_k} \sum_{p=1}^{N_k} X_p \]  

(9)

The algorithm starts with a pre-defined initial number of clusters \( k \) chosen according to some criteria or some heuristic procedure. In each iteration, each cluster is assigned to its nearest cluster center according to the Euclidian distance measure between the two, and then the cluster centers are re-calculated (Rokach and Maimon, 2005) until convergence of the algorithm occurs as per the defined criteria, e.g. when the algorithm exceeds the pre-defined number of iterations or when partitioning error is not going to reduce further on re-allocating cluster centroid, indicating that solution is locally optimal.

The method is known for its low run time, its efficiency in clustering large data set with numerical attributes (Rao and Srinivas, 2008), and simple implementation and interpretation since no parameters (except the number of clusters) are involved. The linear complexity is also one of the reasons for the popularity of the k-means algorithm. More detailed information about the k-means clustering method can be obtained from Ball and Hall (1967) and MacQueen et al. (1967), among others.

2.4. Wavelet-based multiscale entropy methodology

Fig. 2 shows the schematic of the methodology proposed in the present study. The streamflow data from all the stations are standardized (by subtracting the mean and dividing by the standard deviation) and the continuous wavelet transform (CWT) is applied to each time series using Morlet wavelet to obtain wavelet coefficients at different time–frequency scales (2, 4, 8, 16, 32, 64, and 128 months). The entropy is estimated at each scale using the Shannon entropy concept. The entropy estimated at each scale will capture the nature of the process at every scale, since wavelet coefficients are localized in nature and provide a better measure of variance attributed to localized events. Then the multiscale entropy signature from each station is used as the basis to form homogeneous clusters using the k-means clustering technique.

The selection of a suitable number of clusters is an important task while dealing with the k-means clustering. In this study, four validity indexes, namely the Davies–Bouldin (DB) index, the Dunn’s index, the Homogeneity index, and the Separation index, are used to identify the optimum number of clusters for the data.
2.4.1. Davies–Bouldin (DB) index

The Davies–Bouldin (DB) index (Davies and Bouldin, 1979; Kasturi et al., 2003), which represents the internal cluster evaluation criterion, is the most popular and widely used index in hydrology, due to their ability to identify optimal number of clusters that are well separated and compact. The DB index is defined as a function of the ratio of the sum of within-cluster scatter to between cluster separations, given by:

$$DB = \frac{1}{k} \sum_{i=1}^{k} \max_{j \neq i} \left\{ \frac{diam(C_i) + diam(C_j)}{||C_i - C_j||} \right\}$$  \hspace{1cm} (10)

where in this case, the diameter of the cluster is defined as:

$$diam(C_i) = \left( \frac{1}{n_i} \sum_{x \in C_i} ||x - z_i||^2 \right)^{1/2}$$  \hspace{1cm} (11)

where $n_i$ is the number of points and $z_i$ is the centroid of cluster $C_i$. Since the objective is to obtain cluster with minimum intra-cluster distances, a small value of DB is important.

2.4.2. Dunn’s index

The Dunn’s index (Dunn, 1973; Halkidi et al., 2001; Bolshakova and Azuaje, 2003), which is also an internal cluster evaluation criterion, is defined as the ratio of the minimal intra-cluster distance to the maximal inter-cluster distance. The Dunn’s index for $k$ clusters is defined as:

$$DU = \min_{i \neq j} \left\{ \min_{m=1}^{k} \left( \frac{\text{diss}(C_i, C_j)}{\max_{m \neq i, j} \text{diam}(C_m)} \right) \right\}$$  \hspace{1cm} (12)

where $\text{diss}(C_i, C_j)$ is the dissimilarity between clusters $C_i$ and $C_j$, given by:

$$\text{diss}(C_i, C_j) = \min_{x \in C_i, y \in C_j} ||x - y||$$  \hspace{1cm} (13)

and $\text{diam}(C)$ is the intra-cluster function (or diameter) of the cluster, given by:

$$\text{diam}(C) = \max_{x, y \in C} ||x - y||$$  \hspace{1cm} (14)

A large value of Dunn’s index is preferred, as it represents well a compacted cluster.
2.4.3. Homogeneity index and Separation index

The Homogeneity index is calculated as the average distance between each gene expression profile and the center of the cluster it belongs to. Mathematically, it is represented as:

\[ H_{\text{ave}} = \frac{1}{N_{\text{gene}}} \sum_{i} D(g_i, C(g_i)) \]

(15)

where \( g_i \) is the \( i \)th gene and \( C(g_i) \) is the center of the cluster that \( g_i \) belongs to, \( N_{\text{gene}} \) is the total number of genes, and \( D \) is the distance function (Chen et al., 2002).

