

Supplementary materials.

1). Klang

Table 1. Clusters members (months) resulted from the method HACA on its own (traditional HACA) and HACA with topological properties (improved HACA) in Klang station for year 2000 until 2015.

Traditional HACA		Improved HACA		
Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3
Jan-00 until Dec-00		Jan-00 until Dec-00		
Jan-01 until Dec-01		Jan-01 until Dec-01		
Jan-02		Jan-02 until Jul-02		
Apr-02 until Dec-02		Sep-02 until Dec-02		
Jan-03 until Dec-03		Jan-03 until Dec-03		
Jan-04 until Dec-04		Jan-04 until Dec-04		
Jan-05 until Jul-05		Jan-05		
Sep-05 until Dec-05	Feb-02	Mar-05 until Jul-05		
Jan-06 until Dec-06	Mar-02	Sep-05 until Dec-05	Aug-02	
Jan-07 until Dec-07	Aug-05	Jan-06 until Dec-06	Feb-05	Aug-05
Jan-08 until Dec-08	Jun-13	Jan-07 until Dec-07	Sep-15	Jun-13
Jan-09 until Dec-09	Mar-14	Jan-08 until Dec-08	Oct-15	Mar-14
Jan-10 until Dec-10	Sep-15	Jan-09 until Dec-09		
Jan-11 until Dec-11	Oct-15	Jan-10 until Dec-10		
Jan-12 until Dec-12		Jan-11 until Dec-11		
Jan-13 until May-13		Jan-12 until Dec-12		
Jul-13 until Dec-13		Jan-13 until May-13		
Jan-14 and Feb-14		Jul-13 until Dec-13		
Apr-14 until Dec-14		Jan-14 and Feb-14		
Jan-15 until Aug-15		Apr-14 until Dec-14		
Nov-15 and Dec-15		Jan-15 until Aug-15		
		Nov-15 and Dec-15		

Based on Table 1, two clusters formed using traditional HACA approach with cluster 1 and cluster 2 consists months with and without haze respectively. In improved HACA approach, three clusters formed with months without haze (cluster 1), moderate and severe haze (cluster 2 and cluster 3 respectively). In 2002, unhealthy air quality was reported due to peat swamp and forest fires during dry periods between February and March 2002 and from mid-July to September 2002 [1]. Cluster 2 in traditional HACA approach is able to identify February and March 2002 as the months with haze and this is not achieved by HACA with topological properties. However, our improved HACA method is able to identify August 2002 and February 2005 as the months with haze. The clusters formed by the improved HACA approach are also well separated between moderate and severe haze. Note that, for both approaches, July 2014 is not included as a cluster consisting of month with haze and this differs from the main result in the manuscript (see Figure 9). This is due to the weakness of HACA, which only cluster an object once and that object would not be considered again. Thus, the possible previous misclassifications are unable to be corrected by HACA [2].

2). Petaling Jaya

Table 2. Clusters members (months) resulted from the method HACA on its own (traditional HACA) and HACA with topological properties (improved HACA) in Petaling Jaya station for year 2000 until 2015.

