Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

# Community Detection Using Modularity Optimization Method For Catchment Classification

Siti Nur'Ain Amat Idris[1], Siti Aisyah Tumiran[1,*]

[1] Mathematics Visualization (MathViz) Research Group, Faculty of Science and Natural Resources, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia

**ARTICLE INFO**

***Article history:***
Received 25 September 2023
Received in revised form 28 December 2023
Accepted 13 April 2024
Available online 22 May 2024

***Keywords:***
Catchment classification; Community Structure; Modularity; Complex Networks; Streamflow

**ABSTRACT**

A general framework for catchment classification may be helpful for more accurate and efficient modeling of hydrologic systems, as well as to improve communication between hydrology researchers and those in other disciplines. There are plethora numbers of methods applied for catchment classification, but in these years, recent studies are implementing the complex networks concept for classification purposes. The community structure methods which are complex networks-based methods are focus mainly to classify catchments. Hence, the efficiency of complex network ideas, especially using the methods of community structure is examined in this study. Specifically, the modularity optimization method that is one of the community structure methods is applied to classify 218 stream-gauges stations in entire Australia that covers a large variety of hydroclimatic, topographic, geomorphic, soil usage, and climatic parameters. In the present study, the applicability and the efficiency of the community structure concept is validated by the proposed method. The classification of Australian catchments was further assessed with threshold value of 0.8, which resulted formation of nine communities with at least 9 stations in a community which combine to have almost 77% of the total number of stations (165 out of 218). All nine selected communities were also examined in terms of the flow characteristics (i.e. flow mean and flow covariance) and the catchment characteristics (i.e. drainage area, elevation and stream length). The catchment behaviors for each selected communities were also interpreted in terms of distance and correlation relationship, which give some useful insights towards generalization of hydrologic framework.

## 1. Introduction

Catchment classification is needed for a variety of environmental, hydrologic and ecological research, especially for determining the right level of model complexity, extrapolating and interpolating data scheme, and for environmental water assessment. Catchments are streamlined into groups and subgroups according to their key criteria in catchment classification. There are several strategies and techniques for classifying catchments. For examples those based on ecohydrologic and geomorphic factors, river/flow regimes, geostatistical properties, entropy

* *Corresponding author.*
*E-mail address: sitiaisyah.tumiran@ums.edu.my*

https://doi.org/10.37934/araset.45.2.7889

properties, hydrologic similarity indices, scale properties, data-driven methods and data-based mechanistic strategies as well as by hydrological responses and statistical regression analysis [1-3]. As a result, the complexity of various types of catchments were also varied. The primary issue with classifying catchments is that it necessitates studying the network of catchments and taking into account as much as many criterias. The number of catchments, the geographic area, the type and resolution of data (spatial and temporal), all affect the size and shape of such a network [4].

With the advancement of complex network science in recent years [5], community structure concept has been recently applied for studying streamflow as well as other hydrological data sets, especially rainfall [6-9]. Whichever approach is used, the primary idea in classification of catchment is to observe any existence of connection (for example, correlation or similarity indices) may occurs between a pair of catchments and then the strength of those connections will be use for grouping [1]. Numerous complex networks have nodes that organize into distinct groups, each of which becomes more densely coupled with the other networks. The qualities of each individual node and the network as a whole have little bearing on the traits of this group. The "community" is this group, and the "community structure" is this kind of network organization [10]. Walktrap, leading eigenvector, edge betweenness, modularity optimization, greedy algorithm and label propagation, are a few techniques for community structure in complex networks. According to studies based on community structure methods for classification of catchments, there are several studies and method for detecting communities in the network have been developed [7,11,12].

Some of the methods for community structure are based on modularity metrics, which measure the quality or strength of the community, such as edge betweenness [13], greedy algorithm [14], leading eigenvector [15], and multilevel optimization [16]. The present study is focused on the suitability of community structure method which is modularity-based specifically, the modularity optimization method to classify Australian catchments. As Australian catchment is one of the reliable and covering vast mass of various regions which are ideal for a ‘test bed’ to assess the performance for the proposed method. At the same time, this study is carried out to strengthen the assessment of the general suitability of community structure techniques especially modularity-based method for catchment classification. Therefore, applying the modularity optimization method across large areas and multiple river basins, which will almost certainly include a wide range of hydroclimatic, topographic, geomorphic, land use and other broader related features is crucial to study and analyse the effectiveness of such methods in community structure concepts. Therefore, the present study applied the modularity optimization method to monthly 218 streamflow stations in entire Australia. The catchment characteristics such as drainage area, elevation and stream length as well as flow characteristics such as flow mean and flow covariance (CV) will be examined for catchment behaviours understanding based on selected communities. Apart from that, the analysis of classification also will interpret the communities identification based on distance and correlation relationship.

This paper is organized as follows. In Methodology, the Modularity Optimization method is explained. Next section describes the study area of Australian catchments and the details of streamflow data used is provided. The study's findings and the classification of streamflow in Australia are discussed and interpreted in results and discussion section. In final section offers suggestions for more research directions and make some conclusions.

## 2. Methodology

Network is consists of a set of nodes connected by links when represent in hydrological research, streamflow are represented as nodes and links are denoting similarities between the pair of nodes.