The Separation index is calculated as the weighted average distance between cluster centers, and is defined as:

\[ S_{\text{ave}} = \frac{1}{\sum_{i} N_i N_j} \sum_{i} N_i N_j D(C_i, C_j) \]

(16)

where \( C_i \) and \( C_j \) are the center of the \( i \)th and \( j \)th clusters, respectively, and \( N_i \) and \( N_j \) are the number of genes in the \( i \)th and \( j \)th clusters, respectively.

Thus, \( H_{\text{ave}} \) reflects the compactness of the clusters, while \( S_{\text{ave}} \) reflects the overall distance between the clusters. Decreasing \( H_{\text{ave}} \) or increasing \( S_{\text{ave}} \) suggests an improvement in the clustering results (Chen et al., 2002).

3. Study area and data

In this study, streamflow data from the contiguous United States are analyzed to test the effectiveness of the proposed WME methodology for regionalization purposes. The streamflow records are selected from the US Geological survey (USGS) Hydro Climatic Data Network (HCDN). The HCDN dataset contains streamflow observations from the US Geological survey streamgages that are considered to be relatively unaffected by anthropogenic influences, land use changes, measurement changes, and measurement error. The data considered in this study are those observed over a period of 52 years (1951–2002), based on data availability from all stations. It is relevant to point, at this stage, that some recent studies have analyzed streamflow data from the HCDN database in the context of spatial variability and catchment classification, using concepts of nonlinear dynamics (correlation dimension method and false nearest neighbor algorithm) and complex networks (e.g. Sivakumar and Singh, 2012; Sivakumar and Woldemeskel, 2014; Vignesh et al., 2015).

The following criteria are adopted for selecting streamflow stations for this study, so that the data are of sufficient quality for reliable outcomes in regionalization analysis:

- Unregulated (virgin) streamflow: Regulation by multipurpose or flood-control reservoirs reduces peak flows and generally augments low flows. The extent of alterations in natural streamflows caused by regulation and local diversions is variable and difficult to quantify accurately;
- Area between 50 and 2000 sq. km: It has been previously established that catchments that are either less than 50 sq. km or greater than 2000 sq. km do not give good results in case of regionalization study (Zoppou et al., 2002); and
- Nested catchments are not considered even when the above criteria are satisfied.

Based on these criteria, 530 streamflow stations are selected for the study. Fig. 3 shows the locations of streamflow stations used in this study.

4. Results and discussion

The wavelet multiscale entropy is estimated, across seven different scales (i.e. 2, 4, 8, 16, 32, 64, and 128 months), for each of the streamflow series from the 530 stations considered in this study. Fig. 4 shows, for example, the streamflow time series observed from the St. John River at Dickey, Maine (USGS Station #01010500) (Fig. 4a) and the results from the wavelet analysis. The wavelet power spectrum (Fig. 4b) for this series shows the presence of sub-annual and decadal (64 months to 128 months) features apart from the annual cycles. The relative power of each of these features is shown in the wavelet global power spectrum (Fig. 4c). Fig. 4d shows the multiscale entropy plot for the same time series. It shows a peak at the scale of 4–8 months, indicating high variability in that scale, and a low value of entropy at 8–16 months scale, indicating the presence of consistent feature at this scale.

The entropy values across different scales for all the 530 streamflow series are now used as a basis for clustering of the stations. The optimal number of clusters is decided using each of the four validation indexes described above. Fig. 5 presents the values of the four indexes. It can be observed that the Dunn index and the Davies–Bouldin index show an optimum value of 18 and 15 clusters, respectively, whereas the Homogeneity index and the Separation index show that after 14 clusters there is no significant decrease in these values. Therefore, the possible number of clusters can be 14, 15, and 18. Out of these, the cases of 15 and 18 clusters are neglected, as the difference between the separation index and the homogeneity index for 14 clusters against 15 and 18 clusters are not that significant. It is also observed that when working with 15 and 18 clusters, some of the clusters have very few (one or two) stations. This may be one of the reasons that makes the Dunn index to suddenly fall to a low value. When the number of clusters is chosen to be 14, the overall distribution of the stations in the clusters is justifiable. Therefore, the optimum number of clusters is chosen to be 14.