Traditional HACA		Improved HACA	
Cluster 1	Cluster 2	Cluster 1	Cluster 2
Jan-00 until Dec-00			
Jan-01 until Dec-01			
Jan-02 until Jul-02			
Sep-02			
Nov-02 and Dec-02		Jan-00 until Dec-00	
Jan-03 until Mar-03		Jan-01 until Dec-01	
Jun-03 until Dec-03	Aug-02	Jan-02 until Dec-02	
Jan-04 until Apr-04	Oct-02	Jan-03 until Dec-03	
Jul-04	Apr-03	Jan-04 until Dec-04	
Oct-04 until Dec-04	May-03	Jan-05 until Jul-05	
Jan-05 until Jul-05	May-04	Sep-05 until Dec-05	
Oct-05 until Dec-05	Jun-04	Jan-06 until Dec-06	
Jan-06 until Sep-06	Aug-04	Jan-07 until Dec-07	Aug-05
Nov-06 and Dec-06	Sep-04	Jan-08 until Dec-08	Jun-13
Jan-07 until Dec-07	Aug-05	Jan-09 until Dec-09	Mar-14
Jan-08 until Dec-08	Sep-05	Jan-10 until Dec-10	Sep-15
Jan-09 until Dec-09	Oct-06	Jan-11 until Dec-11	Oct-15
Jan-10 until Dec-10	Jun-13	Jan-12 until Dec-12	
Jan-11 until Dec-11	Mar-14	Jan-13 until May-13	
Jan-12 until Dec-12	Sep-15	Jul-13 until Dec-13	
Jan-13 until May-13	Oct-15	Jan-14 and Feb-14	
Jul-13 until Dec-13		Apr-14 until Dec-14	
Jan-14 and Feb-14		Jan-15 until Aug-15	
Apr-14 until Dec-14		Nov-15 and Dec-15	
Jan-15 until Aug-15			
Nov-15 and Dec-15			

As shown in Table 2, two clusters are identified in both methods with cluster 1 represents months without haze and cluster 2 represents months with severe haze. The months with severe haze are mixed up with months without haze in cluster 2 resulted from the traditional HACA approach. An improvement of the cluster members is shown in improved HACA approach with cluster 2 consists only months with severe haze. Similarly, at Shah Alam station (Table 3), the clusters formed are well separated, proving the consistencies of our improved HACA approach.

3). Shah Alam

Table 3. Clusters members (months) resulted from the method HACA on its own (traditional HACA) and HACA with topological properties (improved HACA) in Shah Alam station for year 2000 until 2015.

Traditional HACA		Improved HACA	
Cluster 1	Cluster 2	Cluster 1	Cluster 2
Jan-00 until Dec-00		Jan-00 until Dec-00	
Jan-01 until Dec-01		Jan-01 until Dec-01	
Jan-02 and Feb-02		Jan-02 until Dec-02	
Apr-02 and May-02		Jan-03 until Dec-03	
Jul-02		Jan-04 until Dec-04	
Sep-02		Jan-05 until Jul-05	
Nov-02 and Dec-02	Mar-02	Sep-05 until Dec-05	
Jan-03 until Dec-03	Jun-02	Jan-06 until Dec-06	Aug-05
Jan-04 until May-04	Aug-02	Jan-07 until Dec-07	Jun-13
Jul-04	Oct-02	Jan-08 until Dec-08	Mar-14
Oct-04 until Dec-04	Jun-04	Jan-09 until Dec-09	Sep-15
Jan-05	Aug-04	Jan-10 until Dec-10	Oct-15
Apr-05 until Jul-05	Sep-04	Jan-11 until Dec-11	
Sep-05 until Dec-05	Feb-05	Jan-12 until Dec-12	
Jan-06 until Sep-06	Mar-05	Jan-13 until May-13	
Nov-06 and Dec-06	Aug-05	Jul-13 until Dec-13	
Jan-07 until Dec-07	Oct-06	Jan-14 and Feb-14	
Jan-08 until Dec-08	Jun-13	Apr-14 until Dec-14	
Jan-09 until Dec-09	Mar-14	Jan-15 until Aug-15	
Jan-10 until Dec-10	Sep-15	Nov-15 and Dec-15	
Jan-11 until Dec-11	Oct-15		
Jan-12 until Dec-12			
Jan-13 until May-13			
Jul-13 until Dec-13			
Jan-14 and Feb-14			
Apr-14 until Dec-14			
Jan-15 until Aug-15			
Nov-15 and Dec-15			

Reference

1. Enviro Knowledge Centre. Malaysia Environmental Quality Report 2002. Available online: <https://enviro.doe.gov.my/> (accessed on 25 November 2018).
2. Saxena, A.; Prasad, M.; Gupta, A.; Bharill, N.; Patel, O.P.; Tiwari, A.; Er, M.J.; Ding, W.; Lin, C.T. A review of clustering techniques and developments. *Neurocomputing* **2017**, *167*, 664–681.