Then, a set of nodes that share similarities will create communities [17]. Several methods for detecting communities in networks have been developed. Modularity (Q) value, which measures the strength or the quality of a community are highly hinged in several of these methods.

### *2.1 Modularity Optimization Method*

This modularity optimization method proposed by Blondel [16] uses a similar approach from the greedy algorithm method where initially, each node is placed to its own community. Moreover, as the aforementioned modularity function (or Q value) proposed by Newman and Girvan [15] is a well-known method to measure community structure. To identify community characteristics, the Q value will calculate the difference between the actual number of intra-community links and the expected number of links. Divisions with the highest modularity values have better community structure than divisions with the lowest modularity values. Modularity is a standard objective function in defining network clusters that measures the quality of network division into communities. Networks that have communities based on high modularity values are communities that have strong link connections between nodes in the community but have weak connections between nodes in other communities [18]. Modularity can also have a positive or negative value, where a positive value indicates the presence of community structure, and a negative value indicates the absence of community organization. Hence, the structure of communities can be accurately analysed by finding network partitions with positive and large modularity values [19].

Figure 1 shows a simple undirected network graph with six nodes and seven links where the connection between node $i$ and node $j$ is similar to the connection between node *j* and node *i*. The procedure to apply modularity optimization method is firstly, consider such a network has $n$ vertices/nodes. By dividing a given network into two subgroups let, $s_i = 1$ if vertex $i$ belongs to group 1 and $s_i = -1$ if it belongs to group 2. Then, adjacency matrix will be constructed where if there is exist a link between a pair of vertices $i$ and $j$, it will be denoted with 1 and 0, otherwise. The modularity function to calculate the modularity value (Q value) to each subgroup, used in the present study is shown in Eq. (1).

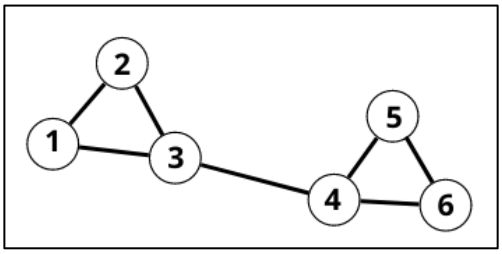



**Fig. 1.** A simple undirected network graph

$$Q = \frac{1}{2m}\sum_{i,j}(A_{ij} - \frac{k_i k_j}{2m})\delta(c_i, c_j) \quad (1)$$

where $A_{ij}$ represents the adjacency matrix, *m* is the count of edges, $k_i$ and $k_j$ represents the degree, from the adjacency matrix, $c_i$ is the count of all rows/columns, and $c_j$ represent a type (or group), and $\delta$ represents the Kronecker delta which is 1 if $c_i = c_j$ (where nodes *i* and j are within same type or group) and as 0, otherwise.

## 3. Study Area

Australia is the smallest continent and one of the largest countries on earth, where it is located in the southern hemisphere between the Pacific ocean and the Indian ocean. Australia's large lands extend approximately 2,500 miles (4,000 km) from west to east and nearly 2,000 miles (3.2 km) from the Cape York Peninsula in the northeast to Wilsons Promontory in the southeast (3,200 km) [20]. In the present study, a total of 218 streamflow monitoring stations across Australia were taken into account for the classification of catchments. The location of the considered streamflow stations are shown in Figure 2. The Hydrological Reference Station (HRS), where the database is maintained by the Australian Bureau of Meteorology (BoM), provided a total of 218 streamflow data. The length of observation and human impacts were two of the many considerations used to construct the HRS database. Zhang [20] provides information on the HRS database's selection by BoM. However, due to the presence of a handful of missing data, several stations from the HRS database were excluded from this study. The flow data used in this study spanned a period of 26 years, that is, from January 1981 to December 2006 and the data is the value for the monthly average. Table 1 shows a summary of the minimum and maximum values for some important features of the station or data, as well as the corresponding station numbers. As seen in Table 1, the 218 streamflow stations and streamflow data observed show enormous variations in their characteristics. For example, the drainage area is between 11.65$km^2$ (4.5 $mi^2$) to 603069.15 $km^2$ (232846.3 $mi^2$).

**Table 1**
Catchment Characteristics of Australian Streamflow Data

| | Minimum | Maximum | Station (State) |
|---|---|---|---|
| Latitude | $-43.14°$ | $-11.83°$ | Minimum: #473 $(\mathrm{TAS})^a$<br>Maximum:#$926002A$ (QLD) |
| Longitude | $115.44°$ | $153.42°$ | Minimum:#610008 (WA)<br>Maximum: #$146012A$ (QLD) |
| Drainage area ($km^2$) | 11.65<br>($4.5\ mi^2$) | 603069.15<br>($232846.3\ mi^2$) | Minimum: #235205 (VIC)<br>Maximum: #$A0030501$ (SA) |
| Elevation ($m$) | 5<br>($16.37\ ft$) | 2181.55<br>($7157.32\ ft$) | Minimum: #$G8140040$ (NT)<br>Maximum: #401012 (NSW) |
| Flow mean ($m^3/s$) | 0.36<br>($12.83\ ft^3/s$) | 182.42<br>($6442\ ft^3/s$) | Minimum: #$A0030501$ (SA)<br>Maximum: #$112002A$ (QLD) |
| Flow standard deviation ($m^3/s$) | 0.944<br>($33.337\ ft^3/s$) | 233.9082<br>($8260.39\ ft^3/s$) | Minimum: #616013 (WA)<br>Maximum: $112002A$ (QLD) |
| Flow CV | 0.471 | 6.12 | Minimum: #226222 (VIC)<br>Maximum: #$G0010005$ (NT) |

*NT* Northern Territory; *TAS* Tasmania; *SA* South Australia; *QLD* Queensland; *WA* Western Australia; *VIC* Victoria; *NSW* New South Wales.