Based on these observations, the 530 streamflow stations are segregated into 14 clusters using the WME based method, as explained earlier. Table 1 shows the number of stations that falls in each of the cluster category, and Fig. 6 shows the geographic locations of the stations in each of the clusters. As can be seen, apart from geographic contiguity, the clustering shows that there is hydrologic similarity in the clusters. It is observed that some of the stations in a given cluster are spread across the study area showing that the basis of clustering is not the geographic proximity. The stations in each of these clusters are further examined for any common characteristics (in terms of multiscale entropy) they may have among themselves.

Fig. 7 shows, for example, the multiscale entropy values for the stations in five selected clusters: Clusters 2, 5, 8, 9, and 12. As can be seen, the multiscale entropy values are, to a great extent, similar within any given cluster. As Fig. 7 shows, the basis of the clustering is the entropy signature of the streamflow observed at all the stations within a given cluster. For example, in Cluster 2 (Fig. 7a), the entropy signatures for all the stations are more or less similar in nature. The peaks in the plots correspond to high value of entropy, which corresponds to high variability of the feature at that scale across time. Furthermore, the pattern of the entropy in a given cluster across all scales for the stations is unique for that cluster but also different from each other.

To make the analysis more meaningful and simpler, the average entropy over a given cluster can be used instead of the individual station-wise entropy values. The average entropy of all member stations of a cluster at each scale is taken as the representative value of entropy for that cluster for the given scale. Therefore,
Fig. 3. Locations of the 530 streamflow stations in the United States.

Fig. 4. (a) Streamflow time series from St. John River at Dickey, Maine (USGS Station #010105). (b) Wavelet power spectrum. (c) Global power spectrum, and (d) Multiscale entropy.
Fig. 5. Selection of optimum number of clusters based on: (a) Dunn index. (b) Davies–Bouldin index. (c) Homogeneity index, and (d) Separation index.

Table 1
Number of stations in each cluster for 530 stations in the United States.

<table>
<thead>
<tr>
<th>Cluster no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of stations</td>
<td>28</td>
<td>50</td>
<td>65</td>
<td>43</td>
<td>67</td>
<td>5</td>
<td>59</td>
<td>7</td>
<td>33</td>
<td>33</td>
<td>52</td>
<td>13</td>
<td>49</td>
<td>26</td>
</tr>
</tbody>
</table>

Fig. 6. Cluster-wise geographical distribution of streamflow stations.
average entropy values at each scale are obtained for each of the cluster and further analysis is carried out. Fig. 8 shows, for example, the average entropy at all scales over all stations in Cluster 5. Cluster 5 is chosen here for illustration, as it contains the maximum number of streamflow stations (67) (see Table 1). As can be seen, variation of the entropy values across scales for all stations in a given cluster is minimum and the WME values for stations in a given cluster display a similar pattern. Similar results are also observed for each of the other clusters. The results also indicate that, in most of the cases, the WME value first decreases to a minimum value for scales $\leq 2$ months. This decrease is possibly due to the absence of significant features at these scales.

The scale of 3–8 months marks a constant increase in WME value indicating a high variability at these scales. For the 9–13 month scale, WME first increases and then decreases. The steep increase in WME observed up to a scale of 11 months can be attributed to a decrease in information and increased randomness of wavelet coefficients for intermediate scales. As all streamflow stations have very distinct annual patterns, the decrease in WME at around 12–13 months scale is well justified. Beyond the 14-month scale, however, the trend in WME becomes highly variable. This may be due to the presence of high irregularity in the presence of features at these scales and, hence, such scales not considered for further discussion.
Based on these observations, four distinct bands are identified for further analysis. Band 1 captures the features up to 2-month scales. Band 2 and Band 3 capture the features from 3 months to 8 months and scales from 9 months to 14 months, respectively. The features having a scale beyond 14 months are grouped under Band 4. Fig. 9 shows the average normalized WME for all the clusters at different bands (the WME values are normalized for better comparisons). It can be seen that there is a clear distinction in the values of WME for different clusters for the first three bands. In view of this, information from the first three bands are further analyzed.

The WME of each cluster is further classified into “High”, “Medium” and “Low” categories, based on the position of the individual WME plot with respect to the mean level for that band. For instance, if the WME of a cluster, for a given band, falls below the mean of WME of all clusters, then that particular cluster is assigned a signature of “Low”. Using this classification, an entropy signature is given to each cluster based on the entropy values in...
the three scale-based bands. For notational simplicity, the classifications “High”, “Medium”, and “Low” are represented by “1”, “0”, and “-1”, respectively. This means, for example, that an entropy signature of (0, -1, 1) would indicate that the cluster has relatively moderate entropy up to 2 months, low entropy for 3–8 months and high entropy value for 9–14 months. Based on these notations, the entropy signature for each of the 14 clusters is given in Table 2.

As a further step in the analysis, an attempt is made to relate the WME values at different scale-based bands to their respective catchment areas. Fig. 10 is a box plot of drainage areas of streamflow stations for each cluster. It is observed that the clusters that have a “High” entropy for the scale 9–13 months (i.e. Clusters 3, 4, 5, 7, and 11) have characteristically small drainage areas. On the other hand, the clusters that is characterized by “Low” entropy for the scale 9–13 months (i.e. Clusters 1, 6, and 8) have large drainage areas. This observation corroborates our general intuition that a catchment with a smaller area will be characterized by high variability and unstable properties.

In addition, an effort is made to identify geographic locations of the stations that fall within each of the clusters, and the following observations are made. Catchments in Cluster 1 are generally located in central lowland area. Cluster 5 shows geographic contiguity around the central lowland, Appalachian plateau, and nearby regions. Clusters 2, 9, 10, 11, 13, and 14 contain catchments having intermediate drainage areas, and all these stations also show moderate values of WME at high scales, with the exception of Clusters 11 and 14, which still show high WME value at high scales. Note that catchments of Clusters 2, 9, 10 are more or less located in the coastal plains, and catchments in Cluster 14 are located in Sierra Mountains, while those in Cluster 11 are located in Ozark plateaus. Furthermore, Clusters 3, 4, 7 have catchments of low drainage area. Cluster 4 shows flavor of geographic contiguity among the stations present in the Piedmont province. Cluster 7 has catchments scattered on the eastern part of the US. Stations in Cluster 3 are located in the New England region and Appalachian plateau.

All these observations indicate the effectiveness of the present methodology for catchment regionalization. However, they still remain somewhat preliminary, as there are other factors too that play a vital role in determining the properties of a catchment. An in-depth analysis is, therefore, clearly needed to confirm the present results. Such an investigation, however, is beyond the scope of the present study, and will be undertaken in a future study.

5. Conclusions

This study has presented a novel method for catchment regionalization using multiscale wavelet entropy. Application of the method to streamflow data from 530 monitoring stations in the contiguous United States offers promising results for regionalization. The results lead to the following highlights:

(a) The proposed k-means coupled wavelet-based multiscale entropy approach for regionalization of hydrologic catchments overcomes some of the limitations of the existing approaches (especially when data are limited) and is robust for hydrologic regionalization.

(b) Wavelet-based multiscale entropy measure captures the variability of streamflow dynamics at each station independently and then allows the formation of homogeneous clusters, which is not based on any prior assumptions.

(c) The wavelet-based multiscale entropy appears to be an important statistic in capturing the catchment characteristics. The 530 stations studied are categorized into 14 clusters, each having a distinct WME pattern across the different scales considered.

(d) Based on the pattern of the average WME for each cluster at the different scales, a characteristic signature is given to each catchment, which provides an approximation of WME of a catchment across scales 1–2, 3–8, and 9–14 months relative to other stations; and

(e) The precise cause of fluctuations in WME at different scales remains to be investigated, although drainage area and topography seem to be possible factors.

Applications of the proposed methodology to data from other parts of the world, with different climatic, hydrologic, and environmental characteristics, would also help verify, and possibly strengthen and confirm, the conclusions drawn here.

Acknowledgments

This research was funded by Department Science and Technology, India through the INSPIRE Faculty Fellowship held by Maheshwaran Rathinasamy. Bellie Sivakumar acknowledges the financial support from the Australian Research Council (ARC) through the Future Fellowship Grant (FT110100328). We thank the two reviewers for the constructive comments and useful suggestions on an earlier version of this manuscript. Their comments helped significantly improve the quality of the manuscript.

References