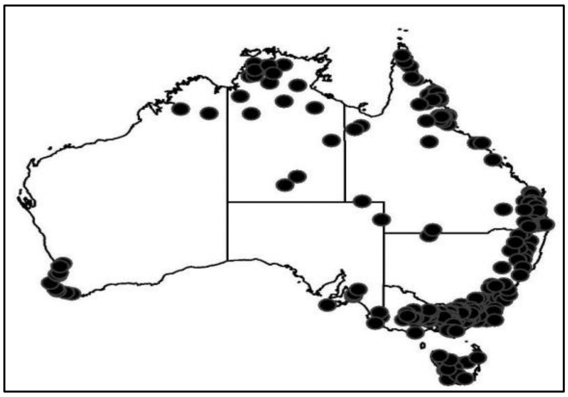


**Fig. 2**. The location of 218 hydrological stream flow monitoring stations in Australia

## 4. Results

The link between node pairs i.e., the station is given to classify catchments based on the Pearson correlation coefficient of catchments in implementing the modularity optimization method to 218 monthly catchment data in Australia. The correlation threshold range (T) is determined based on the examination of streamflow data and other hydrological characteristics using a network-based approach. This analysis is done to better represent the threshold influence value. T = 0.65, 0.7, 0.75, and 0.8 are the threshold levels that have been taken into account. Based on Figure 3 (a-d), different colors are used to visualize communities that have more than 9 stations while the smaller communities are in open circle and have no significance when comparing thresholds. Generally, based on overall observations in Figure 3, a very large count of links will be identified when low value of threshold is generated. As a result, it will form a large-scale community that will cover most of the study area, e.g. communities identified with T=0.65 (Figure 3 (a)). However, this will not help on studying the stream flows variability. As opposed to that, when the threshold value is high, a small number of connected links which eventually breaks the network into smaller and very close neighbors and lead to more remote communities. An accurate amount of number of communities with number of stations in Table 2 is obtained to show the improvement of physical explanations and interpretations for the classification of catchments, as well as a better understanding of hydrological similarities. According to Table 2, when the threshold rises, the number of stations for the largest communities drops while it rises for communities with just a few catchments (one and two catchments). However, the number of communities are also varies when the threshold value increases.

Communities identified based on threshold value, T=0.8 (Figure 3(d)) has been selected to allow for a better interpretation of catchment characteristics when referred to the classification by the division of catchments based on regions and boundaries. In particular, a total of 9 communities with at least nine stations were studied. By referring to Table 3, based on T = 0.8, there are six communities that have more than 10 stations of which 11, 16, 17, 24, 25 and 43 stations. These stations are combined to make up approximately 62% from total stations in the network (136 from 218) and the 9 largest communities (out of 48) have over 77% of the total number of stations when combined (165 from 218). This count has shown that, either the distance between the communities or whether they are located in different basins or regions, it can be suggested that every catchment in a large community has a strong relationship with other catchments in that community.

As referred to Table 2, almost 68% of the total count of identified communities (48) is consist of communities with only one catchment (32 out of 48) but only covers about 16% of the total count of stations (32 out of 218). Therefore, it is often assumed that each catchment in this tiny community has no link to other catchments or only has a tenuous connection to them, regardless of whether the community is present in the same basin or area. As a result, this study can be further explored by connecting the communities that have been identified with catchment or flow characteristics like station drainage area, station elevation and station flow length (as the characteristics of catchment), as well as the station flow mean and station flow coefficient of variation (CV) (as the characteristics of flow).

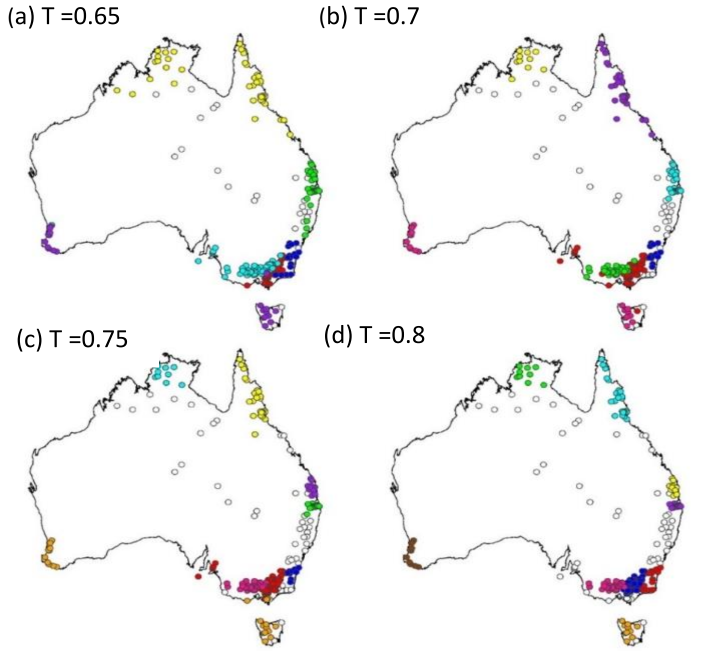



**Fig. 3 (a-d).** Classification using modularity optimization method for streamflow from Australia at four different correlation threshold values

**Table 2**
Count of communities identified in Australia using Modularity Optimization method at four correlation threshold values (T=0.65, 0.7, 0.75 and 0.8). NC is the number of communities, NS is the number of stations and NSC is the number of stations in the communities identified

| T = 0.65 | | | T = 0.7 | | | T = 0.75 | | | T = 0.8 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| NSC | NC | NS | NSC | NC | NS | NSC | NC | NS | NSC | NC | NS |
| 1 | 8 | 8 | 1 | 12 | 12 | 1 | 24 | 24 | 1 | 32 | 32 |
| 2 | 3 | 6 | 2 | 6 | 12 | 2 | 3 | 6 | 2 | 2 | 4 |
| 3 | 1 | 3 | 5 | 1 | 5 | 3 | 1 | 3 | 3 | 3 | 9 |
| 7 | 1 | 7 | 7 | 1 | 7 | 4 | 2 | 8 | 4 | 2 | 8 |
| 18 | 1 | 18 | 10 | 1 | 10 | 9 | 1 | 9 | 9 | 1 | 9 |
| 20 | 1 | 20 | 14 | 1 | 14 | 10 | 2 | 20 | 10 | 2 | 20 |
| 31 | 1 | 31 | 20 | 1 | 20 | 17 | 1 | 17 | 11 | 1 | 11 |
| 33 | 1 | 33 | 27 | 1 | 27 | 24 | 1 | 24 | 16 | 1 | 16 |
| 42 | 1 | 42 | 29 | 1 | 29 | 26 | 2 | 52 | 17 | 1 | 17 |
| 50 | 1 | 50 | 39 | 1 | 39 | 55 | 1 | 55 | 24 | 1 | 24 |
| | | | 43 | 1 | 43 | | | | 25 | 1 | 25 |
| | | | | | | | | | 43 | 1 | 43 |
| Total | 19 | 218 | Total | 27 | 218 | Total | 38 | 218 | Total | 48 | 218 |

The relationship between catchment characteristics (drainage area, elevation mean and stream length) and flow mean from 9 selected communities (165 stations) is presented in Figure 4(a-c). By referring to Figure 3 (d), a total of nine largest communities are selected to be considered throughout the analysis. There are three communities located in the south-east (Community 1 is in red, Community 12 is in blue, and Community 34 is in pink). Then, in the north and northeast parts, community 22 (colored in green) and community 24 (colored in cyan). Meanwhile, in the eastern region, there are two communities (Community 26 is in yellow and Community 28 is in purple). There are also communities 38 (colored in orange) located in Tasmania and Community 44 (colored in brown) is located in the south-west region.

As seen in Figure 4(a), the relationship between drainage area and flow mean does not indicate a linear relationship. It can be seen that some stations from some communities are scattered, especially communities 22 (green), 24 (cyan), 28 (purple), 34 (pink), 44 (brown) and 38 (orange), however one particular community which is 12 is more clustered. Whereas, for the relationship between (stream length against flow mean) as in Figure 4(b), it is almost similar distribution as in Figure 4(a) which is not linearly correlated. It is interesting to observe that in Figure 4(c), when elevation mean against flow mean, there can be seen visible variability when in terms of elevation and flow mean which is almost a straight line especially for communities 1 (red), 12 (blue) and 34 (pink) which all communities are mostly located in the southeast region. So, it can be suggested that the factor of geographical proximity is most likely contributed to the relationship between the watersheds within the community.

As seen in Figure 5, the relationship between drainage area (Figure 5 (a)) and stream length (Figure 5(b)) against flow CV are more sparser than against flow mean (Figure 4(a-b)). The relationship between elevation mean and flow CV (Figure 5 (c)) are more stretcher than against flow mean (Figure 5(c)). Most communities such as communities 22 (colored in green), 24 (colored in cyan), 26 (colored in yellow), 28 (colored in purple), and 44 (colored in brown) are sparser which are still able to form as communities. This seems to indicate that this probably caused by the geographic proximity and the closeness factor within the catchments in each community where each of the community are formed in their respective regions and therefore indicating the efficacy of community structure method especially the modularity optimization method application for classifying catchments.

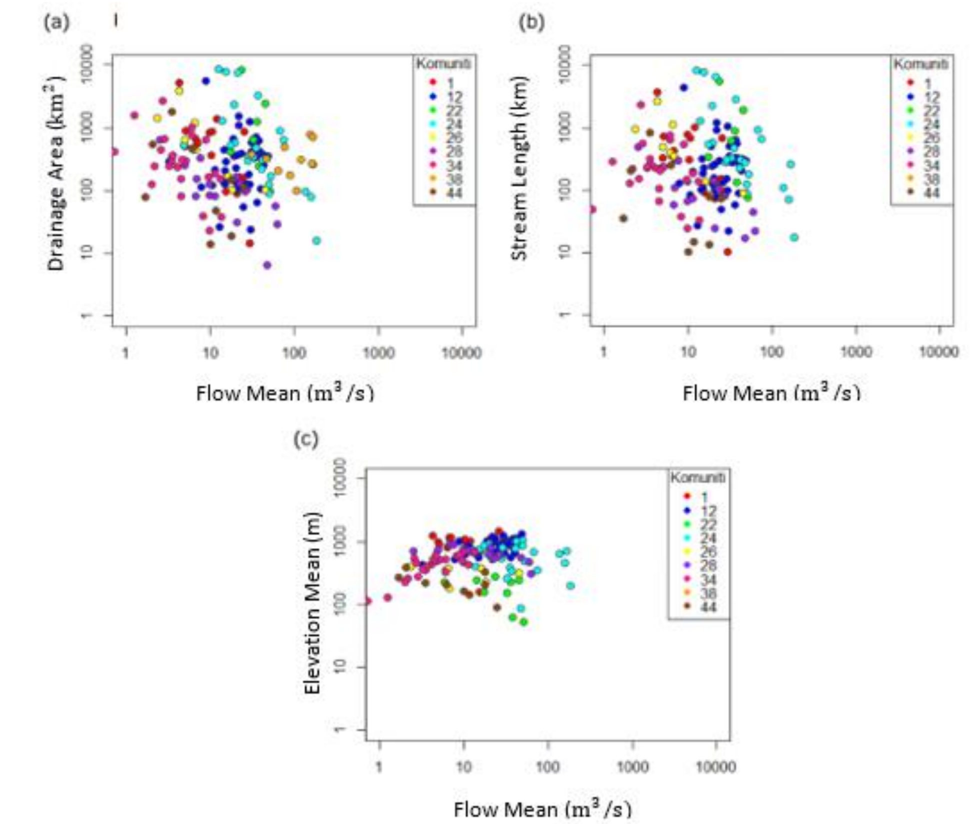



**Fig. 4 (a-c).** Relationship between, (a) drainage area, (b) stream length, and (c) elevation mean with flow mean for 9 communities identified in Australia

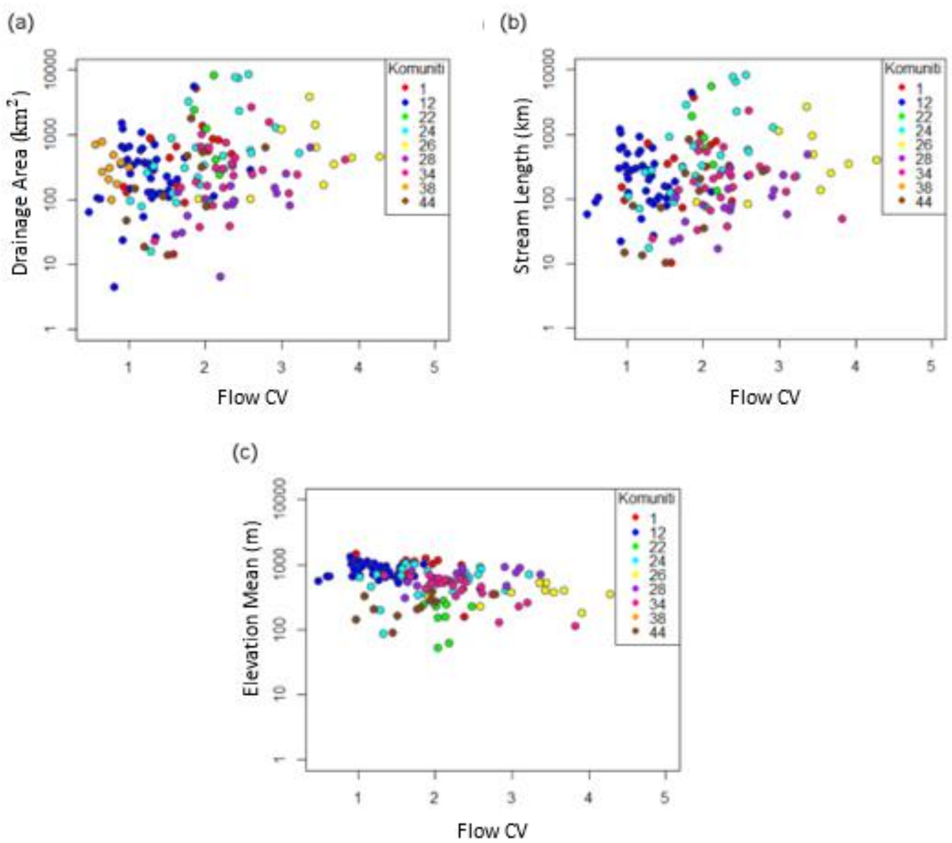
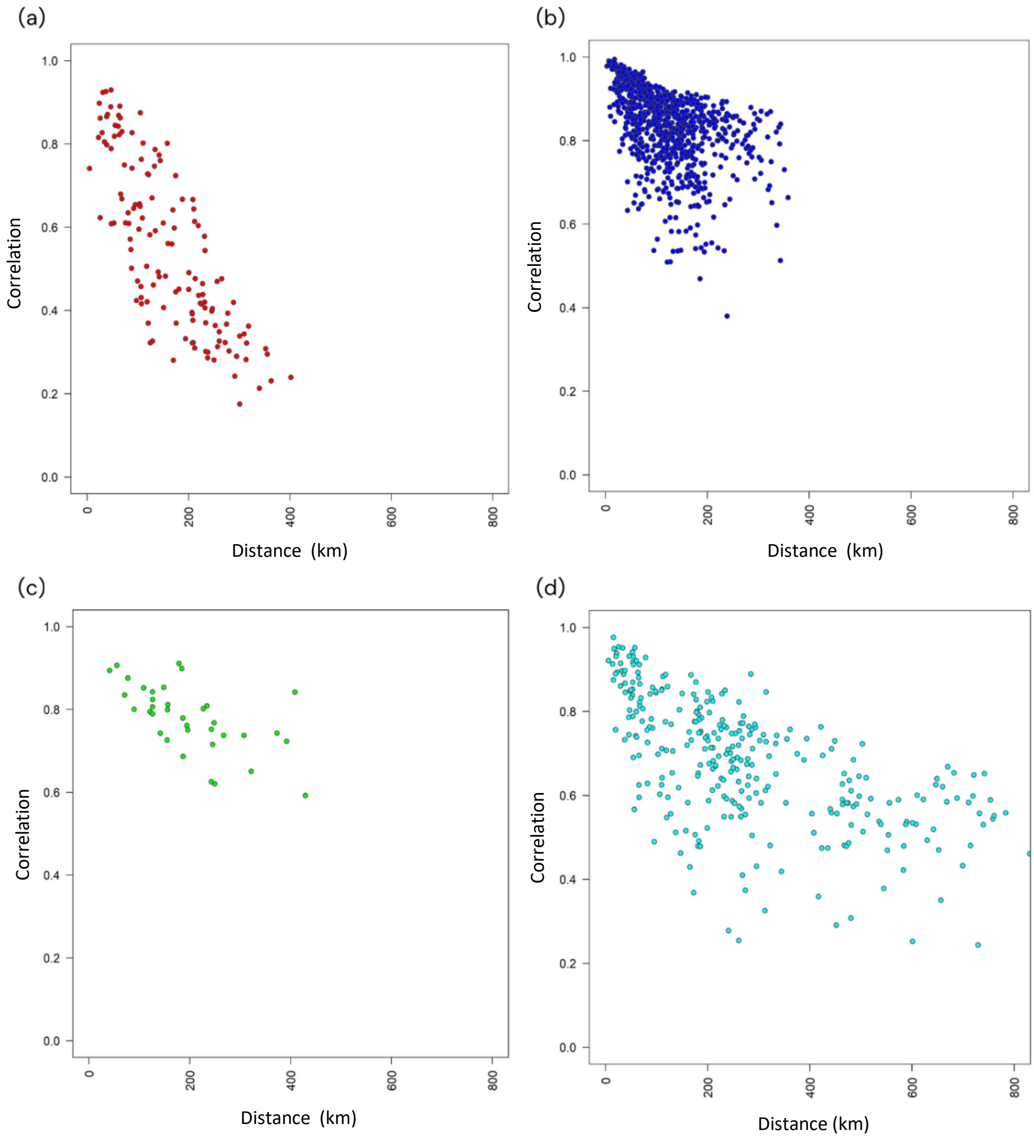



**Fig. 5 (a-c).** Relationship between, (a) drainage area, (b) stream length, and (c) elevation mean with flow CV for 9 communities identified in Australia

By analyzing relationship of distance and correlation within community for the 9 communities identified, the use of the modularity optimization method for classification is also investigated. The linearity of distance and correlation for the 9 communities are compared as shown in Figure 6. Communities 1 (red), 12 (blue), 22 (green), 24 (cyan), 34 (pink), and 44 (brown) can all be observed to maintain comparatively stronger correlations as the distance grows (Figure 6(a), (b), (c), (d), and (g)). Given that very high correlations might facilitate linkages between stations that are spread across great distances, it is not unexpected that community 24 (cyan) is enormous and spans vast distances. Despite this, links are observed in communities 26 (yellow), 28 (purple), and 38 (orange) (Figure 6(e), (f), and (h)), which may be because of the close proximity. Therefore, geographical closeness and the river system may also be crucial considerations for classifying catchments. The count for stations in each community is substantially lower and the patterns are more sparse for the community 44 (brown) (Figure 6(i)). These communities can develop no matter how far apart they are. This appears to imply that the stations cover great distances due to the significant connections. Overall, the modularity optimization technique and its capacity for catchment classification according to connectedness as its foundation, without relying on links in streamflow but rather without previous knowledge of the catchment physics, has shown to be beneficial.



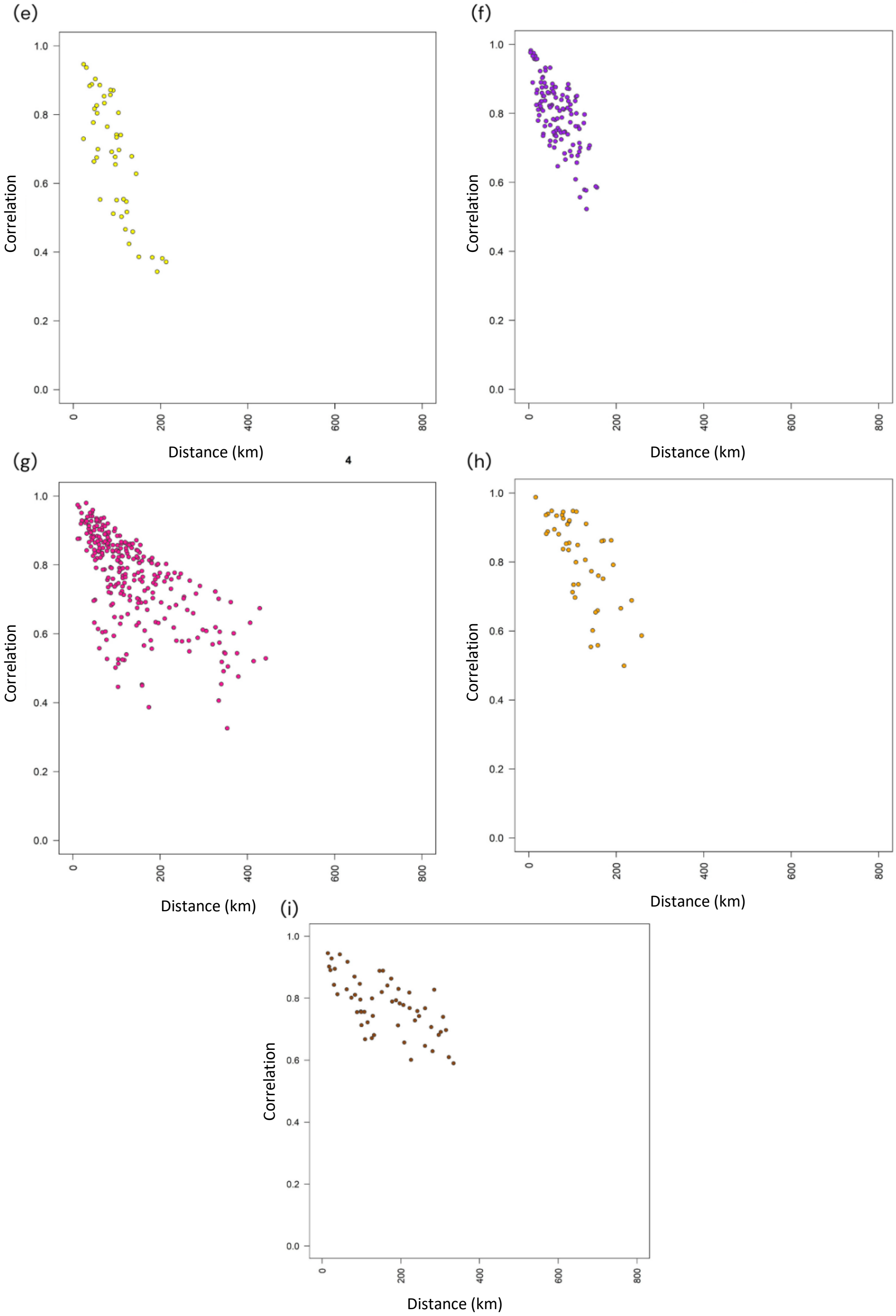



**Fig. 6 (a-i).** Relationship of distance and correlation for 9 communities identified in Australia, following to the color scheme in **Fig. 3(d)**

## 5. Conclusions

In general, the results of the classification that has been carried out on large areas covering this variety of climates have shown some points. Among them, a small count of communities consist of big number of streamflow stations. For example, nine of communities identified have combined to cover almost 77% of the catchment area. In addition, a large count of communities consists of a few stations in them, which means that almost 68% of the total communities has only at most two stations in it, representing only about 20% of the streamflow stations in entire Australia. The relationship between correlation and distance as well as the identified community investigation of some important catchments and flow properties such as the drainage area, elevation, stream length, flow CV and flow mean have provided some insightful discussions. The results also show that the same correlation threshold value T=0.8 is ideal to observe the monthly streamflow data for catchment classification in Australia, with other application of community structure methods, which may be discuss in near future to assess more in depth of the usefulness and the suitability of method of community structure for catchment classification, as a whole.

This is the first attempt to use the concept of modularity optimization to the classification of catchments across Australia. The evaluation of modularity optimization methods, particularly for Australian regions, has shed some light to justify the suitability of methods for classifying catchments. The results are encouraging, demonstrating the applicability and effectiveness of modularity optimization methods and community structure methods in catchment classification. However, as the modularity optimization method or any other community structure methods that implemented modularity metrices should be implemented with caution. This is because the modularity function has limitations and disadvantages in terms of scale resolution limit problem (variability of network sizes) which is common in catchment network. The improvement can be made by substituting the modularity function by modularity density function (or D value) to combat the limitation of scale resolution limit problem occurred when one applied the modularity-based classification approach. Therefore, improving the existing community structure methods especially the ones that are modularity-based methods are needed to ensure that the community detection techniques using the concept of community structure stay relevant and applicable towards general hydrology framework development.

**Acknowledgement**
This work was supported by UMS, Research Grant No. SLB2105 under Research Management Centre (RMC), Universiti Malaysia Sabah, Malaysia.

**References**

[1] Sivakumar, Bellie, Vijay P. Singh, Ronny Berndtsson, and Shakera K. Khan. "Catchment classification framework in hydrology: challenges and directions." *Journal of Hydrologic Engineering* 20, no. 1 (2015): A4014002. https://doi.org/10.1061/(asce)he.1943-5584.0000837

[2] Li, Qingsong, Huishan Gao, Shan Chai, Zongyu He, and Guangxing Ji. "Attribution analysis of different driving forces on vegetation and streamflow variation in the Jialing River Basin, China." *Open Geosciences* 14, no. 1 (2022): 1513-1523. https://doi.org/10.1515/geo-2022-0443

[3] Du, Shuping, Shanhu Jiang, Liliang Ren, Shanshui Yuan, Xiaoli Yang, Yi Liu, Xinglong Gong, and Chong-Yu Xu. "Control of climate and physiography on runoff response behavior through use of catchment classification and machine learning." *Science of The Total Environment* 899 (2023): 166422. https://doi.org/10.1016/j.scitotenv.2023.166422

[4] Olden, Julian D., Mark J. Kennard, and Bradley J. Pusey. "A framework for hydrologic classification with a review of methodologies and applications in ecohydrology." *Ecohydrology* 5, no. 4 (2012): 503-518. https://doi.org/10.1002/eco.251

[5] Sivakumar, B., and F. M. Woldemeskel. "Complex networks for streamflow dynamics." *Hydrology and Earth System Sciences* 18, no. 11 (2014): 4565-4578. https://doi.org/10.5194/hessd-11-7255-2

[6] Fang, Koren, Bellie Sivakumar, and Fitsum M. Woldemeskel. "Complex networks, community structure, and catchment classification in a large-scale river basin." *Journal of hydrology* 545 (2017): 478-493. https://doi.org/10.1016/j.jhydrol.2016.11.056

[7] Tumiran, Siti Aisyah, and Bellie Sivakumar. "Catchment classification using community structure concept: application to two large regions." *Stochastic Environmental Research and Risk Assessment* 35 (2021): 561-578. https://doi.org/10.1007/s00477-020-01936-4

[8] Xu, Yiran, Fan Lu, Kui Zhu, Xinyi Song, and Yanyu Dai. "Exploring the clustering property and network structure of a large-scale basin's precipitation network: a complex network approach." *Water* 12, no. 6 (2020): 1739. https://doi.org/10.3390/w12061739

[9] Halverson, M. J., and S. W. Fleming. "Complex network theory, streamflow, and hydrometric monitoring system design." *Hydrology and Earth System Sciences* 19, no. 7 (2015): 3301-3318. https://doi.org/10.5194/hess-19-3301-2015

[10] Girvan, Michelle, and Mark EJ Newman. "Community structure in social and biological networks." *Proceedings of the national academy of sciences* 99, no. 12 (2002): 7821-7826. https://doi.org/10.1073/pnas.122653799

[11] Tumiran, Siti Aisyah, and Bellie Sivakumar. "Catchment classification using community structure concept: application to two large regions." *Stochastic Environmental Research and Risk Assessment* 35 (2021): 561-578. https://doi.org/10.1007/s00477-020-01936-4.

[12] Tumiran, Siti Aisyah, and Bellie Sivakumar. "Community structure concept for catchment classification: a modularity density-based edge betweenness (MDEB) method." *Ecological Indicators* 124 (2021): 107346. https://doi.org/10.1016/j.ecolind.2021.107346

[13] Newman, Mark EJ. "Detecting community structure in networks." *The European physical journal B* 38 (2004): 321-330. https://doi.org/10.1140/epjb/e2004-00124-y

[14] Clauset, Aaron, Mark EJ Newman, and Cristopher Moore. "Finding community structure in very large networks." *Physical review E* 70, no. 6 (2004): 066111. https://doi.org/10.1103/physreve.70.066111

[15] Newman, Mark EJ. "Modularity and community structure in networks." *Proceedings of the national academy of sciences* 103, no. 23 (2006): 8577-8582. https://doi.org/10.1073/pnas.0601602103

[16] Blondel, Vincent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008, no. 10 (2008): P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

[17] Newman, Mark EJ. "The structure and function of complex networks." *SIAM review* 45, no. 2 (2003): 167-256. https://doi.org/10.1137/S003614450342480

[18] Li, Wenye, and Dale Schuurmans. "Modular community detection in networks." In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, vol. 22, no. 1, p. 1366. 2011. https://doi.org/10.5591/978-1-57735-516-8/IJCAI11-231

[19] Newman, Mark EJ. "Modularity and community structure in networks." *Proceedings of the national academy of sciences* 103, no. 23 (2006): 8577-8582. https://doi.org/10.1073/pnas.0601602103

[20] Zhang, Xiaoyong Sophie, Gnanathikkam E. Amirthanathan, Mohammed A. Bari, Richard M. Laugesen, Daehyok Shin, David M. Kent, Andrew M. MacDonald, Margot E. Turner, and Narendra K. Tuteja. "How streamflow has changed across Australia since the 1950s: evidence from the network of hydrologic reference stations." *Hydrology and Earth System Sciences* 20, no. 9 (2016): 3947-3965. https://doi.org/10.5194/hess-2015-464